

8-7-2020

Discontinuous Hierarchical Data Visualization Method Based on ThemeRiver Model

Yuangang Zhen

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China;

Chen Yi

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China;

Liu Ying

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China;

Ruijun Liu

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, 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.

Discontinuous Hierarchical Data Visualization Method Based on ThemeRiver Model

Abstract

Abstract: For traditional ThemeRiver model which has poor performance support to characteristics of discrete data, *a kind of visualization method based on improved ThemeRiver model was put forward, which was applied to discontinuous data. All data was put into several time slots and was counted respectively, and new data was added in the time slot which lack data. Gaussian model was adopted to fit curve.* After ordering layout, selecting themes' color and arranging the labels, a layout was proposed which has prediction function and can show the hierarchical characteristic. For the hierarchy, *the same color family was used to show the same hierarchical data. The solution was applied to pesticide residues data*, and the result show that it's helpful for user to monitor pesticide residues in a period of time and the method can provide a basis for forecasting warning of pesticide residue to related people.

Keywords

information visualization, hierarchical data, ThemeRiver, pesticide residues data, time series

Recommended Citation

Zhen Yuangang, Chen Yi, Liu Ying, Liu Ruijun. Discontinuous Hierarchical Data Visualization Method Based on ThemeRiver Model[J]. Journal of System Simulation, 2015, 27(10): 2460-2466.

一种基于 ThemeRiver 模型的非连续层次数据可视化方法

甄远刚, 陈谊, 刘莹, 刘瑞军

(北京工商大学食品安全大数据技术北京市重点实验室计算机与信息工程学院, 北京 100048)

摘要: 针对传统 ThemeRiver 模型对非连续数据特征表现支持较差的问题, 提出一种改进的非连续数据 ThemeRiver 可视化方法。对数据点划分时间段分别进行统计, 补充新的数据点, 构造数据点集, 利用高斯模型曲线拟合, 经过主题布局排序, 颜色选择, 标签分布布局成一种具有预测功能且能够展示层次特征的新型主题河流模型。在层次结构上, 利用相同颜色序列表示同层次数据。将该方案应用于农残数据中, 有助于对一段时间内的农残数据进行有效监测, 为农药残留预测预警提供依据。

关键词: 信息可视化; 层次数据; 主题河流; 农残数据; 时间序列

中图分类号: TP391.9

文献标识码: A

文章编号: 1004-731X (2015) 10-2460-08

Discontinuous Hierarchical Data Visualization Method Based on ThemeRiver Model

Zhen Yuangang, Chen Yi, Liu Ying, Liu Ruijun

(Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China)

Abstract: For traditional ThemeRiver model which has poor performance support to characteristics of discrete data, a kind of visualization method based on improved ThemeRiver model was put forward, which was applied to discontinuous data. All data was put into several time slots and was counted respectively, and new data was added in the time slot which lack data. Gaussian model was adopted to fit curve. After ordering layout, selecting themes' color and arranging the labels, a layout was proposed which has prediction function and can show the hierarchical characteristic. For the hierarchy, the same color family was used to show the same hierarchical data. The solution was applied to pesticide residues data, and the result show that it's helpful for user to monitor pesticide residues in a period of time and the method can provide a basis for forecasting warning of pesticide residue to related people.

Keywords: information visualization; hierarchical data; ThemeRiver; pesticide residues data; time series

引言

大规模层次数据可视化是现今的一大热点^[1-2], 通常情况下, 此类数据同时还具有时间属性, 如何在保持层次结构清晰展现的基础上, 将时变特征表现出来是现今的一大挑战。展示层次数据的可视化

方法主要有结点-链接法和空间填充法 2 种。结点-链接法主要通过线段和结点连接的方式展现层次结构特性, 优点是展现层次关系清晰明了, 但是空间浪费严重。此外, 当数量太大时, 由于屏幕空间不足, 十分容易出现结点重叠的现象。空间填充法的典型代表是 Treemap, Treemap 能够最大限度的利用空间, 同时能够通过面积显示数据的属性值, 但是在展现层次结构方面不如结点链接法直观。无论是结点链接法还是空间填充法, 都很难在时间序列上合理的布局展示。



收稿日期: 2015-06-14 修回日期: 2015-07-24;
基金项目: “十二五”国家科技支撑计划项目(2012BAD29B01-2); 国家科技基础性工作专项(2015FY111200);
作者简介: 甄远刚(1989-), 男, 河北, 硕士生, 研究向为信息可视化与可视分析; 陈谊(通讯作者 1963-), 女, 北京, 博士, 教授, 研究方向为信息可视化与可视分析。

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

• 2460 •

ThemeRiver 模型是一种典型的处理时序数据的可视化的结构样式, 能够将得到的大量时序数据集集合在一起, 按时间顺序构造一个类似“河流”的可视化显示方式, 通过河流的流动趋势和涌流(current)的坡度可以反映主题强度信息随时间的变化情况, 这种方式能够解决大规模数据的时序特性, 且无论主题如何变化, 其走向总是清晰可见, 一目了然。但主题河流能够展示的信息过于单一, 通常无法展现数据的层次特性, 对于细节的把握也不尽如人意。

随着社会的飞速发展, 各种社会生活问题也随之暴露出来。大数据时代, 我们可以通过现代化的手段, 如数据挖掘, 机器学习等, 将数据分析技术与先进的信息可视化技术^[3]相结合^[4], 帮助人们掌握并分析领域内相关数据。

民以食为天, 食品安全越来越受到国家和人民的重视而且食品安全问题受到广泛的关注, 当前, 食品工业的发展计划在国家“十一五”和“十二五”中明确提出了需要解决食品安全的问题。本文选择对食品安全领域的农产品中农药残留量检测结果数据(以下简称农残数据)进行可视化研究, 通过对农残数据样例进行分析发现, 农残数据是典型的层次数据, 并且具有明显的时间属性, 包括地域划分、产品种类和农药种类等层次属性, 对于农残数据的分析需要关注残留量的数值的数值变化, 因此需要关注采样时间。

综上所述, 农残数据具有典型的层次属性和时间特性, 采用 ThemeRiver 模型可以对时序数据进行有效展示, 结合层次数据可视化的一些方法就可以得到一种能够同时解决层次结构和时序特征的可视化方案。因此, 本文提出一种具有时变特性的大规模层次数据可视化方案, 通过将传统的 ThemeRiver 可视化模型进行改进并利用一些交互技术, 借鉴 Treemap 结点布局的思想, 形成一个具有能够显示层次特性的河流图。

1 相关工作

1.1 ThemeRiver 模型

ThemeRiver 模型是 Susan Havre 和 Lucy Nowell 在 2000 年提出的一种以河流为隐喻展示文本数据随时间流逝主题发生变化的可视化模型^[5-6], 在文本可视化领域有着重要的地位, 在会议话题^[7]、交通数据^[8]、搜索关键词排序^[9]等领域都有着十分广泛的应用。ThemeRiver 模型的优势在于能够同时显示多个主题随时间发展的演化经历, 提取与时间相关的数据的规律, 缺陷在于不能够表示数据更多的属性和信息, 并且不能够在图中直观的表现具体数值。经过近十几年的研究, 众多专家学者对 ThemeRiver 缺陷进行了改进。Martin Wattenberg 等人将 ThemeRiver 与平行坐标相结合构建了 History Flow^[10]用于分析维基百科被编辑的过程。此后, Martin Wattenberg 等人又提出一种基于 web 的统计美国新生儿名字流行趋势的系统 NameVoyage^[11], 以颜色的递进代表名字使用频率的范围。此后, Wattenberg 等人又 NameVoyager 进一步改进, 将其应用到社会数据分析中^[12]。刘世霞等人开发了一种话题增强型的 TIARA 系统^[13], 这个系统将 ThemeRiver 模型和标签云(Tag Clouds)技术相结合, 能在展现话题强度变化的同时, 突出显示每个时间段内各个话题内的关键字的出现情况。Wu 等人利用 ThemeRiver 能够展现数据趋势的优势, 将主题河流应用于臭氧层空洞情况的监测^[14]。

同时, 由于“河流”图的形式不仅能够展示连续数据变化情况, 同时能够利用河流的宽度与边缘线的倾斜程度表现事情的发展情况, 很多专家学者利用河流的隐喻来表现事情随时间的变化^[15-17], 例如用于表现事件时间序列的 LifeFlow^[18]。另外, 一些主题河流的变形也引起了人们的注意, 如用于展现话题新闻变化的新型河流图 EventRiver^[19]与 CloudLines^[20]。

1.2 时序层次数据可视化

时序数据与层次数据都是一些特定数据的固有属性,二者并不矛盾,然而同时能够表现层次变化和数值变化依然是现在数据可视化的一个难点。Dou W 等人^[1]与 Cui W^[2]等人致力于研究层次话题随时间的变化情况。特别说明的是,他们都用到了河流的表现形式。

2 改进的 ThemRiver 模型

传统的可视化方法在解决层次数据的时候,往往无法兼顾其时间属性。而单独对时间进行可视化的方法,往往无法直观的展示层次结构特性。本文采取一种层次河流的可视化方案,将树图中空间填充表达层次的思想与 ThemeRiver 模型相结合,同时能够反映数据的层次关系与随时间变化的趋势。

2.1 数据处理

现实生活中原始数据存在数据量大,数据结构复杂,离群点多,数据缺失,数据错误等缺陷,这为可视化过程造成了很多困难。因此,在设计可视化方案之前需要对数据进行预处理,保证数据能够有效展示。

本文采用的数据是层次结构不太复杂(太过复杂的层次结构会在交互的过程中使用户迷航,一般层次结构不超过 4 层),数据量较大,且随时间变化的非连续数据。以农药残留数据为例,农药残留数据是由各地检疫检验局上传,具有数量大,上传时间不固定,具有典型层次结构等特点。由于数据采集是随机的,时间并不连续,有时一段时间内数据采集过多,也可能一段时间内样品数量较少,这样会导致一段时间内的数据并不连续,且缺失的采样点会造成可视化布局的失误(如果数据不进行处理,未采样时间点的采样值按照 0 来计算明显是错误的)。通常情况下采样属性如表 1 所示。

表 1 数据属性

属性	说明
SAMPLENAME	样品名称
SAMPLEAREA	大区
SAMPLEPROVINCE	省份
SAMPLEPREFECTURE	市县
SAMPLEFROM	来源
SAMPLETIME	采样时间
DETECTIONITEM	检出农药名称
DETECTIONRESULTS	检测数值
MRL	农药门限值

采样的数据往往过于集中,且很多时候含有很多的离群点(数据噪声)。例如连续多日进行检测,此时数据量在很多地方十分密集,此时,若直接将 these 数据进行可视化分析,较少的时间区间点可能会使得后期曲线拟合出现较大的失误,同时不利于对数据进行预测。在 Themriver 原始文献中采取所有数据点全部保留的方式,这种方法对于连续性的数据较为有利,对于离散数据,尤其是在获得数据较少的情况下(由于人为因素导致采样数据较少),完全按照真实的数据则会导致绘制出来的河流图不能够反映数据的真实规律。本文提出一种对原始数据进行处理的改进算法。其算法定义如下:设总布局时间跨度为 T ,时间粒度为 G ,时间结点 $t_i, i \in (0, n)$,其中 $n=T/G$,设置门限个数为 k , k 由用户指定。以第 i 个时间结点为例,如图 1 所示,若第 i 个时间结点周围数据点个数 M 多于 k (以 $k=3$ 为例),则所有数据结点保留,若不足 k (临界点公用),则通过对最近不同年份相同时间内的结点的取平均操作进行结点的添加,使其结点个数达到门限值 k ,最后形成的结点集 S 。添加结点的具体方法为:以图 1 中 t_{i-1} 段为例,设三年同时段的数据按照出现时间点频率进行排序,出现时间点频率最高的数据集为 D ,对 D 内的数据进行平均操作,并将结果添加到当前年份的时间段内对应的时间点,如图 1 中蓝色点所示。若仍不足 k ,则选择低于最高频率的时间点数据集进行相同的操作进行添加,直到添加该区域内数据点数目达到 k 。

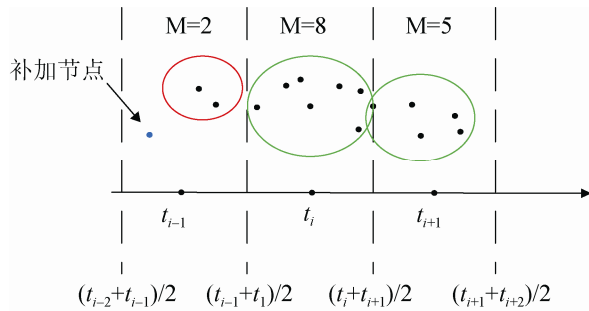


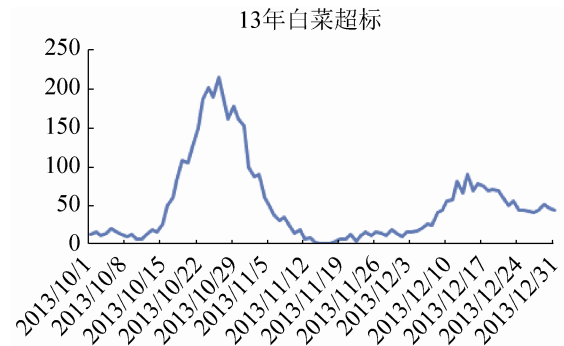
图 1 结点集 s 的形成过程

2.2 曲线拟合

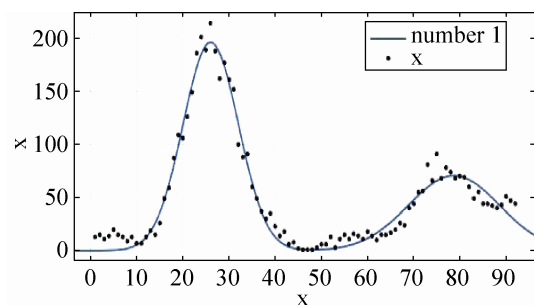
在传统时序数据可视化方法中, 解决较为离散的数据的方法是折线图, 不仅能够显示数据细节, 而且能够反映数据走势, 以 2013 年 10 月到 2013 年 12 月份全国白菜超标例数的显示为例, 图 1(a) 为传统的折线图。很清楚的发现, 在 10 月 22 日前后超标数量达到最高, 但当数据量大时, 数据中可能出现的异常数值, 例如极高点, 此时, 图形中就会出现尖锐的图形, 这种图形无论是准确性和美观上都无法满足人们的需求。

并且, 大数据条件下, 主题河流展现的是总体数据的趋势情况, 对于细节的把握虽然不理想, 但也正是这样的特点, 能够让人们快速的发现数据的规律, 定位感兴趣的时间点。因此, 相比折线图, 经过曲线拟合的曲线图更能够反映数据的规律且符合人们的审美。同时, 优秀的曲线拟合方法, 能够起到对未知数据的预测的效果。最初, ThemeRiver 模型采用的是三次样条插值拟合曲线的方式进行绘制, 保留所有原始数据, 这种绘制方法虽然能够最大的保证数据的真实性, 但是对于“脏数据”十分敏感, 错误的数字将会对图形的形状影响很大。高斯模型和傅里叶模型是较为常见的两种拟合方法, 本文分别采用高斯模型与傅里叶变换模型上述数据进行绘制实验。图 1(b) 为采用高斯模型(二次)进行拟合的结果, 拟合方法见公式(1), 图 2(c)为采用傅里叶模型(三次)进行拟合的结果, 拟合方法见公式(2)。实验结果表明, 图中共有 92 个数据点, 在拟合的准确度表现上, 高斯模型表现

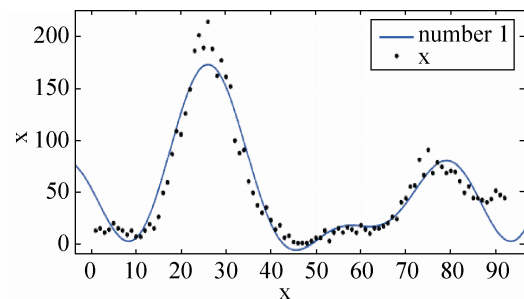
的更为出色, 能够保证 90% 以上的数据点均匀的分布在曲线的两边。因此, 本文采用高斯模型进行曲线的绘制过程。



(a) 折线图



(b) 高斯模型



(c) 傅里叶模型

图 1 绘制 2013 年白菜超标情况的三种图形

二阶高斯拟合:

$$f(x) = a_1 e^{-((x-b_1)/c_1)^2} + a_2 e^{-((x-b_2)/c_2)^2} \quad (1)$$

本例中, 经过测试, $a_1 = 196$, $b_1 = 26.03$, $c_1 = 8.472$, $a_2 = 70.51$, $b_2 = 78.7$, $c_2 = 13.93$ 。

三阶傅里叶拟合:

$$f(x) = a_0 + a_1 \cos(wx) + b_1 \sin(wx) + a_2 \cos(2wx) + b_2 \sin(2wx) + a_3 \cos(3wx) + b_3 \sin(3wx) \quad (2)$$

本例中, $a_0=53.85$, $a_1=-13.37$, $b_1=-39.82$, $a_2=-46.07$, $b_2=-21.28$, $a_3=0.301$, $b_3=2.5$, $w=1.974$ 。

2.3 布局算法

ThemeRiver 图形其实是堆叠图 (Stacked Graphs) 的一种变形, 具有堆叠图的典型特征。不仅能够展现时间序列上主题的变化, 而且通过区分能够使每个主题具有各自独立的意义。堆叠图的布局问题也是专家学者研究的重要部分, Lee Byron 等人^[5]认为处于美学考虑, 堆叠图的布局过程不宜过于摆动, 同时过多的摆动容易会使得用户花费更多的时间找寻自己感兴趣的主体。因此, 采用传统 ThemeRiver 的零基础坐标的方式能够有效去除图中过于尖锐的形状, 同时能够减少图形中斜坡的斜率过大的问题。

本文采用 ThemeRiver 模型与类似 Treemap 中的 Slice and Dice 相结合的思想进行布局。首先, 对于各个大的主题, 用一种颜色进行表示, 在颜色内部, 用不同的饱和度的同一色系的颜色表示内部结点, 其中布局的方法采用刚开始先来后到从下到上的布局方式, 对一段时间后插入的新的主题计算宽度进行比较并重新布局, 这样, 每个层次采用堆叠的方式进行叠加, 最后形成一幅完整的河流图。

2.3.1 布局顺序

本文采用的一种计算主题重要性, 依照重要性对新加入主题布局的方法, 这种方法同时兼顾了主题之间的顺序性与连贯性。方法如下:

Step1: 开始时对已存在的主题按照一定规则进行排列, 对所有主题由下向上进行布局。

Step2: 随后发起的主题, 计算其瞬时主题强度 w , 此时与最底下的主题强度相比, 若比最底下的主题的宽度小则将该主题插入在所有主题的底下, 否则置于所有主题的最上方。

这样做的好处是: 其一, 不用重新进行拆分布局, 减少交叉。其二, 能够保证中间位置的主题是较旧主题, 两边主题是较新主题。同时, 位于整体

布局上方主题是下降主题(刚开始主题强度较高), 下方主题是上升主题(刚开始主题强度较低), 有助于用户快速的找到自己感兴趣的主体并且对于主题的先后顺序有一个宏观的展示。其过程以下列伪代码描述, 其中 n 为河流中所有涌流(current)的个数:

布局顺序算法

current cur[n]; //河流总数为 n

for (int i = 0; i < n; i++)

{

if (cur[i].time == 0) layoutUp(cur[i-1]);

else{

calculate(cur[i].width);

calculate(cur[i-1].width);

cur[i].time.width > cur[i-1].time.width?

layoutUp(highest(cur)):layoutDown(lowest(cur));

}

}

每个层次结构按照上述方法布局完成后, 可按照 Treemap 中的 Slice and Dice 的布局方式对各个大的层次结构以此布局, 在农药残留检测的可视化视图中小层次结构展现的是各个省市, 而大的层次结构展现的是各个大区。

2.3.2 颜色选择

对于每个主题的颜色分配, 传统的 ThemeRiver 采用了相似主题采用相同色系的不同颜色进行绘制, 且不同主题间要颜色区分鲜明。因此, 本文遵循传统 ThemeRiver 的配色原则, 用颜色差异较大的不同色系显示不同主题。而内部的主题结构用相同色系不同饱和度的颜色进行布局。约定颜色越深表示主题存在时间越长, 由布局顺序限制, 通常情况下, 两侧的主题由于存在时间较短, 因此颜色浓度逐渐变浅。如图 2 所示, 布局结果为某大区各个省的 2013 年全年白菜超标情况布局结果。

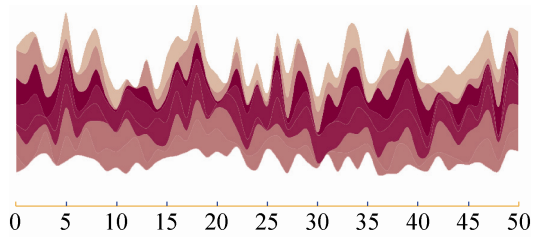


图 2 布局结果
(某大区 2013 年各个省白菜超标情况布局结果)

2.3.3 标签分布

标签分布要遵循两个原则,一个是不能够对其他层遮挡,其次们能够将每个涌流进行明确的区分。在涌流较小的情况下,标签可以直接分布在每个涌流层上面。当涌流过多时,这样的标签就没有地方能够显示。此时需要交互技术的支持,通过鼠标滑动和点击能够提示标签信息。

当主题过多时,每个子主题的表现会收到显示区域的影响,主题河流的涌流(current)越多增多,视觉布局混乱。采用交互技术能够有效减小视觉杂乱的影响,用纯色代表每个主题,通过点击某个主题,其他主题颜色淡去,放大该主题显示内部子主题是一个较为简单的解决方法。如图 3 所示。

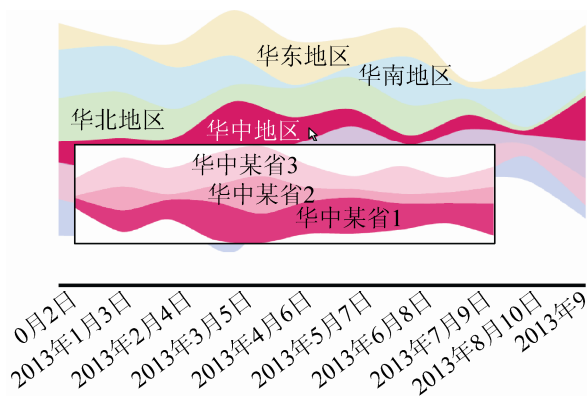


图 3 标签分布

3 应用与评价

3.1 农残数据中的应用

使用模拟的农残数据,分别利用雷达图和本文提到的 ThemeRiver 可视化方法进行展示农药在同一个省市不同地区的超标情况,通过简易直观的分析加上交互手段能够较为容易的对农药超标情况

进行监测,同时对得出相应的结论。

采用雷达图的表现方式有助于对周期模式的挖掘,通过对每一年的每个月的内的数值进行分类,找出能够代表本月份的数值,在雷达图上进行显示。以 2011-2014 年的数据为例,如图 4 上方雷达图所示。用不同的颜色表示不同的年份,能够清晰的看出,白菜超标频次较高的月份为八月份和十月份,而一月份超标例数较低。出现这种现象的原因有很多,有可能当月检测力度不够,也可能白菜上的农药经过一冬天的降解,农药浓度减小,使得超标的频次降低。通过在雷达图上选择不同年份不同月份(可以连续选择,默认为全部)可显示这段时间内的主题河流,图 4 下方展示的是 2013 年 10 月份的 A 省和 B 省白菜超标数据。通过点击每个涌流能够显示详细的数据信息。

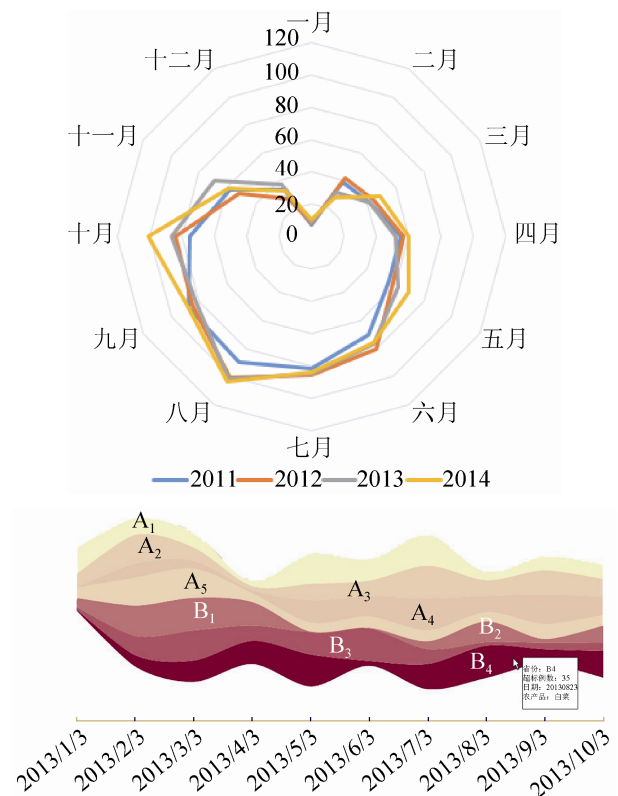


图 4 2013 年 10 月 A 地区和 B 地区白菜超标主题河流

3.2 评价

评价一个“河流图”好坏的标准主要有两个方面。其一,可读性^[21]。其二,美学标准。

可读性的标准在于“河流图”每一层的布局顺序。本文采用的 ThemeRiver 的布局方式借鉴了 Treemap 中的 Slice and Dice 的布局方法,初始布局由下到上。充分满足可读性标准。采用颜色的深浅不仅能够区分每个子主题,同时能够给研究人员一个对主题存在时间的大致感知。

美学标准在于是否能够吸引用户,本文采用了颜色分明的配色方案和简单易用的交互手段,基本符合用户的审美要求。

4 结论

本文通过改进 ThemeRiver 模型中的曲线拟合和配色方案,参考 Treemap 中结点的布局表示层次关系和顺序的思想,提出了一种能够表现非连续数据的新型时序层次数据可视化方法。实验结果表明,这种可视化方案能够充分利用 ThemeRiver 的特点对时序数据的发展趋势进行可视化展示,同时能够显示层次结构的变化特点,此外,能够对未知时间的数据进行合理的推测。该方案具有以下特点: 1) 即使是在数据值有较小部分错误(或缺失)的情况下,依然能够进行合理的绘制; 2) “河流”图具有主题布局连续的特点,利于对数据进行查找; 3) 在数据层次结构更新时布局变化较小,利于主题进行长期追踪。运用层次河流在处理农药残留量数据可视化时,保持了数据原本的地理层次特性,同时能够对时间属性进行合理的利用,能够令使用者方便的进行对时间规律的把握。此外,引入雷达图进行辅助显示,有助于快速定位感兴趣的时间区域,同时对时间的周期性规律的探查也有所帮助。通过实验分析,层次河流不仅仅可以用在农残数据可视化分析上,还可以推广至其他领域具有时序特征的层次化的数据信息的可视分析应用中,特别是对于数据量大具有时序规律的数据,例如每年各景点人流量数据、各地天气情况等,使用层次河流均可以达到一个较好的可视化效果。

参考文献:

[1] Dou W, Yu L, Wang X, *et al.* Hierarchical topics:

Visually exploring large text collections using topic hierarchies [J]. Visualization and Computer Graphics, IEEE Transactions on(S1077-2626), 2013, 19(12): 2002-2011.

[2] Cui W, Liu S, Wu Z, *et al.* How hierarchical topics evolve in large text corpora [J]. Visualization and Computer Graphics, IEEE Transactions on(S1077-2626), 2014, 20(12): 2281-2290.

[3] Liu S, Cui W, Wu Y, *et al.* A survey on information visualization: recent advances and challenges [J]. The Visual Computer (S0178-2789), 2014, 30(12): 1373-1393.

[4] Andrienko N, Andrienko G. A visual analytics framework for spatio-temporal analysis and modeling [J]. Data Mining and Knowledge Discovery (S1384-5810), 2013, 27(1): 55-83.

[5] Havre S, Hetzler E, Whitney P, *et al.* Themeriver: Visualizing thematic changes in large document collections [J]. Visualization and Computer Graphics, IEEE Transactions on(S1077-2626), 2002, 8(1): 9-20.

[6] Havre S, Hetzler B, Nowell L. ThemeRiver: Visualizing theme changes over time [C]// Information Visualization, 2000, InfoVis 2000, IEEE Symposium on. USA: IEEE, 2000: 115-123.

[7] Dörk M, Gruen D, Williamson C, *et al.* A visual backchannel for large-scale events [J]. Visualization and Computer Graphics, IEEE Transactions on(S1077-2626), 2010, 16(6): 1129-1138.

[8] Guo H, Wang Z, Yu B, *et al.* TripVista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection [C]// Pacific Visualization Symposium (PacificVis), 2011 IEEE. USA: IEEE, 2011: 163-170.

[9] Shi C, Cui W, Liu S, *et al.* RankExplorer: Visualization of ranking changes in large time series data [J]. Visualization and Computer Graphics, IEEE Transactions on(S1077-2626), 2012, 18(12): 2669-2678.

[10] Wattenberg M, Viégas F. Beautiful history: Visualizing Wikipedia[M]// Beautiful Visualization: Looking at Data Through the Eyes of Experts. America: O'Reilly Media, 2010: 175-192.

[11] Wattenberg M. Baby names, visualization, and social data analysis [C]// Information Visualization, INFOVIS 2005, IEEE Symposium on. USA: IEEE, 2005: 1-7.

[12] Wattenberg M, Kriss J. Designing for social data analysis [J]. Visualization and Computer Graphics, IEEE Transactions on (S1077-2626), 2006, 12(4): 549-557.

(下转第 2474 页)

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

• 2466 •