DOI:10.13203/j.whugis20130794　　文章编号:1671-8860(2015)08-1080-08

一种基于标签云的位置关联文本信息可视化方法

华一新[1]　李　响[1]　赵军喜[1]　王丽娜[1]　张　晶[1]

1　信息工程大学地理空间信息学院,河南 郑州,450052

摘　要:互联网的广泛应用产生了越来越多与地理空间位置关联的文本信息。现有地理信息系统一般通过外部链接来浏览这些数据,需要频繁的缩放、漫游和点击操作,而其他方法又难以有效表达出空间位置关系。提出了一种基于标签云的位置关联文本信息可视化方法——标签云地图,给出了标签云地图的设计思路和实现流程,并以腾讯微博的真实数据集为例建立了原型,重点研究了点状和面状地理要素的 Cartogram 生成算法,关键字和词频的提取算法,面向不同尺度和不同时间的标签云显示规则的标签位置生成算法。实验表明,该方法能够帮助用户从大量的位置关联文本信息中快速感知并把握信息的总体特征和发展趋势。

关键词:标签云;标签云地图;文本信息可视化;位置关联;可视化

中图法分类号:P289; P208　　**文献标志码**:A

现实生活中存在大量与地理位置或者区域相关联的文本信息,比如一个地方的地志,某个地点发生的新闻,以及网络上与地理位置关联的微博、微信等信息。这些信息被称为位置关联(英文中一般称之为"geo-referenced"或者"geotagged")文本信息。随着众包、自发地理信息等概念的流行[1],出现了大量与地理位置关联的文本信息。如何进行有效的可视化,为信息的探索和发现提供有力支持成为需要解决的问题。

传统的位置关联文本信息可视化方法主要有以地理信息为主的可视化方法和以文本信息为主的可视化方法。

传统的地理信息系统软件(如 ArcGIS、SuperMap 等)根据不同类型的文本信息进行可视化。结构化的文本信息作为地理要素的属性信息存储在关系表中,点击某一个地理要素时,与之关联的文本信息会以数据表格的形式呈现出来。对于非结构化的文本信息,则采用一种外部链接的方法,即该地理区域保存了所有与之关联的文本存储位置,当点击该区域时,由相应的外部程序(如记事本、浏览器等)或者内嵌在地理信息系统软件中的插件(Office,Adobe 插件)打开文本。

这种可视化方法非常适合表达传统的"几何+属性"结构的地理信息,但是对于互联网中大量非结构化的位置关联文本信息,用户难以从这种可视化的形式中发现和探索有用信息。

大百科全书软件(如微软的 Encarta、维基百科以及百度百科等),采用的是与地理信息系统软件截然不同的思路,以文字为主体,文字所关联的地理空间位置则由偏安一隅的地图来表示,如图 1 所示。这种表示方法过于侧重表达文本信息,空间信息的表达过于简略。

基于此,本文提出了一种基于标签云的位置关联文本信息可视化方法。

图 1　以文本信息为主的可视化方法

Fig. 1　The Visualization Method for Text Information

1　基于标签云的位置关联文本信息可视化方法

作为非空间文本信息表达的有效方法,标签

收稿日期:2013-12-17

项目来源:国家自然科学基金资助项目(41401467,41271450,41371383);国家科技支撑计划项目(2012BAK12B02)。

第一作者:华一新,教授。主要从事地理信息系统应用相关工作。E-mail:13607680789@163.com

通讯作者:李响,博士。E-mail:Helloj2ee@126.com

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

云(tag cloud 或 word cloud)最早以“潜意识文档(subconscious file)”一词出现在 Douglas Coupland 的《Microserfs》一书中[2]，此后经 Flickr 网站首次应用之后便得到广泛使用。很快，标签云也用在了空间文本信息可视化中。

1.1　标签云和传统地图相结合法

文献[3]最早将标签云应用到地理信息可视化研究中，根据调查报告中巴黎地标物出现词频的高低，将它们以不同大小的字体标注在巴黎地图上，形成巴黎的心理认知地图。文献[4]通过标签云的方法，将具有地理标签的海量照片信息与地图关联，并进行可视化。此后，也有学者在文献[4]的基础上利用 mash-up 工具将标签和标签云叠加在地图上[5,6]。文献[7]以新闻和科技文献为研究对象，通过标签云的方法进行地理可视化从而及时感知公众健康情况。以上方法将标签云和传统地图很好地结合，但标签云会和地图上原有的注记产生冲突，容易使用户读图困难，甚至产生错误的理解，如图 2 所示。

图 2　旧金山拍摄的照片标签云和旧金山地图相叠加的结果[4]

Fig. 2　Some Photo Tags Conflicted with Labels on San Francisco Map

1.2　标签云填充地图法

部分研究学者注意到了标签云和传统地图结合的不足，于是省略了地图上原有的地理要素，通过标签云来填充地图。

文献[8]中绘制的“世界简史地图”，数据来源于维基百科中各国的历史文章，然后用其中最具代表性的一个单词描述一个国家，总共用 176 个单词通过变形和填充，绘制了一幅世界地图。这幅“世界简史地图”不仅制作精美，而且单词高度凝炼了一个国家的简史，比如中国版图就是用“Dynasty(朝代)”单词变形绘制而成。文献[9]提出了一种 Taggram 地图，和文献[8]的“世界简史地图”不同的是，该图用了多个单词，按照该单词在网站中的(比如 Flickr)流行程度以不同字体、大小填充在相应的国家版图中。

上述两种地图虽然有不同程度的变形，但是均保留了国家版图的主要特征。由于省略了大量的地理要素，标签云和地图结合得也非常完美，标签显示非常清晰。但是上述两种方法存在以下三点不足：

(1) 两种方法均只适用于面状要素，而不适合点状要素；

(2) 文献[8]中的“世界简史地图”虽然精美，但为手绘地图，并不适合计算机自动生成；

(3) Taggram 保留了行政区划形状的真实性，且里面的标签排列是随机生成的，因此标签的位置很容易让读图者产生误解。本文按照文献[9]的方法，根据百度百科中“中国”词条和中国地图形成了一幅 Taggram，如图 3 所示。在图 3 中，读图者很容易误解“北京”是在中国的西部，也容易联想“政党”、“泥石流”和“漠河”这样的词汇是否和台湾岛存在潜在的联系。

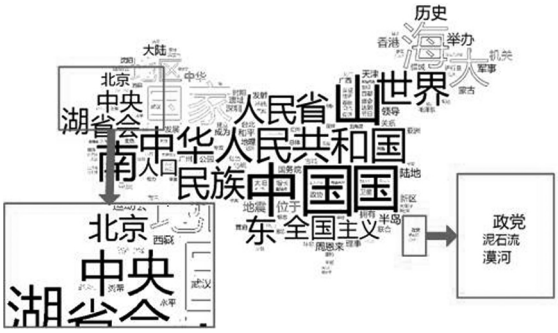


图 3　通过百度百科“中国”词条和中国行政区划图形成的 Taggram

Fig. 3　A Tag Clouds in a Region of China Generated from Articles of Baidu Encyclopedia About China

2　基于标签云的可视化方法

本文针对上述方法的不足，将标签云和 Cartogram 地图相结合，提出了一种新的基于标签云的位置关联文本信息可视化方法。

2.1　Cartogram 地图

Cartogram 有多种中文译法，比如统计地图、拓扑图或者示意地图等，为避免歧义，本文统一称之为 Cartogram。它是一种根据某种属性值将对象形状进行夸大或缩小的地图，能保持位置的相对正确，基于属性进行夸张变形，直观地传递某种特定信息。Cartogram 可以分为 4 类：离散 Cartogram，连续 Cartogram，多林 Cartogram(由英国利兹大学的多林发明，并以他的名字命名，又称之为圆形 Cartogram，与之类似的还有矩形 Carto-

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

gram）以及伪 Cartogram（Pseduo Cartogram）[10,11]。

离散和连续的 Cartogram 和原地图相比，尽管存在变形，但它们是在原地图的基础上根据属性值的不同对原有地理要素扩充或者缩小，而多林 Cartogram 则直接以圆形或者矩形取代原有地理要素。

伪 Cartogram，是一种并不完全遵守 Cartogram 变化法则的 Cartogram。最为著名的伪 Cartogram 由 Waldo Tobler 博士于 1986 年提出，这种 Cartogram 并不依据属性值对地理要素放大或者缩小，而是保持各要素变化前后的正交性（若变化前要素 A 在要素 B 的正北方，那么变化后要素 A 仍然要保持在要素 B 的正北方）的一种变形[12]。

2.2 标签云和 Cartogram 相结合法

本文借鉴了多林 Cartogram 和伪 Cartogram 的思想，利用规则格网设计并生成了一种适合标签云叠加的 Cartogram。该 Cartogram 将点状和面状地理要素根据不同权值转换成离散圆，同时又保持了正交性的特点，然后将标签云和这种 Cartogram 相结合，生成了一种新的可视化地图作品，本文将其称之为标签云地图。

以中国及其周边 19 个国家的百度百科词条为例，具体说明该可视化方法。图 4(a)是普通的行政区划图，以各行政区划的中心点生成离散单位圆（图 4(b)），将这些离散单位圆按照水平和垂直方向重新布局，使其不相互压盖（图 4(c)），并且保持相对位置的准确性。按照权重的不同，为离散单位圆设置大小不同的直径或者填充不同的颜色，图 4(d)是依据百度百科对各国描述的文字数量所计算的权重。然后在不同单位圆之间建立连接关系，图 4(e)中将陆上边界相邻的国家之间以直线相连。图 4(f)～4(i)是依不同比例尺对各单位圆进行显示控制，当标签云地图随比例尺放大时，首先显示出国名（图 4(f)）、再依次显示出国名和 50 个标签（图 4(g)）、国名和 100 个标签（图 4(h)）、国名和 200 个标签（图 4(i)）。

与 Taggram（图 3）相比，标签云和 Cartogrom 相结合的方法既能适用于面状要素又能适用于点状要素。由于标签云并不按照地理要素的真实形状填充，而是统一按照圆形进行填充，避免了用户对标签位置产生误解。同时，由于保持了相对的正交性，以及采用直线相连表示邻接关系，使得地理要素的相对位置和拓扑关系得以保留。

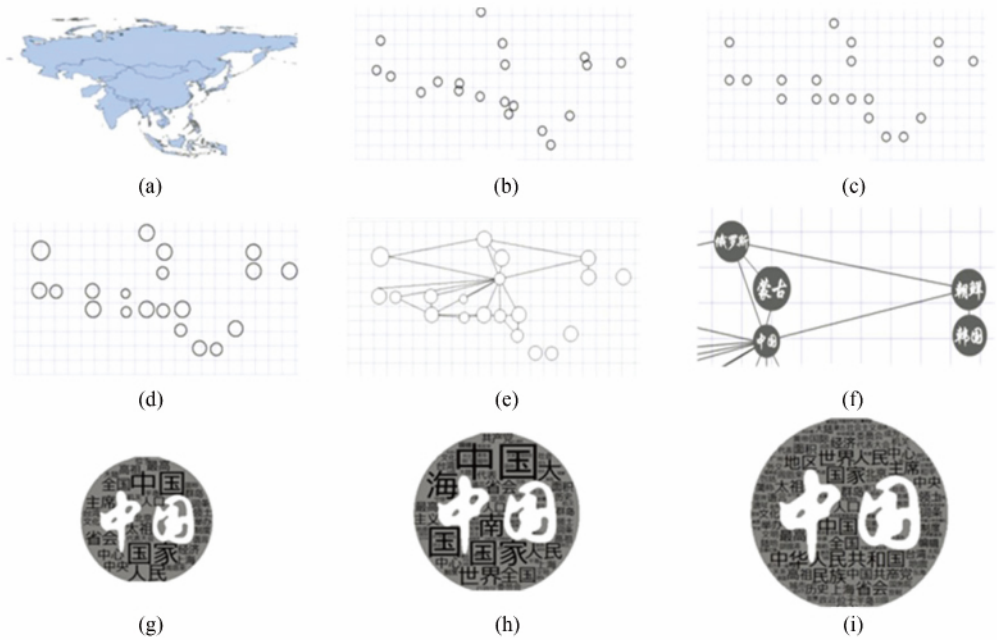


图 4 标签云地图实现流程

Fig. 4 Implementation of a Tag Cloud Map

2.3 标签云地图生成的基本流程

标签云地图生成的整体流程如图 5 所示。标签云地图的数据源是普通地图和与地理位置关联的文本信息。我们需要按照不同的指标（如旅游景点的点评数量、国土面积等）来生成 Cartogram。由于涉及的地理要素主要是点、面要素，因此需要分别设计点要素 Cartogram 生成算法和面要素 Cartogram 生成算法。针对大量的非结构化文本信息，通过词法分析和过滤来提取关键词和相应的词频，最后依据标签位置生成算法，生成

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

标签云地图。整个流程涉及三个核心算法，分别是 Cartogram 生成算法、汉语自动分词算法以及标签位置生成算法。其中汉语自动分词已经有较为成熟的商业化产品（http://ictclas.org/index.html），标签位置生成算法的研究也较为成熟，已有多个商业化产品，如 Wordle 和 Tagxedo 等。

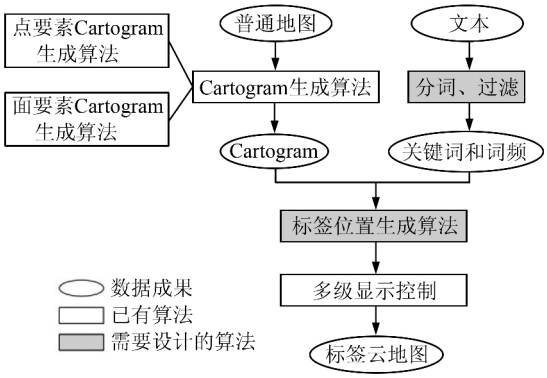


图 5　标签云地图的生成整体流程

Fig. 5　An Overall Process of Implementation of a Tag Cloud Map

因此，本文根据标签云地图的特殊需求，以腾迅微博部分城市位置微博文本数据为实验对象，重点解决地理点要素和面要素的 Cartogram 生成算法以及标签位置生成算法中显示规则模型的设计。图 6 是某一天全国部分大城市的标签云地图，在该比例尺条件下，显示了当天每个城市频率最高的三个词。

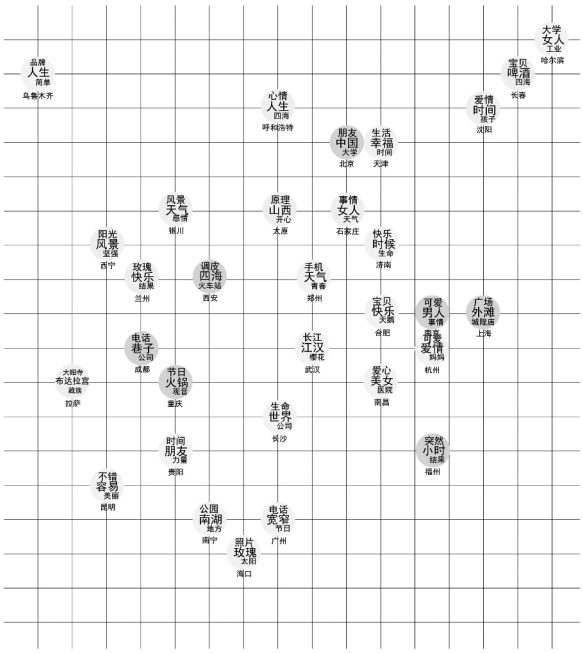


图 6　全国部分大城市标签云地图

Fig. 6　Top Three Words in Big Cities of China with the Tag Cloud Map Method

3　标签云地图的设计和实现

3.1　Cartogram 的生成算法

用户在微博网站上发布的部分消息会同时包含其位置信息，如城市和城市街区。在小比例尺地图上，可将城市看作点状要素，在大比例尺地图上，城市街区可看作面状要素。因此，本文中 Cartogram 算法主要关注点状和面状要素。针对点状要素，该算法的首要规则是所有点都分布于规则的网格交叉点上，这样便于浏览，可实现有序的可视化布局。同时，这也是与认知地图学的结论相吻合的，因为人们倾向于在水平方向或者垂直方向上来记忆位置之间的关系[13]。基于 Jo 和 Ryu 的算法思想[14]，我们对原本分布密度不规则的两个相邻点之间的距离进行简化，根据两点之间夹角 θ 的大小，仅保留两点中 X 轴和 Y 轴方向上距离较大的点，并且将较大的距离调整为标准单位 1，较小的简化为 0。即如果两个点之间水平方向的距离比较大，可以认为两个点在同一水平线上，垂直方向的距离简化为 0。具体如图 7 所示。

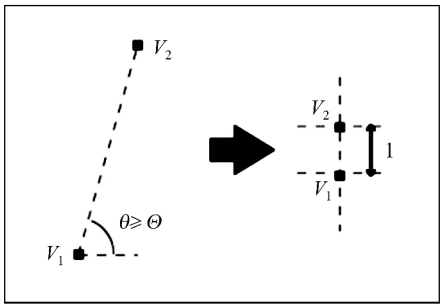


图 7　两点之间距离的简化

Fig. 7　Simplification of the Distance Between Two Points

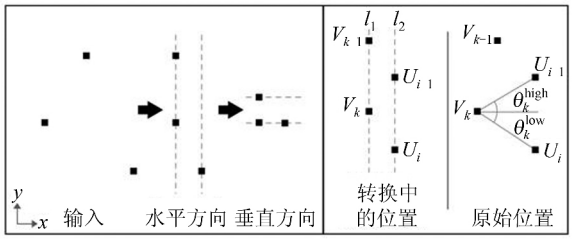
该算法首先进行横向压缩，然后进行纵向压缩（图 8(a)），方向从左至右、从上至下。假设 l_1 是一个 n 个 X 坐标相同的位置点的集合，l_2 是与 l_1 在 X 轴上相邻的一个集合。θ_k^{high} 和 θ_k^{low} 定义为 l_1 上的点 V_k 和 l_2 上的分别位于点 V_k 的上面和下面且距离 V_k 最近的两个点 U_{i-1} 和 U_i 的夹角（图 8(b)）。只有当所有夹角都不小于阈值角度 Θ（设为 45°）时，l_1 和 l_2 两个数据集合便可以压缩成相同的 X 值。在 Y 轴上重复该过程完成纵向压缩。这样就将所有位置点置于规则单元网格的交叉点上，如图 9 所示。

对于面要素，首先根据其中心点位置将所有面要素转换成点要素。然后，实施点要素的 Cartogram 生成算法。最后，将位置具有相邻关系的

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

点用直线连接起来，如图10所示。

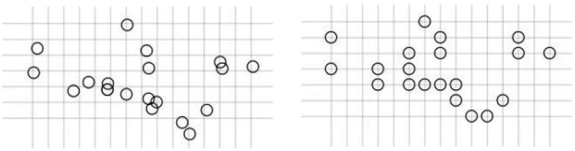
$$\min\{\theta_k^{\text{high}},\theta_k^{\text{low}} \mid k=1,2,\cdots,n\} \geqslant \Theta$$



(a)压缩序列 (b)定义与Θ比较的角度

图8 压缩序列与定义与Θ比较的角度

Fig. 8 Compression Sequence and the Definition of Two Angles to be Comparedto Θ



(a) 所有点的原始位置 (b) 重新调整后所有点都分布于网格交叉点上

图9 所有点的原始位置和重新调整后位置

Fig. 9 Actual Positions and Reallocated Positions of all the Points

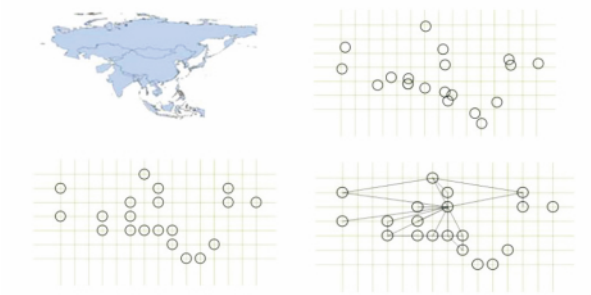


图10 面要素的Cartogram算法实现

Fig. 10 A Cartogram Algorithm for Area Features

3.2 关键字和词频的提取技术

很多流行的微博网站(如新浪和腾讯)都会提供API接口。根据这些API接口，可以获取用户发布的信息。如图11所示，我们基于腾迅微博API编写了一个小工具，该工具可以通过经纬度获取时间、地点、用户名、粉丝数量、转发数量、评论数量以及全文内容等信息，并且直接存储在本地数据库中。获取的文本数据可以直接通过现有的工具如ICTCLAS进行分词和过滤(详见http://www.ictclas.org)。最终可以获得文本的关键词和词频统计，如图12所示。

3.3 标签云显示规则的设计

关于标签云的生成已经有很多成熟的算法和工具(如Wordle，Tagxedo等)。因此Cartogram

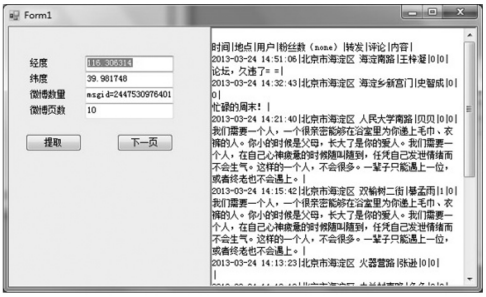


图11 获取微博信息的工具

Fig. 11 A Tool for Retrieving Messages from Microblog

	A	B	C
1	word	uniCount	inCount
2	(总词数)	348	0
3	难过	15	0
4	对方	8	0
5	明康	8	0
6	俗话	5	0
7	北京	5	3
8	女儿	5	0
9	儿子	4	0
10	商家	4	0
11	房地产	4	1
12	世界	3	0
13	北京大学	3	0
14	员工	3	0
15	朋友	3	0

图12 文本的关键词及词频

Fig. 12 Keywords and Their Frequency

和标签云结合的关键在于显示规则的设计。目前主要有两种显示规则，一种是面向不同尺度，另一种是面向不同时间。

3.3.1 面向不同尺度的显示规则

该规则是用离散的模型表达不同尺度上的相同对象，如图13所示。从国家级别到地区级别，我们设计了4种不同的离散模型(图13(a)～13(d))。用户逐渐放大地图，比例尺越来越大，标签云显示的内容会愈加详细。首先显示出来的是所有城市，每一城市用模型一表示，接着放大至模型二，最后是模型三。如果用户继续放大至城市级别，便会显示出城市的不同地区，我们用模型四来表示。在模型四中，相邻的地区用直线连接起来。随着用户继续放大，每一个地区又会重复该过程。

(a) 模型一 (b) 模型二 (c) 模型三 (d)模型四

图13 使用不同的模型表达相同的对象

Fig. 13 Providing Various Models to Represent the Same Object

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

当然，不同地区的信息量是有差异的，为了表示出这种差异，首先应用归一化的方法计算出每个地区所对应的标准信息量，考虑到要素之间可能产生的压盖，可采用不同的颜色来表示。

具体的计算过程如下：

$$M_{\text{normalized}} = \frac{M - M_{\min}}{M_{\max} - M_{\min}},$$

$M_{\text{normalized}} \in [0, 1/3)$ 时，模型用蓝色填充；$M_{\text{normalized}} \in [1/3, 2/3)$ 时，模型用黄色填充；$M_{\text{normalized}} \in [2/3, 1)$ 时，模型用红色填充。此处 M 表示地区的信息量，即每个地区发表的位置微博的总条数。对于每一模型，关键字使用模型填充颜色的相近色系来表示，如图 13(d)所示。每种模型的显示级别、最大词频数和标签大小值与屏幕大小、信息载负量等多个因素有关，以各要素尽量不压盖、显示清晰为目标，给出每个模型的实验参数，如表 1 所示。

表 1　每种模型的实验参数

Tab. 1　Each Model's Empircal Parameters

模型	最大级别	最小级别	最大词频数	标签大小/px
模型一	1∶19 百万	1∶4.75 百万	3	60
模型二	1∶4.75 百万	1∶2.38 百万	12	120
模型三	1∶2.38 百万	1∶0.59 百万	24	180
	1∶0.59 百万	1∶0.29 百万	3	60
模型四	1∶0.29 百万	1∶0.10 百万	12	120
	1∶0.10 百万		24	180

3.3.2　面向不同时间的显示规则

本文设计了两种时间标签云的显示方法。第一种类似于"sparkclouds"的思想[15,16]。用户移动鼠标至某一个关键词上时，它就会浮动出来并且放大显示。文字下面的波线图表示的是在一段时间内该关键词出现的频率，如图 14 所示。

第二种方法使用"瀑布"的隐喻，随时间变化的文本以瀑布飞流直下的形式分布。用户点击图上任何一个模型，右栏就会显示出一个"瀑布"式的标签云，如图 15 所示。

图 14　波线显示的关键词出现的频率

Fig. 14　A Polyline Appears Under Text to Represent It's Frequency of Recent Days

通过该方法，我们发现了一些在普通地图或微博上都难以发现但是有趣的事情。比如，在 2013 年 3 月 21 日这一天，很多在北京地区的人

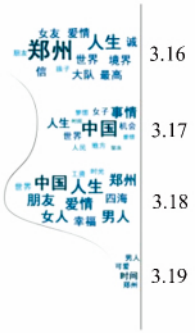


图 15　随时间变化的"瀑布"式标签云显示

Fig. 15　Using a "Waterfall" Methaphor to Represent Tags Change over Time

对一个名叫孟晓苏的人很感兴趣，而很多在贵阳的人却对一个名叫张妮娜的人关注颇多，如图 16 所示。在互联网上可以查找到关于孟晓苏和张妮娜的部分详细资料。

图 16　北京和贵阳在同一天的标签云

Fig. 16　Tag Clouds of Beijing and Guiyang on the Same Day

孟晓苏是原中国房地产开发集团有限公司的总裁，他在该时间段内连续发表了很多关于国务院出台的调控楼市新"国五条"的言论。该时间段内在北京地区，孟晓苏作为一个高频词出现，可能正好反映出北京市民对于房价问题和房地产相关政策法规的关心。张妮娜是一个平面模特，于 2009 年走红于网络，在该时间段内她发布了一系列照片，但她仅在贵阳地区作为一个高频词出现，这很可能与她是贵阳本地人有关。

4　结　语

标签云是关键词的视觉化描述，是文本分析处理结果的一种简单而高效的可视化表达方法，而地图则是表达地理位置的重要方法。因此，作为两者结合的标签云地图不仅提供了文本信息的描述，更是对地理空间数据进行表达和探索的地理可视分析的一种新尝试。它的特点总结如下：(1) 滤除了很多普通地图上不相关的细节信息，只保留了主要的信息，不同尺度下信息的详略程度不同；(2) 适用于点状要素，也适用于面状要素；(3) 标签云没有使用行政区域的轮廓，避免了

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

由于标签位置而产生的误解;便于用户浏览与地理位置关联的文本信息,减少了一些不必要的操作,并能够帮助用户在大量的位置关联文本信息中把握信息的总体特征和趋势。

本文给出了标签云地图设计和实现整体流程,并以腾讯微博的真实数据集为例,应用全国33个城市一个月内的微博消息数据设计并实现了一个原型。但这还不足以证明该方法的有效性,今后还需要收集更多的数据继续该研究,随着数据量的增加,进一步完善显示规则模型的设计。同时,Cartogram实现算法的实施效率也有待进一步改进。除此以外,还需要通过可用性测试来评价标签云地图的可用性,以及它与传统的地图可视化方法的比较。

参 考 文 献

[1] Li Deren, Qian Xinlin. A Brief Information of Data Management for Volunteered Geographic Information[J]. *Geomatics and Information Science of Wuhan University*, 2010, 25(4): 379-383(李德仁, 钱新林. 浅论自发地理信息的数据管理[J]. 武汉大学学报·信息科学版, 2010, 25(4): 379-383)

[2] Coupland D. Microserfs: A Novel [M]. New York: Harper Collins, 2011

[3] Milgram S. Psychological Maps of Paris [J]. *Environmental Psychology: People and Their Physical Settings*, 1976: 104-124

[4] Jaffe A, Naaman M, Tassa T, et al. Generating Summaries and Visualization for Large Collections of Geo-referenced Photographs [C]. The 8th ACM International Workshop on Multimedia Information Retrieval, Santa Barbara, California, USA, 2006

[5] Slingsby A, Dykes J, Wood J, et al. Interactive Tag Maps and Tag Clouds for the Multiscale Exploration of Large Spatio-temporal Datasets[C]. The Information Visualization, 11th International Conference, California, USA, 2007

[6] Wood J, Dykes J, Slingsby A, et al. Interactive Visual Exploration of a Large Spatio-temporal Dataset: Reflections on a Geovisualization Mashup [J]. *Visualization and Computer Graphics, IEEE Transactions on*, 2007, 13(6): 1 176-1 183

[7] Maceachren A, Stryker M, Turton I, et al. Health GeoJunction: Place-time-concept Browsing of Health Publications [J]. *International Journal of Health Geographics*, 2010, 9(1): 23-26

[8] Elmer M. Laconic History of Our World Map [OL]. http://maphugger.com/post/38323044556/laconic-history-of-the-world-2012-my-first, 2012

[9] Dinh-Quyen N, Schumann H. Taggram: Exploring Geo-data on Maps through a Tag Cloud-Based Visualization [C]. The 14th Information Visualisation (IV) International Conference, London, 2010

[10] Tobler W. Thirty Five Years of Computer Cartograms [J]. *Annals of the Association of American Geographers*, 2004, 94(1): 58-73

[11] Dorling D. Area Cartograms: Their Use and Creation [OL]. http://www.dannydorling.org/wp-content/files/dannydorling_publication_id1448.pdf, 2013

[12] Tobler W R. Pseudo-Cartograms [J]. *Cartography and Geographic Information Science*, 1986, 13(1): 43-50

[13] Tversky B. Distortions in Memory for Maps [J]. *Cognitive Psychology*, 1981, 13(3): 407-433

[14] Hyungeun J. Placegram: A Diagrammatic Map for Personal Geotagged Data Browsing [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2010, 16(2): 221-234

[15] Lee B, Riche N H, Karlson A K, et al. SparkClouds: Visualizing Trends in Tag Clouds [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2010, 16(6): 1 182-1 189

A Tag Cloud-based Visualization for Geotagged Text Information

HUA Yixin[1] *LI Xiang*[1] *ZHAO Junxi*[1] *WANG Lina*[1] *ZHANG Jing*[1]

1 Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450052, China

Abstract: With the wide application of Internet, there is more and more geotagged text information. Nevertheless, such data is currently browsed in Geographical information systems through external links, requiring frequent zooming, panning, and click operations. Meanwhile, other methods can not effectively express spatial relationships. A tag cloud mapping method—a tag cloud-based visualization for geotagged text information is presented in this paper. We provide an overall implementation process for the method and design and implement a prototype based on the real-life data sets of Tencent Microblog. We focus on designing cartogram algorithms for point and polygon features, the ex-

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

traction of keywords and frequency and the display rule models for different scales and different times in the label placement algorithms. This method can help users rapidly make sense of the important concepts and grasp main features and developing trends in a large amount of geotagged text information.

Key words: tag cloud; tag cloud map; text information visualization; geotagged; visualization

First author: HUA Yixin, PhD, professor, specializes in application of geographic information system. E-mail: 13607680789@163.com

Corresponding author: LI Xiang, PhD. E-mail: wln_map@126.com

Foundation support: The National Natural Science Foundation of China, Nos. 41401467, 41271450, 41371383; the National Science & Technology Support Program of China, No. 2012BAK12B02.

(**上接第** 1074 **页**)

[16] Wu Xiongbin, Cheng Feng, Yang Zijie, et al. Broad Beam HFSWR Array Calibration Using Sea Echoes [C]. CIE'06 International Conference, Shanghai, 2006

[17] Digital Radio Mondiale (DRM); System Specification[OL]. http://www.drm.org/wp-content/uploads/2014/01/DRM-System-Specification-ETSI-ES-201-980-V4.1.1-2014-01.pdf, 2014

[18] Cheng Feng, Wu Xiongbi. A Passive Channel Adjustmet Method Based on a Non-linear Antenna Array [OL]. http://www.freepatentsonline.com/EP1879043.html, 2008

Receiving Array Calibration Using DRM Frequency Reference Signal of Direct Wave for HF Passive Bistatic Radar

SHAO Qihong[1,2]　*WAN Xianrong*[1]　*YOU Jun*[1]　*CHENG Feng*[1]　*KE Hengyu*[1]

1 School of Electronic Information, Wuhan University, Wuhan 430079, China

2 School of Electrical & Electronic Engineering, Hubei University of Technology, Wuhan 430068, China

Abstract: Based on the characteristics of the digital radio mondiale (DRM), a new method of array calibration for high-frequency passive radar (HFPBR) by using frequency reference signal of direct wave is proposed. Firstly the signal structure of DRM is introduced, while the feasibility of take frequency reference signal as the calibration source is demonstrated. Secondly, taking into account the electromagnetic environment and time-varying non-stationary propagation characteristics, a method of choosing high quality direct wave period by the shift-invariant antenna pairs is proposed. Finally, assist by the known DRM transmitter position, we utilized this method deal with measured direct wave data which propagate in surface wave mode and sky wave mode, compared the result of array calibration value with known single frequency signal source. The results showed that the amplitude and phase compensation method in DRM-based HFPBR is effective.

Key words: passive radar; HF radar; frequency reference signal; array calibration

First author: SHAO Qihong, PhD candidate, specializes in signal processing of wireless communication and passive radar. E-mail: qihongshao@163.com

Corresponding author: WAN Xianrong, PhD, professor. E-mail: xrwan@126.com.

Foundation support: The National Natural Science Foundation of China, Nos. 61331012, U1333106, 41106156, 41074116; the Fundamental Research Funds for the Central Universities, No. 2012212020214.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net