

8-7-2020

## Algorithm of Double Level Visual Modeling Based on Data of Pesticide Residues

Ruijun Liu

*1. Beijing Technology and Business University, Beijing 100048, China;;3. Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, China;*

Chen Yi

*1. Beijing Technology and Business University, Beijing 100048, China;;3. Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, China;*

Xu Yang

*2. Peking University, Beijing 100871, China;;*

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

---

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

---

# Algorithm of Double Level Visual Modeling Based on Data of Pesticide Residues

## Abstract

**Abstract:** Based on the study of pesticide residues detection result data, which has the characteristics of geographical level and the level of classification of fruits and vegetables, in node links and radial loop algorithm visualization model, the double level structure, radiation chain graph *CircleLink*, was designed for analysis of pesticide residues data in visual modeling algorithm. The visual modeling algorithm was applied to the actual pesticide residues detection results data for the simulation analysis and conclusions. Analysis of the characteristics of pesticide collection, distribution of agricultural products and supermarket pesticide residues of pesticide residues was applied.

## Keywords

pesticide residues, node-link, radial ring, CircleLink

## Recommended Citation

Liu Ruijun, Chen Yi, Xu Yang. Algorithm of Double Level Visual Modeling Based on Data of Pesticide Residues[J]. Journal of System Simulation, 2015, 27(9): 2144-2149.

# 基于农残数据的双层次可视化建模算法

刘瑞军<sup>1,3</sup>, 陈谊<sup>1,3</sup>, 徐扬<sup>2</sup>

(1.北京工商大学, 北京 100048; 2.北京大学, 北京 100871;  
3.北京工商大学食品安全大数据技术北京市重点实验室, 北京 100048)

**摘要:** 基于研究农药残留量检测结果数据中具有地理层次和蔬果分类层次的特点, 在节点-链接和放射环可视化算法模型的基础上, 设计了分析农残数据中双层次结构的可视化建模算法放射链图(CircleLink)。该方法可视化展示农残信息的分布, 为相关信息的可视分析提供支持。将可视化建模算法应用于实际的农药残留量检测结果数据当中, 进行建模与仿真分析, 并得出结论。该方法有效的分析了农药采集的特征, 农药残留的分布和超市中农产品农药残留的情况等, 为农产品中农药残留量的检验检疫提供技术支持。

**关键词:** 农药残留; 节点-链接; 放射环; 放射链图

中图分类号: TP391.4

文献标识码: A

文章编号: 1004-731X (2015) 09-2144-06

## Algorithm of Double Level Visual Modeling Based on Data of Pesticide Residues

Liu Ruijun<sup>1,3</sup>, Chen Yi<sup>1,3</sup>, Xu Yang<sup>2</sup>

(1. Beijing Technology and Business University, Beijing 100048, China; 2. Peking University, Beijing 100871, China;  
3. Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, China)

**Abstract:** Based on the study of pesticide residues detection result data, which has the characteristics of geographical level and the level of classification of fruits and vegetables, in node links and radial loop algorithm visualization model, the double level structure, radiation chain graph *CircleLink*, was designed for analysis of pesticide residues data in visual modeling algorithm. The visual modeling algorithm was applied to the actual pesticide residues detection results data for the simulation analysis and conclusions. Analysis of the characteristics of pesticide collection, distribution of agricultural products and supermarket pesticide residues of pesticide residues was applied.

**Keywords:** pesticide residues; node-link; radial ring; CircleLink

## 引言

大数据是近年来计算机领域的一个热点话题。随着计算机的普及, 应用面的拓展, 每日都有海量的数据产生。但是, 单纯地存、取和使用数据, 是不能将数据潜在的价值所挖掘出来的。通过数据挖

掘和分析的算法, 可以从海量的数据中发现一定的规律, 帮助人们对数据的结果做出决策。

使用数据挖掘分析和算法, 往往需要有针对性的专业知识。当且仅当具有对数据的规律、背景和表现有一定的理解, 并且对算法的原理有所认识的情况下, 才能有效地使用数据挖掘的方式对数据进行深层的分析。而可视化, 通过将抽象的数据, 以易于被人们理解的具象的图形图像表现出来, 让即使不具有专业知识的人, 通过简单的解释说明和交互手段, 也可以对数据进行有效的分析。



收稿日期: 2015-06-14 修回日期: 2015-07-20;  
作者简介: 刘瑞军(1982-), 男, 河北唐山, 讲师, 研究方向为虚拟现实、可视化; 陈谊(1963-), 女, 北京, 教授, 硕导, 研究方向为信息可视化与可视分析; 徐扬(通讯作者 1981-), 男, 湖北武汉, 副教授, 研究方向为知识管理、信息生命周期管理。

<http://www.china-simulation.com>

• 2144 •

层次数据可视分析,是发展较为全面的可视分析方法之一。层次数据的可视化算法,主要有节点-链接算法、空间填充算法和混合算法 3 种。其中,节点-链接算法可以清晰地表现数据层次结构,显示大量的节点数据;空间填充方法具有很高的空间利用率,能够突出节点权值的大小;混合算法,由于结合的算法不同,会具有不一样的特性。

## 1 相关研究

农药残留量检测结果数据,以下简称农残数据,用于反映民生话题中的食品安全问题。方便、快捷和有效地分析农残数据,发现数据背后隐藏的规律,保障蔬菜水果中农药使用符合规范具有重要的意义。

数据采集和分析的过程如图 1,首先由各地的检测人员制定采集计划,然后根据计划到其所属的超市或农贸市场采集检测样品,再将样品带回检科院进行检测。根据不同的农产品,检测可以按不同的检测方法,检测不同的农药。根据检测的结果,按照不同国家的检测标准,分析检测结果;按照农药的分类,分析高毒和剧毒的农药使用情况等。

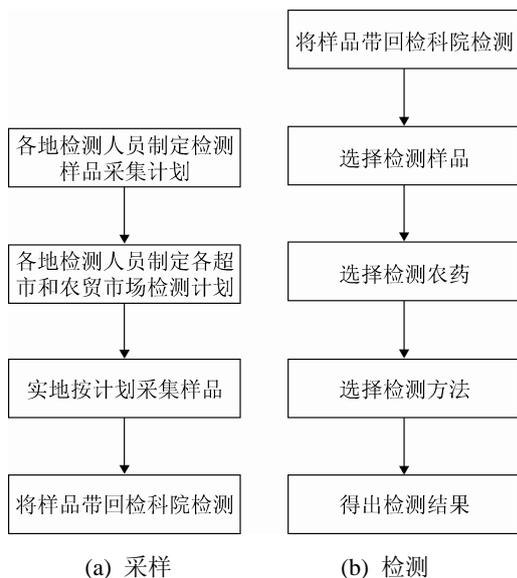


图 1 农药检测过程

在对农残数据分析的过程中,发现其具有 3 个层次结构:1) 地理层次结构;2) 农产品种类层次

结构;3) 农药种类层次结构。在使用传统的层次数据可视化算法时,只能针对其中的一种层次结构进行分析。如果将多个层次结构重新组合为一颗具有多层次结构的树,如图 2,在使用单种可视化算法的时候,必须使用颜色标注或者其他方式才可以区分每一种层次结构。

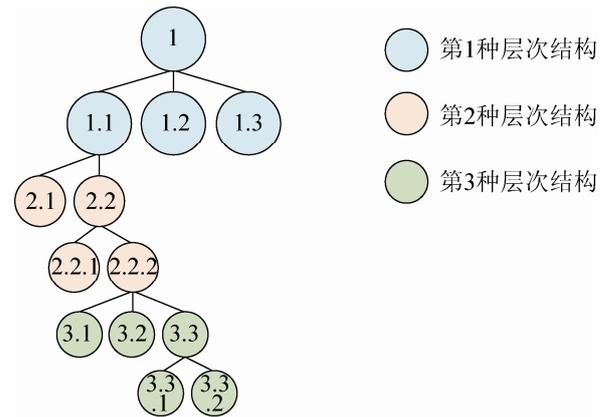


图 2 一棵树中多层次结构

针对农残数据中具有多个层次结构的特点,以节点-链接算法和放射环算法为基础,提出了基于农残数据的可视化建模算法放射链图(CircleLink),实现具有多个层次结构的数据的可视分析。该算法具有以下 3 个特点:1) 有较高的空间利用率;2) 能够清晰地表达数据的层次结构;3) 可以表现节点的权值大小。

使用节点-链接的方式对层次数据进行可视化显示,可以清晰地显示数据的层次结构,使用径向的方式进行显示,更加凸显层次结构。

传统的 Radial tree<sup>[1]</sup>将圆心定义为根节点,根据距离圆心的长短与树的层次深浅对应,每一层的子节点在其父节点包含的弧度内。Hyperbolic Tree<sup>[2]</sup>通过将 Radial tree 的直线链接优化成曲线链接,并使用“焦点+上下文”的交互技术,用户可以实现拖拽,收缩和扩展在。而 Hyperbolic Tree<sup>[3]</sup>不仅可以在 2D 显示中使用,在 3D 视图中也可应用该技术。TVBRT<sup>[4]</sup>将传统的 Radial Tree 加以改进,侧重于对数据进行层次分析,适合于展示对具体度量值更为关心的数据集。

通过环状显示,放射环具有以下几个特点:保持数据的整体视野;层次结构显示清晰,并易于理解;允许对分支目录进行展开,实现上下文同时分析;使用空间填充,具有良好的空间利用率;易于对比全局和细节数据的关系和差别。

在布局的过程中,可以在颜色和结构上对放射环进行优化, Radial Clustergrams<sup>[5]</sup>由于层次过深时,叶子节点的形状过于细长, TimeRadarTrees<sup>[6]</sup>由一个主要的放射环显示整体树形结构,通过在圆环上或者圆环外生成一个或多个规模较小的放射环实现对节点信息的展开显示。而在 3D 维度, 3D sunburst wheel<sup>[7]</sup>将 3 维显示模式通过横切,纵切,间隔均分这三种形式,可通过设置不同的切割方式生成自定义的符合数据特点的“wheel”状视图。Stacked 3D-Sunbursts<sup>[8]</sup>对层次结构中,分析具有多维数据属性的数据类型。

## 2 CIRCLELINK 算法

### 2.1 数据建模

对于农残数据,本文选择地理层次和蔬果分类层次进行建模。分析农残数据的按地域分布的使用情况,分析农药残留量在不同的蔬菜水果之中的使用情况。1) 地理层次选择的维度从省、市、区再到具体的超市和农贸市场; 2) 样品的分类层次为农产品分类、蔬菜或水果中种类分类和蔬菜水果具体名称; 3) 叶子节点取具体农产品检出农药的频次。

### 2.2 可视化仿真算法

#### 2.2.1 径向布局

对于单个层次的节点数据,采用节点-链接算法布局时,为了突显数据的层次结构,将采用圆心向外径向的方式进行节点布局,如图 3。将根节点布局位置作为圆心,每个相邻的层次布局之间,距离圆心的距离累加  $S_k$  的长度。对于处于同一层次  $i$  中的兄弟节点,其对应的数据结构的深度为  $h$ , 则其圆环之间的距离为  $(h-1)S_k$ 。

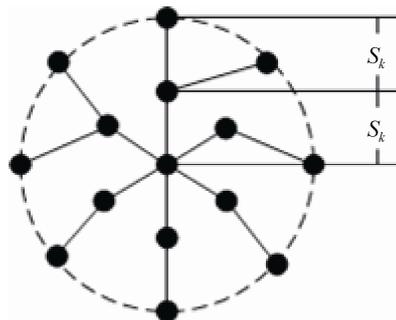


图 3 径向显示

对于处于  $h$  深度的节点,通过将兄弟节点按存储顺序,从圆心正上方  $(h-1)S_k$  的圆环上开始顺时针布局,从视觉效果上明显区分每一个层次,每一个环上的点都属于树中同一层次的节点。

若使用直线连接,则会存在如图 4(a)中红线的交叉现象。为了避免由于连线之间的交叉导致连线凌乱,影响数据层次的显示效果。对圆心、链接的子节点和父节点不在一条直线上,选择使用 4 阶的贝塞尔曲线进行连接。将父节点  $P_0$  作为连线的起点,将子节点  $P_3$  作为连线的中点。则贝塞尔曲线仍需要 2 个辅助点  $P_1, P_2$ , 其取点方式为:若父节点  $P_0$  与圆心的连线与  $P_3$  所处的圆环交点为  $P_p$ , 则  $P_1, P_2$  为  $P_p$  与  $P_3$  之间弧度的三等分的位置点。再由贝塞尔曲线公式(1),进行曲线计算,生成图 4b 中的曲线。

$$B(t) = P_0(1-t)^3 + 3P_1(1-t)^2 + 3P_2t^2(1-t) + P_3t^3 \quad (1)$$

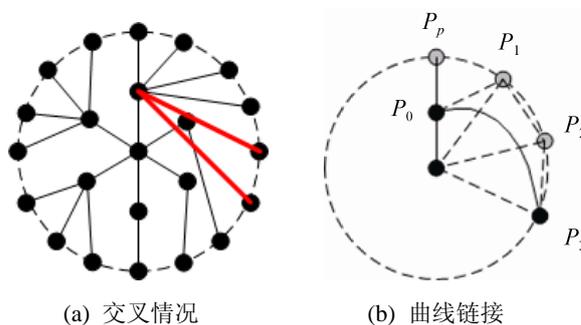


图 4 曲线链接方式

使用曲线优化之后的节点链接线条,可以有效避免因节点过多引起的连线交叉的情况,使节点链接布局可以保持凸显层次关系的特性。

### 2.2.2 放射环

对于图 5(a)的树形结果, 根据图 5(b)中圆的面积进行放射环布局, 布局结果如图 5(c)所示。

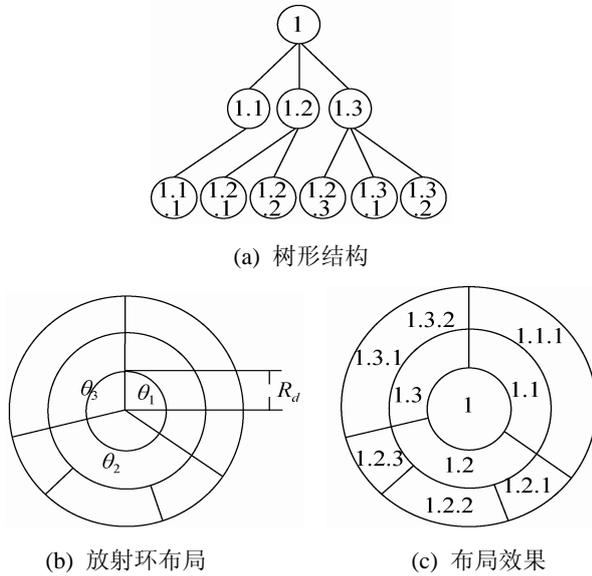


图 5 放射环布局

其中放射环布局过程为: 若放射环中用于布局的树形结构深度为  $h$ , 放射环可布局的圆形空间半径为  $R_f$ , 将圆形空间以圆心为中点按深度进行分割, 每个深度的圆环离圆心的距离为  $R_d = R_f / h$ , 则相邻层次的数据的布局之间距离  $R_d$  的长度。

子节点的分割方式按照其父节点所占的圆环的角度进行映射分割, 分割方式为按角度和权值进行等比映射, 如公式(2)。假设其父节点有  $k$  个子节点, 其对应的权值为  $\{v_1, v_2, v_3, \dots, v_k\}$ , 则其对应的角度记为  $\{\theta_1, \theta_2, \theta_3, \dots, \theta_k\}$ , 按公式(2)计算每一个节点  $i$  对应的角度值。对于属于同一个父节点下的兄弟节点, 兄弟节点深度为  $h_b$ , 若其父节点所占有的角度范围为  $[\theta_s, \theta_e]$ , 则在该角度范围内, 兄弟节点之间按照角度分割  $[(h-1)R_d, hR_d]$  的圆环面积。

$$\theta_i / v_i = \sum_{u=1}^k \theta_u / \sum_{u=1}^k v_u \quad (2)$$

### 2.2.3 点链接放射环

若数据具有 2 个层次结构 A 和 B, 其中具有

根节点的属于层次结构 A, 如图 6 所示。则先以层次关系 A 作为顶层结构, 当树形层次结构伸展至 A 的叶子节点 A.1, A.2 和 A.3 时, 用 A 的叶子节点作为层次关系 B 的父节点, 其中 A.1=B1, A.2=B2 和 A.3=B3。对于层次关系 A 使用节点-链接方法显示, 层次关系 B 使用放射环算法布局。

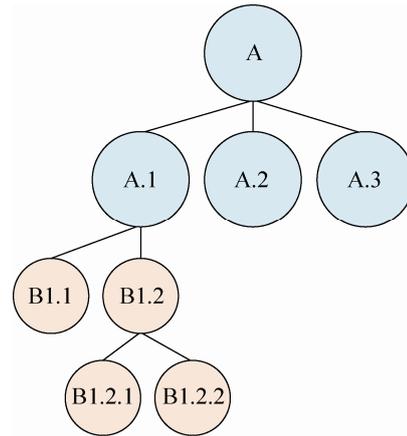


图 6 双层次结构数据

当对于层次关系 B 使用放射环布局时, 首先计算放射环中心点距离其父节点圆环的直径距离  $R_c$ 。若布局的兄弟节点中, 权值最大的节点为  $Q_m$ , 按照公式(2)中权值进行角度映射, 可以计算出其对应的可分割的角度为  $\theta_m$ 。若需要布局的放射环可布局的圆形半径为  $R_m$ , 则  $R_c = R_m + (h-1)S_k$ , 其中  $h$  为该放射环的父节点在整个布局的树形结构中所处的深度。用三角函数如公式(3)计算  $R_m$ 。确保放射环与其父节点和父节点的兄弟节点不想交, 并留有连线的空间, 避免连线过于混乱。

$$R_m = R_c \times \sin \theta_m \quad (3)$$

若兄弟节点之间对应的角度为  $\{\theta_1, \theta_2, \theta_3, \dots, \theta_k\}$ , 则其节点布局放射环对应的圆形半径为  $\{R_1, R_2, R_3, \dots, R_k\}$ , 其中每一个  $R_i$  根据公式(4)进行等比计算。

$$R_i = R_m \times \theta_i / \theta_m \quad (4)$$

根据半径为  $R_c$  圆环, 从第 1 个节点的所代表的放射环开始布局放射环的圆心。若第一个节点可

分割的角度为  $\theta_1$ ，则以  $\theta_1$  的角平分线与水平线垂直的方式，确定其布局圆心所处的位置  $O$ ，以  $O$  为圆心布局第一个半径为  $R_1$  的放射环。

当确定了第一个放射环布局的位置后，按角度分割的方式，沿顺时针顺序布局所有兄弟节点的放射环圆形，如图 7。

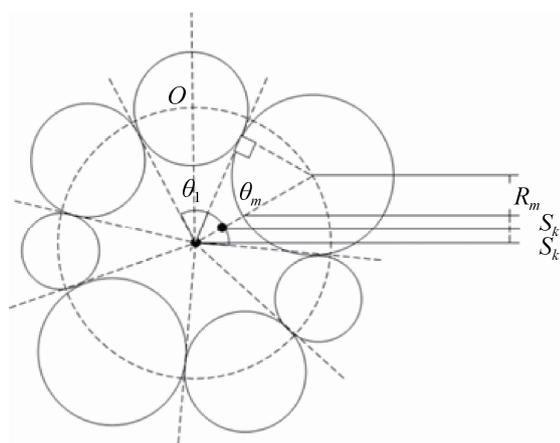


图 7 第二层次关系放射环布局

### 3 实际数据仿真应用

对于图 8 中的可视化仿真结果图，取的为真实的农残数据。对数据地理层次结构的建模为：省、城市、地、超市；对于农产品的层次结构建模为：蔬菜或水果分类、蔬菜或水果中的种类、蔬菜水果的具体名称。放射环面积对应的权值为农药残留量的检出频次。

图 8 中两种层次关系衔接的节点为超市中蔬菜和水果农药残留情况，其中放射环中颜色代表不同的层次层级；在同一个父节点下的兄弟节点填充放射环时，按节点的顺序使用渐进色进行填充。

通过图 8 可以清晰地区分 2 种层次结构的情况，通过节点链接情况可以看出来，该在 5 个城市中都对该超市连锁店进行了样品采集，说明可视化的数据具有一定的代表性。从放射环的面积可以看出来，检出农药残留量的频次在每个超市中的情况区别并不大，说明检测样品的采集较为平均，但是对于某些超市的农药检出频次还是比较高的，例如 11 号超市，需要对 11 号超市的进货来源加强监控。

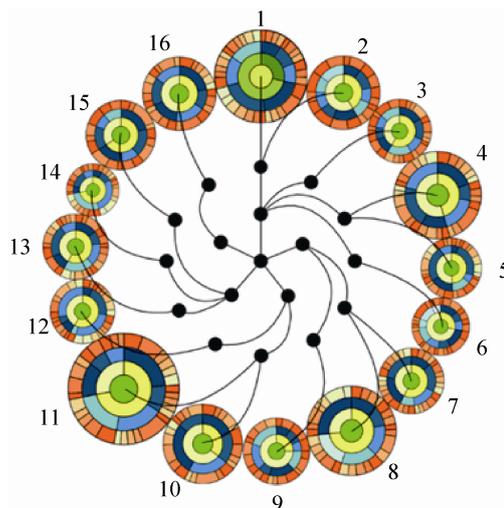


图 8 某省 5 市 12 个地区 16 个超市蔬果抽样结果可视化仿真分析图

再分析放射环中蓝色的层次结构的可视化仿真效果，蔬菜水果分类的品种数量不一致，说明采样的种类为随机采样，而非均衡采样，所以每个超市采样的蔬果样品都可能是不一样的。通过对比橙色层次结构，放射环中叶子节点的面积可以发现，某些蔬菜水果的农药检出量较大，占据了放射环中最外环较大的面积比例。针对检出量大的数据，可以反向追踪其数据的来源。

对比 1 号超市和 4 号超市的数据，两者在检出频次上相差不大。但实际上 1 号超市的样品的种类和数量都是多于 4 号超市的，两者对比之下，1 号超市的农产品平均检出的农药频次是较低的，证明 1 号超市的农产品农药残留情况要低于 4 号超市。

### 4 结论

本文针对农残数据中，存在地理、蔬果分类和农药分类的层次结构特点，进行数据建模并提出了 Circlelink 的仿真可视化算法。Circlelink 主要采用节点-链接和放射环结合的思想，引入了曲线链接的形式进行链接线条优化进行仿真可视化算法设计。通过 Circlelink，使节点链接的形式能够在保持数据层次结构的同时突出节点的权值大小；清晰地区分数据中不同的层次结构。

将可视化仿真技术应用于实际的农药残留量检测结果数据的分析之中,可以有效地帮助检察院分析农药使用情况的层次分级和使用情况,支持后续对不同层次类别的数据源进行有效区别监控。

对于农药残留量检测结果数据,本文仿真分析针对的为数据中地理层次、蔬果层次和农药层次的层次结构,农药检出量的频次情况。农残数据中还存在许多维度的属性值,因此,对于数据中其他维度的属性,例如采样时间、检测时间的的时间维度属性,中国 MRL 检测标准、欧盟 MRL 检测、日本 MRL 检测标准等多维属性的建模分析。

而本文提出的 CircleLink 可视化仿真技术可以尝试应用于其他具有多层次结构的数据建模分析中,例如应用分析具有地理和物品分类的网购数据。

### 参考文献:

- [1] Burch M, Konevtsova N, Heinrich J, *et al.* Evaluation of traditional orthogonal and radial tree diagrams by an eye tracking study [J]. Visualization and Computer Graphics (S1077-2626), 2011, 17(12): 2440-2448.
- [2] Lamping J, Rao R. The Hyperbolic Browser: A Focus + Context Technique For Visualizing Large Hierarchies[J]. Journal of visual languages and computing (S1045-926X), 1996, 7(1): 33-55.
- [3] Herman I, Melançon G, Marshall M S. Graph visualization and navigation in information visualization: a survey[J]. Visualization and Computer Graphics (S1077-2626), 2000, 6(1): 24-43.
- [4] 孙宁伟, 赵瑜, 刘勇, 等. TVBRT: 一种基于 Radial Tree 的具有度量属性的多变量时态数据可视化方法 [J]. 计算机科学, 2014, 41(6): 5-11.
- [5] Agrafiotis D K, Bandyopadhyay D, Farnum M. Radial clustergrams: visualizing the aggregate properties of hierarchical clusters [J]. Journal of chemical information and modeling (S1549-9596), 2007, 47(1): 69-75.
- [6] Burch M, Diehl S. Timeradartrees: Visualizing dynamic compound digraphs [C]// Computer Graphics Forum. Massachusetts, USA: Blackwell Publishing Ltd, 2008, 27(3): 823-830.
- [7] Schulz H J, Hadlak S, Schumann H. The design space of implicit hierarchy visualization: A survey [J]. Visualization and Computer Graphics, IEEE Transactions on (S1077-2626), 2011, 17(4): 393-411.
- [8] Schedl M, Knees P, Widmer G, *et al.* Browsing the web using stacked three-dimensional sunbursts to visualize term co-occurrences and multimedia content [C]// Proceedings of the 18th IEEE Visualization Conference. Los Alamitos, USA: IEEE Computer Society Press, 2007: 281-289.



## 《系统仿真学报》荣获“2014 中国最具国际影响力学术期刊”证书

由中国学术期刊(光盘版)电子杂志社与清华大学图书馆联合成立的中国学术文献国际评价研究中心,以美国汤森路透 Web of Science 收录的 1.2 万余种期刊为引文统计源,首次研制发布了 2012《中国学术期刊国际引证年报》(CAJ-IJCR 年报)。第一次给出了我国 5600 余种中外文学学术期刊总被引频次、影响因子、半衰期等各项国际引证指标,并采用了新的国际影响力综合评价指标 CI 对期刊排序,发布了“中国最具国际影响力学术期刊”(排序 TOP5%)和“中国国际影响力优秀学术期刊”(排序 TOP5-10%),在国内外学术界产生了较大反响。之后,2013 年版年报,将引文统计源期刊扩展到 1.44 万多种。目前,2014 版国际、国内年报与 TOP5% 和 TOP5-10% 期刊的遴选业已完成,《系统仿真学报》被列入“2014 中国最具国际影响力学术期刊”行列。

我学报连续 2 年被列入 TOP5% 国内一流的中国最具国际影响力学术期刊,走向世界,进入国际一流,指日可待!