doi:10.3772/j.issn.2095-915x.2017.04.007

情报学视角下的文本可视化应用

- 1. 中国科学技术信息研究所 北京 100038;
- 2. 西南大学计算机与信息科学学院 重庆 400715

郭传斌 1 刘琦岩 1 赵婧 1 袁嘉莲 2 李明倩 2 望俊成 1

摘要 文本可视化可以帮助人们在海量的文本文档中快速获取知识,随着信息技术的发展和信息需求的增加,文本可视化的应用愈加广泛。作为情报学的主要研究对象,文本可视化对基于文献资料的情报研究具有重要的价值。本文结合情报研究任务和文本可视化的应用场景,将27个案例分为基于文本概览、文本主题、文本情感和文本关系的四类应用,对近期文本可视化的研究状况进行了分析和总结。最后针对可视化方法在情报学的应用提出了建议,以期为相关研究人员提供参考。

关键词:情报学,文本可视化,信息可视化

中图分类号: G35

开放科学(资源服务)标识码(OSID)



Study on Text Visualization from the Information Science Perspective

- 1. Institute of Scientific and Technical Information of China, Beijing 100038, China;
- 2. School of Computer and Information Science, Southwest University, Chongging 400715, China

GUO ChuanBin¹ LIU QiYan¹ ZHAO Jing¹ YUAN JiaLian² LI MingQian² WANG JunCheng¹

Abstract Text visualization can help people acquire knowledge rapidly in vast amounts of text documents. With the development of information technology and the growing demand for information, text visualization have been widely used. As the main research object of information science, text visualization plays an important role in intelligence research based on the literature. In this paper, 27 recent studies were summarized according to overview, theme, emotion and relationship by combining intelligence analysis tasks and text visualization scene. Finally some suggestions for the development of text visualization in information science are proposed, which aim to provide reference for researchers.

Keywords: Information science, text visualization, information visualization

基金项目:本研究得到中国科学技术信息研究所科研项目预研基金项目"基于用户视角的政策信息可视化的研究"(项目编号:YY2016-05)和国家社会科学基金青年项目"专利信息的生命特征揭示和老化规律研究"(项目编号:12CTQ025)的资助。作者简介:郭传斌(1990-),硕士研究生;刘琦岩(1964-),经济学博士,研究员,研究方向:科技发展战略与政策研究,赵婧(1987-),硕士,研究方向:科研管理;袁嘉莲(1997-),研究方向:信息管理;李明倩(1997-),研究方向:信息管理;望俊成(1984-),通讯作者,情报学博士,副研究员,研究方向:可视化、专利分析,Email:wangjo@istic.ac.cn。

1 引言

实现对信息的理解是情报学研究的终极目的。传统的情报分析方法大都是基于科技文献的知识发现,随着大数据概念的提出及大数据时代的到来,数据量爆炸式的增长使得人们对其处理和理解的难度日益增大,情报学研究范畴已经明显的呈现出大数据的典型特征。在这一阶段,情报学研究中涉及的数据类型复杂,非结构性数据所占比重明显增大,数据处理难度加大,对智能型数据分析工具依赖性增强,对分析结果可视化呈现需求提高[1]。

情报研究的步骤包括情报搜集、资料整理、情报分析和形成情报产品等,相对于音频、视频等数据,无论是信息来源、分析过程还是成果的展示,目前最便于获取、应用最多的都基于文本信息。在海量的文本数据面前,传统的分析方法非常低效,文本可视化可以将文本中复杂的或者难以通过文字表达的内容和规律以视觉符号的形式表达出来,同时向人们提供与视觉信息进行快速交互的功能,快速获取和理解大量的信息^[2]。1990年可视化作为一个独立的领域出现,2007年可视化领域的研究主题由科学可视化转向信息可视化^[3],在此期间文本可视化的研究内容、方法和应用场景在不断地丰富,为人类洞察数据的内涵、理解数据蕴藏的规律提供了重要的手段。

我国学者从1998年开始关注情报研究中的可视化问题^[4],2003年周宁^[5]等人率先进行了文本型文献信息可视化的研究,其后信息检索可视化成为研究热点,2005年陈悦^[6]等人介绍了国外知识图谱研究的最新进展,之后知识图谱

和知识可视化广泛用于信息计量的研究。近20 年来,图情领域的可视化研究不断发展深化, 形成了馆藏数字资源可视化、信息检索可视 化、可视化分析技术与方法、学科知识结构分 析、可视化理论研究5个热点^[7,8],最近两年, 基于主题模型的可视化逐渐成为趋势^[9-11]。数据 时代的情报分析任务对可视化提出了更高的要 求,因此,笔者从情报学视角对近年来文本可 视化的研究状况进行了分析和总结,希望为可 视化方法在情报学的应用开拓思路。

2 理论与案例来源

2.1 可视化用于情报分析的基础

在情报研究中, 普遍接受的基础模型是"数 据、信息、知识、智慧"层次模型,以数据为基 层架构, 按照信息流顺序依次完成数据到智慧的 转换, 可视化方法满足的正是对从数据到知识的 探索。根据2014年Sacha提出的可视分析模型[12], 这种探索过程可以分为人和计算机两部分。在计 算机的部分中,数据被绘制为可视化图表,同时 也通过模型进行整理和挖掘。可视化图表既可以 显示原始数据的特性,也可以显示模型的结果。 在人的部分,用户通过三种循环的交互分析进行 反馈, 在探索循环中, 用户通过模型输出和可视 化图表寻找数据中可能存在的模式,基于此采取 一系列行动,例如改变参数,去产生得到新的模 型输出和新的可视化图表。在验证循环中、人们 通过模式洞察到数据的特点,产生可能的猜测。 最后,在验证循环之上有知识循环,不断的收集 验证循环中已被验证的猜测,总结为知识。

doi:10.3772/j.issn.2095-915x.2017.04.007

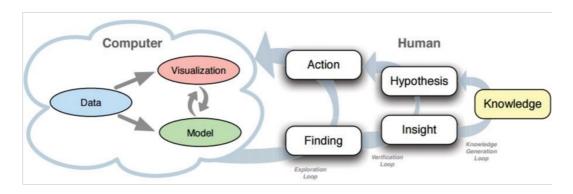


图1 可视分析的知识产生模型

因此,可视化提供了一种满足知识探索的 方法论,可以应用到信息组织、信息检索、信 息分析、信息服务的方方面面。从情报研究工 作的角度来看,可视分析结合了定性和定量的 方法,包含完整的信息分析流程,可以支持情 报研究人员更有效的分析理解信息。

2.2 文献调研与分类

可视化是一门综合了计算机图形学、数据挖掘、人机交互等领域的综合学科,相关研究众多,重要研究一般集中在会议发表。可视化领域的三大会议为IEEE VIS(含InfoVis、SciVis、VAST三个核心子会议)、IEEE EuroVis、IEEE PacificVis,两个顶级期刊IEEE TVCG和ACM TOG,所以笔者选择了这三个会议和两个期刊作为文献调研的重点。2007年是科学可视化和信息可视化研究的分水岭,所以选择以上文献中2007年之后的论文,使用"text visualization"作为关键词进行筛选,经过对相关文献的阅读和梳理,最终选择了27个相关的文本可视化案例。

情报研究的主要任务是跟踪、比较、预测、评价,涉及文本信息的提炼概括、分析推理,结合文本可视化的应用场景和情报研究的

任务,对27个研究案例从视图、特点和应用进行了分析总结,具体参见表1。为便于其中的可视化方法向情报研究迁移,将案例分为基于文本概览的应用、基于文本主题的应用、基于文本情感的应用和基于文本关系的应用四类。基于文本概览的应用旨在用可视化文本的主要内容,是对文本最基础的概括;基于文本主题的应用旨在揭示文本中语义层面的主题,观察文本中主题的演化;基于文本情感的应用旨在可视化文本的情感信息,将抽象的数据映射为直观的视觉特征;基于文本关系的应用旨在探索文本之间的复杂关系,寻找其中隐含的知识。

3 情报视角下的文本可视化应用

3.1 基于内容概览的应用

情报研究目的是为决策服务,将信息转化 为决策知识,情报分析成果的呈现方式在很大 程度上影响用户的体验和认知。随着文本数据 的爆发式的增长,可视化研究中最重要的命题 就是如何快速地了解和获取一个文本数据集中 所涵盖的主要内容,因此有大量的研究集中在 这方面。

表1 文本可视化案例

案例	主要视图	特点	应用
Wordle2009 ^[14]	标签云	自定义空间布局的标签云	文本概览
Geo Clouds2016 ^[15]	标签云	融合 GIS 信息的标签云	文本概览
Cui2010 ^[16]	标签云	动态标签云	文本概览
MorphableWordClouds2015 ^[17]	标签云	时序文本标签云	文本概览
RadCloud2016 ^[18]	标签云	多个标签云合并为一个视图	文本概览
Wordtree2008 ^[19]	树状图、标签云	树形结构表达文本词汇的前缀关系	文本概览
PhraseNets2009 ^[20]	网络图、标签云	含有语义关联的标签云	文本概览
SentenTree2016 ^[21]	树状图、标签云	含句法结构的社交短文本标签云	文本概览
LiteratureFingerprinting2007 ^[22]	像素图	文本的统计特征映射为视觉特征	文本概览
TEXTDNA2016 ^[23]	像素图	语言的时空分布和演化	文本概览
TIARA2009 ^[25]	河流图、标签云	文档集的主题变化	文本主题
TextFlow2011 ^[26]	河流图、标签云	文档集的主题变化	文本主题
ThemeDelta2015 ^[27]	桑基图、标签云	主题词的动态趋势变化	文本主题
$Xu2013^{[28]}$	河流图、标签云	社交文本主题的竞合关系	文本主题
EVORiver2014 ^[29]	河流图、标签云	社交文本主题的竞合关系	文本主题
IdeaFlow2015 ^[30]	桑基图、标签云	社交文本主题	文本主题
TopicPanorama2014 ^[31]	网络图	多源文档全局视角的主题关系演变	文本主题
Jiang2015 ^[32]	桑基图、标签云	跨领域的学科研究主题演变	文本主题
Oelke2009 ^[34]	像素图	时序文本情感信息布局方法	文本情感
PEARL2014 ^[35]	河流图	个人社交文本情感分析工具	文本情感
Assady2016 ^[36]	泡泡图	动态的意见变化展示	文本情感
CiteSpace ^[37]	网络图	学科热点和研究趋势分析	文本关系
PivotPaths2012 ^[39]	网络图	文本中实体的多维关系分析	文本关系
PaperVis 2011 ^[40]	网络图	引用文献的可视化	文本关系
TopicNets2012 ^[41]	网络图	跨学科知识探索与主题发现	文本关系
VAiRoma 2015 ^[42]	旭日图、河流图、热力图	罗马历史的文本分析和时空分析	文本概览
CiteRivers2015 ^[43]	河流图、像素图	文献主题和作者关系的综合分析	文本关系

通过可视化对文本关键信息的概览,能够减少视觉复杂性,从而获得对内容的快速理解。在文本内容的可视化技术中,标签云(tag-cloud,或词云word cloud)是最简单、最常用的一种展现形式^[13]。标签云通常采用词袋模型、TF-IDF等算法度量文本的词频,将文本中的关键词汇按照一定的原则排列在给定的视

图空间,利用字体的大小、字体、字体粗细和颜色等等反映词汇分布的差异,从而获得对文本内容的快速概览。Wordle(2009)^[14]在空间布局上对标签云进行了扩展,通过调整关键词在视图空间的填充路径,提高了空间利用率和外观艺术性。GeoCloud(2016)^[15]是Wordle的改进,将关键词信息和文本的地理空间信息结

doi:10.3772/j.issn.2095-915x.2017.04.007

合起来自动化地生成指定地理图形样式的标签 云,在提高视图空间利用率的同时保持了携带 信息的完整性。静态的标签云适用于展示整体 内容的概览,而通过交互和动画的方式能展示 更丰富的信息。Cui2010^[16]设计了可交互的动 态标签云,对时间序列的文本数据集进行可视 化,通过一个时间轴的窗口,既展示关键词数 量的变化,又能拖动时间轴动态展现关键词随 时间的动态变化。Morphable Word Clouds^[17]结 合时间线展示了一种可用于叙事的标签云,布 局时将每一个关键词视为一个刚体,在时间线 上以动力学模型约束,在紧凑的空间里达到美观的效果,进一步提升了可读性。

在很多情况下需要同时比较多个文本的内容,但是将多个标签云排列在一起在视图上会造成认知的困难。RadCloud^[18]通过将多个文本的标签云合并为一个视图解决了这个问题,它采用径向布局将所有词汇放在一个圆圈内,文字大小和位置对应词向量高低,圆弧的颜色对应文本的来源,词汇下方的色块显示词汇在各个主题分布的比例,通过文字颜色可以直观的看出词汇最主要的来源。

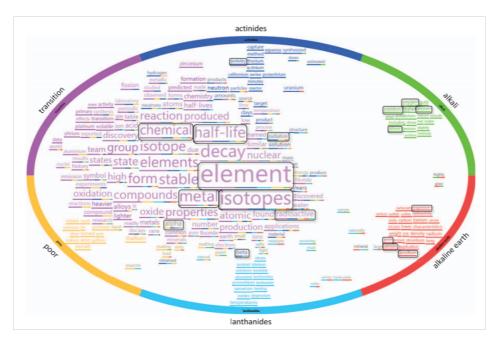


图2 维基百科中六个金属类别的文章的RadCloud,每类文章下面包含100个词汇

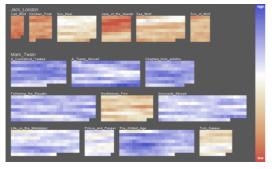
相比传统文本而言,社交网络上存在大量重复的短文本,为标签云加上文本间的关联能够刻画含句法细节的全局信息。基于Wordtree (2008)^[19]和Phrasenets (2009)^[20],SentenTree (2016)^[21]提供了一种新的可视化方式,针对社交网络文本数据集,通过模式生成算法获取模式列表构建树形结构,在

此基础上对分割的单词使用力导向布局算法调整,根据单词顺序和方向进行约束,如图 3展示的是twitter上2014年世界杯的第一场比赛的第一个进球的数据,这种可视化方式保留了基本的句法结构,对社交网络文本的快速概览的同时通过交互能够实现细节层面的探索。



图3 2014年世界杯第一个进球的twitter数据SentenTree

除了关键词总结性的文本内容外,通过 文本的特征还可以探索发现内容的分布和演 化。Literature Fingerprinting^[22]提出了文献指 纹方法帮助用户了解某一特征在全文中的分 布规律,将特征在整个文本中的分布用一系 列像素图表达,如图4左所示,句子的平均长 度是文本度量特征,可以辨识作者的写作风格。科罗拉多大学的Szafir(2016)^[23]提出了文本基因的方法,通过对上百万大规模文档集的分析,比较词汇的分布差异,研究了过去350年英语文档中常用词汇的变化,如图4右所示。



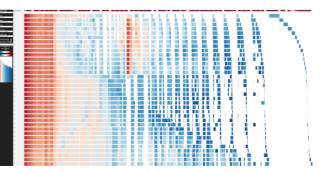


图4 LiteratureFingerprinting(左)与TextDNA(右)

3.2 基于文本主题的应用

情报分析的过程中,研究人员需要在大量信息中寻找所隐含的知识,如学科的研究热点、趋势、在相关学科之间的演变,以及网络舆论热点的产生、传播和消亡等,这些信息需要深入到文本的语义层面进行挖掘。文本中所包含的主要语义信息可以用主题来概括,在此基础上通过可视化可以研究文本之中主题的分布、主题之间的关系以及随时间变化的趋势。主题抽取一般基于主题模型的方法,包括向量空间模型和统计语言模型。

主题河流(Themeriver)是一种经典的展现文本集合中主题演化的可视化方法,它采用河流作为可视化原语来编码文档集合中的主题

信息,将主题隐喻为时间上不断延续的河流,每一层代表一个主题,河流的宽度代表当前主题的热度,这种方法提供了一个宏观的主题演化视角,帮助用户观察主题的产生、发展和衰落^[24]。主题内容通常使用具体的关键词或概念来表示,和标签云结合是最典型的展示方式。微软亚洲研究院的刘世霞团队和IBM的崔为炜团队在2009年和2011年分别开发了TIARA^[25]和TextFlow^[26]系统,通过无缝集成的LDA模型抽取文本主题将其展现在河流图中,并将每个主题下的关键词显示在每条线条中,提供主题变化的情况。在此基础上,ThemeDelta^[27]使用桑基图取代河流图展示主题的流动,使主题变化更加直观明了。

doi:10.3772/j.issn.2095-915x.2017.04.007

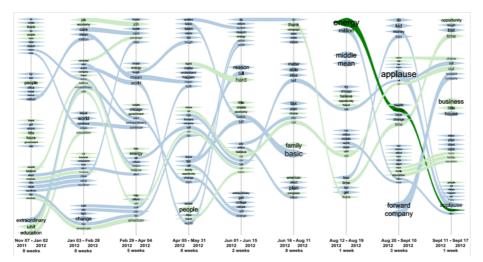


图5 2012美国大选奥巴马与罗姆尼演讲主题的ThemeDelta可视化

对于情报研究人员来说,在社交网络上追踪、分析大型事件中观点的爆发、传播是重要的研究课题,可视化提供了很好的研究方法。Xu(2013)[28]将社交网站上的群体分为媒体(Media),草根(Grassroots),政治人物(Political Figures),通过时间线视图反映不同主体之间的竞争以及在竞争过程中意见领袖所扮演的角色,如图6所示,借用主题河流图的形式来展示每个主题的发展和衰落,河流宽度表示该主题在对应时刻的热度。不同类

型的意见领袖用不同颜色的线,线条穿插在不同的河流中,隐喻在对应的主题中产生了一定的影响力。线条的分裂或合并表示意见领袖们的关注点在不同话题的分裂与合并,不同主题之间竞争力的转换采用径向视图来展示。为了分析具体的语义内容,系统还提供了词云,并且在词云的圆周边界上,用户可以对主题和意见领袖进行筛选。EvoRiver(2014)[29]和IdeaFlow(2015)[30]也采用了类似的设计来反映社交网络上的主题变化。

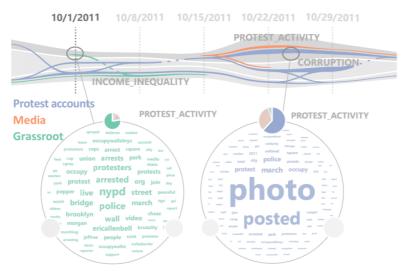


图6 "占领华尔街"运动期间twitter网站上主题竞争可视化

使用主题河流图和桑基图能够直观的看出 文本中主题的变化,但是缺乏主题之间关系全 局的认识。TopicPanorama^[31]围绕主题网络关 系图构建了一个径向填充的圆形视图,来展示 来自不同信息源中主题之间的全局关系。Jiang (2015)^[32]采用了桑基图、标签云和统计图 表,对TOG、TVCG、TKDE三个期刊的研究主题进行了可视化研究,以三个期刊的研究主题代表各自领域的发展和趋势,在这些主题之间建立一个层次和进化关系,将主题的演变映射为桑基图的流向,散点图显示主题的MDS投影,与流图结合起来展示主题之间的相似性。

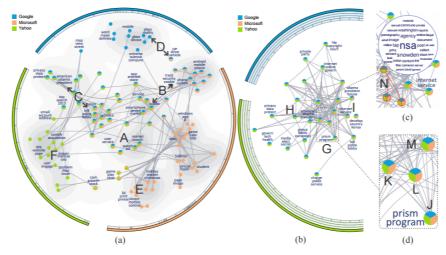


图7 TopicPanorama可视化

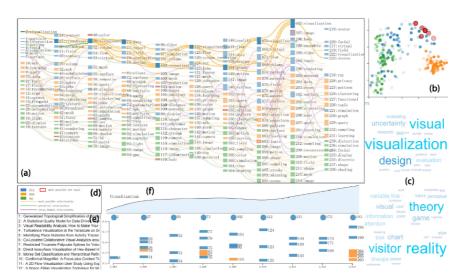


图8 三个跨学科领域的跨域主题可视化

3.3 基于文本情感的应用

文本内容信息中蕴含了描述人类主观喜 好、赞赏、感觉的情感信息,文本的情感分 析(又称意见挖掘)可视化是将文本抽象量 化的过程,通过可视化把量化的数据转换为 直观的感受。情感分析一般通过文本挖掘技 术提取文本中的带有情感色彩的主观性信 息,按照正向和负向不同的权重进行量化分

doi:10.3772/j.issn.2095-915x.2017.04.007

析、处理、归纳和推理,能够体现主观评价和意见倾向^[33],适用于舆情监测、市场调研、信息预测等等。

社交网站为公众提供一个多元开放的舆论环境,其中的用户可以充分表达自己的观点意见,也是情报分析人员重要的信息源。Oelke(2009)[34]采用矩阵热力图可视化顾客对商品

的反馈信息,以颜色作为视觉特征表示顾客对商品的喜好程度。PEARL(2014)[35]是针对个人的社交网络多维情感分析工具,基于个人的社交网络数据,通过文本监测不同时间个人的情绪波动,总结个人的情感风格,在情绪类型、强度、主题多个层次进行可视化展示,并支持通过交互探索分析一个人的情绪变化。

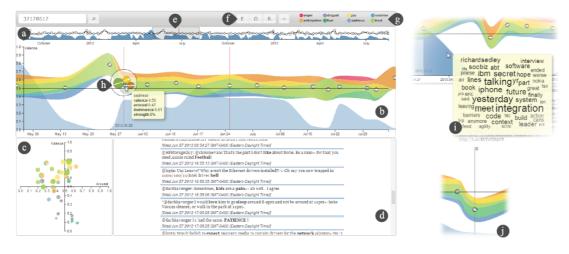


图9 PEARL支持多种交互,提供情绪波动曲线、河流图、推文标签云、情绪象限等视图

Assady (2016) [36]引入了一种动态的可视 化分析方法来分析多方会谈中各方的行为模 式,统计每个发言人的语言,抽象相关的文本 特征,如图10所示,采用动画的视觉隐喻探索 随时间推移每个发言人的观点变化,以协助政 治研究人员探索对话的动态。

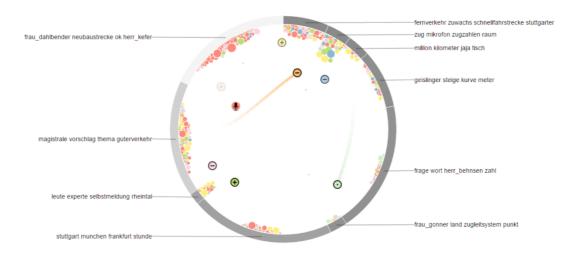


图10 著名的"斯图加特21项目"全民公投各方立场动态可视化

3.4 基于文本关系的应用

情报分析的核心是对知识的发现和洞察, 无论是单个文档还是海量的文本数据集,文本 之间的关系都广受关注。从可视化的角度观察 文本之间的关系能够发现以往的方法难以发现 的隐藏信息。通常文本关系可视化的内容包括 文本间的引用关系、网页的超链接关系等直接 关系,以及主题相似性等潜在关系。

在图情领域, 多数探索文本间关系的可视

化研究都以网络图作为呈现形式。对科学文献的组织、分析、探索是情报研究的重要任务,在传统的文献计量学中,通过网络图谱进行共词分析、共引分析、耦合分析是典型的分析方法,学术界开发了UciNet、Pajek、Gephi、Citespace^[37]等等一系列的可视化分析软件,图11为利用Citespace制作的大数据相关文献的期刊引用叠加图^[38],从中可以清晰地看出大数据科学研究的主要研究领域、知识的流动方向等。

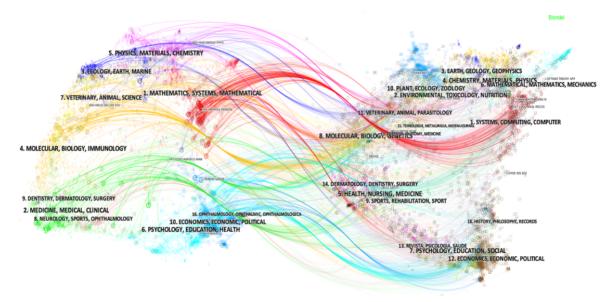


图11 大数据相关文献的期刊引用叠加图

PivotPaths^[39], PaperVis^[40], TopicNets^[41] 通过人机交互的方式让用户自由地探索文本间的关系,突破单一视角从多个角度观察,更自然、高效、符合人的认知心理。ViaRoma^[42] 集时空分析与文本分析为一体,通过展示文本的层次关系和主题,能够让用户快速了解罗马历史进程中的事件关系。CiteRivers^[43]是一个非常完善的集成应用,结合其他学者的研究,扩展新的分析方法,设计了全新的对文献的内容、属性以及引用模式的可视化分析系统。如

图12所示,通过河流图的方法来展示文献,根据文献的属性或者相似度,文献被划分到某个聚类(流)中;每个聚类(流)又由不同时间的块组成,河流的宽度则代表文献的数量;每个块中的论文抽取出相应的关键词云,在引用属性面板中展示统计文献的引用年龄和引用熵。面板下方像素图对应每个主题中统计出的最高产的10位作者,并展现其中的变化,这一设计可以同步观察文献和作者的关联变化。

doi:10.3772/j.issn.2095-915x.2017.04.007

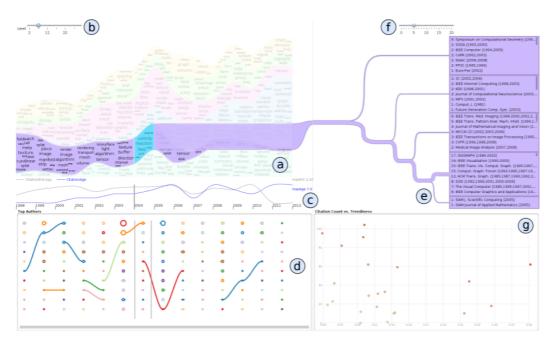


图12 可视化领域1998~2011年文献的CiteRivers可视化

4 总结

综上,基于文本概览的可视化具有对文本 内容进行高效的分析处理的优点,能够直观呈 现文本的有效信息,但也损失了语义层面上大 量信息;基于文本主题的可视化,优点是能够 动态观察主题的演化,然而主题抽取的有效性 和可信度有待考量;基于文本情感的可视化可 用于网络舆情监控、商业调查,最大的优点是 把抽象的情感转化为视觉的特征,但是这种映 射方式是否符合人的直觉需要可视化理论的进 一步的支持;基于文本关系的可视化能够揭示 文本中蕴含的复杂关系,不过数据关系代表的 内在本质依旧需要专业研究人员的解读,而且 过于复杂的可视化结果也造成了认知的困难。

通过以上案例分析可以看出可视化在处理 复杂信息时兼具可读性和解释性,可视化方法 和情报研究人员的专业分析相结合能产生很大 的价值,在情报分析工作中具有很强的应用性。目前的可视化软件尚不能满足大多数普通情报分析人员的需要,研究人员应加强情报分析可视化应用的探索和实践,借鉴可视化领域的技术和方法支撑情报分析任务。

将情报研究的需求与可视化有效结合,应 以用户为中心的进行系统设计与开发,笔者认 为可分为三个方面:首先,情报研究工作的数 据来源复杂,将非结构化的异构数据转换为结 构化数据是分析基础,基础的文本处理工具可 以满足挖掘文本信息的初步需求;其次,面向 研究人员开发具有通用功能的界面友好、交互 性强的综合性文本可视化分析工具,降低情报 获取、组织和分析的复杂程度,提高工作的效 率和智能性;最后,对于从事专业情报分析工 作的研究人员来说,模块化、功能化的应用, 更有利于针对不同的任务敏捷开发适用性强的 专业分析工具,以深入地研究复杂性、抽象性 的信息。

文本的可视化尚存在较多问题有待进一步 研究,但是可以预见,随着可视化技术趋于成 熟和智能,可视化的方法将更紧密地融入情报 的研究工作中。

参考文献

- [1] 曾建勋, 魏来. 大数据时代的情报学变革[J]. 情报学报, 2015(1): 37-44.
- [2] 唐家渝, 刘知远, 孙茂松. 文本可视化研究综述[J]. 计算机辅助设计与图形学学报, 2013, 25(3): 273-285.
- [3] Isenberg P, Isenberg T, SedImair M, et al. Visualization as Seen through its Research Paper Keywords[J]. IEEE Transactions on Visualization and Computer Graphics, 2017, 23(1): 1-1.
- [4] 张进, 陈远. 论情报检索可视化过程中信息节点的歧义性问题[J]. 情报学报, 1998(3): 175-179.
- [5] 周宁, 文燕平, 刘玮. 文献信息可视化研究[J]. 情报学报, 2003, 22(4): 468-471.
- [6] 陈悦, 刘则渊. 悄然兴起的科学知识图谱[J]. 科学学研究, 2005, 23(2): 149-154.
- [7] 邱均平, 李小涛, 董克. 图情领域可视化研究的发展、 演化与创新[J]. 图书情报工作, 2014(13): 125-131.
- [8] 赵蓉英, 郭月培. 近10年我国信息可视化研究的计量分析[J]. 情报科学, 2014(5): 3-6.
- [9] 詹建, 高民权. 基于主题河的网络舆情可视化关联分析方法[J]. 情报资料工作, 2014, 35(6): 17-22.
- [10] 赵一鸣, 张进, 黎苑楚. 基于多维尺度模型的潜在主题可视化研究[J]. 情报学报, 2014, 33(1): 45-54.

- [11] 孙国超, 徐硕, 乔晓东. AToT模型可视化工具开发 [J]. 情报工程, 2016, 2(4): 20-29.
- [12] Sacha D, Stoffel A, Stoffel F, et al. Knowledge Generation Model for Visual Analytics[J]. IEEE Transactions on Visualization & Computer Graphics, 2014, 20(12): 1604-1613.
- [13] Kaser O, Lemire D. Tag-Cloud Drawing: Algorithms for Cloud Visualization[J]. Computer Science, 2007.
- [14] Viegas F B, Wattenberg M, Feinberg J. Participatory Visualization with Wordle[J]. IEEE Transactions on Visualization & Computer Graphics, 2009, 15(6): 1137-44.
- [15] Buchin K, Creemers D, Lazzarotto A, et al. "Geo Word Clouds." 2016 IEEE Pacific Visualization Symposium (PacificVis). IEEE, 2016: 144-151
- [16] Cui W, Wu Y, Liu S, et al. Context-preserving, Dynamic Word Cloud Visualization.[J]. Computer Graphics & Applications IEEE, 2010, 30(6): 42-53.
- [17] Chi M T, Lin S S, Chen S Y, et al. Morphable Word Clouds for Time-Varying Text Data Visualization[J]. IEEE Transactions on Visualization & Computer Graphics, 2015, 21(12): 1415-26.
- [18] Burch M, Lohmann S, Beck F, et al. RadCloud: Visualizing Multiple Texts with Merged Word Clouds[C]// International Conference on Information Visualisation. IEEE, 2014: 108-113.
- [19] Wattenberg M, Viégas F B. The Word Tree, an Interactive Visual Concordance[J]. IEEE Transactions on Visualization & Computer Graphics, 2008, 14(6): 1221.
- [20] Van Ham F, Wattenberg M, Viégas F B. Mapping Text with Phrase Nets[J]. IEEE Transactions on Visualization & Computer Graphics, 2009, 15(6):

探索与研究

DISCOVERY AND RESEARCH

doi:10.3772/j.issn.2095-915x.2017.04.007

1169-76.

[21] Hu M, Wongsuphasawat K, Stasko J. Visualizing Social Media Content with SentenTree[J]. IEEE Transactions on Visualization & Computer Graphics, 2017(1): 621-630.

[22] Keim D A, Oelke D. Literature Fingerprinting: A New Method for Visual Literary Analysis[C]// Visual Analytics Science and Technology, 2007. VAST 2007. IEEE Symposium on. IEEE, 2008: 115-122.

[23] Szafir D A, Stuffer D, Sohail Y, et al. TextDNA: Visualizing Word Usage with Configurable Colorfields[J]. 2016, 35(3): 421-430.

[24] 巫英才, 崔为炜, 宋阳秋,等. 基于主题的文本可视分析研究[J]. 计算机辅助设计与图形学学报, 2012, 24(10): 1266-1272.

[25] Liu S, Zhou M X, Pan S, et al. Interactive, Topic-based Visual Text Summarization and Analysis[J]. ACM Transactions on Intelligent Systems & Technology, 2009, 3(2): 543-552.

[26] Cui W, Liu S, Tan L, et al. TextFlow: Towards Better Understanding of Evolving Topics in Text[J]. IEEE Transactions on Visualization & Computer Graphics, 2011, 17(12): 2412-21.

[27] Gad S, Javed W, Ghani S, et al. ThemeDelta: Dynamic Segmentations over Temporal Topic Models[J]. IEEE Transactions on Visualization & Computer Graphics, 2015, 21(5): 672-85.

[28] Xu P, Wu Y, Wei E, et al. Visual Analysis of Topic Competition on Social Media[J]. Visualization & Computer Graphics IEEE Transactions on, 2013, 19(12): 2012-2021.

[29] Sun G, Wu Y, Liu S, et al. EvoRiver: Visual Analysis of Topic Coopetition on Social Media[J]. IEEE

Transactions on Visualization & Computer Graphics, 2014, 20(12): 1753-62.

[30] Zhong Y, Liu S, Wang X, et al. Tracking Idea Flows between Social Groups[J]. Computer Science, 2015.

[31] Liu S, Wang X, Chen J, et al. TopicPanorama: A Full Picture of Relevant Topics[C]// Visual Analytics Science and Technology. IEEE, 2015: 183-192.

[32] Jiang X, Zhang J. A Text Visualization Method for Cross-domain Research Topic Mining[J]. Journal of Visualization, 2016, 19(3): 1-16.

[33] 赵妍妍, 秦兵, 刘挺. 文本情感分析[J]. 软件学报, 2010, 21(8): 1834-1848.

[34] Oelke D, Hao M, Rohrdantz C, et al. Visual Opinion Analysis of Customer Feedback Data[C]// Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on. IEEE, 2009: 187-194.

[35] Zhao J, Gou L, Wang F, et al. PEARL: An Interactive Visual Analytic Tool for Understanding Personal Emotion Style Derived from Social Media[C]// Visual Analytics Science and Technology. IEEE, 2015: 203-212.

[36] El-Assady M, Gold V, Acevedo C, et al. ConToVi: Multi-Party Conversation Exploration using Topic-Space Views[J].Computer Graphics Forum. 2016, 35(3): 431-440.

[37] Chen C, Ibekwe - Sanjuan F, Hou J. The Structure and Dynamics of Cocitation Clusters: A Multiple - Perspective Cocitation Analysis[J]. Journal of the Association for Information Science and Technology, 2010, 61(7): 1386-1409.

[38] 李杰. Citespace中文指南[EB/OL]. [2016-11-25]. http://cluster.ischool.drexel.edu/~cchen/citespace/

manual/CiteSpaceChinese.pdf.

[39] Dörk M, Riche N H, Ramos G, et al. PivotPaths: Strolling through Faceted Information Spaces[J]. IEEE Transactions on Visualization & Computer Graphics, 2012, 18(12): 2709-18.

[40] Chou J K, Yang C K. PaperVis: Literature Review Made Easy[J]. Computer Graphics Forum, 2011, 30(3): 721-730.

[41] Gretarsson B, Donovan J, Bostandjiev S, et al. TopicNets: Visual Analysis of Large Text Corpora with

Topic Modeling[J]. ACM Transactions on Intelligent Systems & Technology, 2012, 3(2): 565-582.

[42] Cho I, Dou W, Wang D X, et al. VAiRoma: A Visual Analytics System for Making Sense of Places, Times, and Events in Roman History[J]. IEEE Transactions on Visualization & Computer Graphics, 2016, 22(1): 210-219.

[43] Heimerl F, Han Q, Koch S. CiteRivers: Visual Analytics of Citation Patterns[J]. Visualization & Computer Graphics IEEE Transactions on, 2016, 22(1): 190-199.