**武汉大学**

**硕/博士研究生学位论文开题报告**

论 文 题 目:**泛在网络空间数据可视化研究**

研 究 生 姓 名 李冬琳

学 号 2014282050150

导 师 姓 名 李精忠

系 所 资源与环境科学学院

专 业 测绘工程

研 究 方 向 空间数据可视化

入 学 时 间 2014年9月

毕 业 时 间 2016年7月

注：本表可复印，可另加附页

**一、立论依据**

|  |
| --- |
| **（论文的研究意义、国内外研究现状分析、附主要参考文献）**  **（一）研究意义**：  随着信息技术的发展，传感网、互联网、通信网、行业网等网络系统构成的泛在网络产生了海量的带有地理空间信息的数据，泛在网络又极大地丰富了人类获取信息的手段和内容， 使任何人在任何地方都能通过泛在网络获取需要的任何信息， 称为泛在信息。而这些海量数据背后的信息才是真正的价值，据统计，泛在信息中80%带有空间信息，所以挖掘泛在网络环境下的空间信息是一项有重大意义的研究课题。当前数据挖掘中，可视分析是一项重要的分析手段，可视分析结合了可视化、人机交互和自动分析，并使数据分析过程透明化。用户通过人机交互技术评价、修改和改进自动分析模型， 从而得到新的自动分析结果。由人来定义分析任务和识别复杂的模式， 由机器来存储和分析大量的数据，这样就让人和机器实现优势互补，能够非常有效直观的挖掘出数据背后的价值，发现数据背后隐藏的规律。  泛在网络空间数据，  **（二）国内外研究现状**：  **国外研究现状**：  国外对可视化的研究比国内起步早，成果丰富。Thom等人利用Twitter用户发布的信息用点标注的方法来实现可视化，并利用算法将局部信息和全球信息进行对比来探测异常事件。Marcus等人通过算法对Twitter数据进行聚合，实现了交互式事件可视化，帮助用户梳理整个事件的走向。Zhao等人将文本的叙述结构语义以树的形式进行可视化，同时展现了相似度统计、修辞结构、以及相应的文本内容。Krüger等人开发的轨迹可视化系统允许用户交互式定义轨迹的起止点，实现轨迹可视化。Zeng等人计算了乘客从一个公交或地铁站点出发到达其他站点的时间花费，并实现了多种公共交通状况专题的可视化形式，比如交通可达性等专题。Tominski等人利用时空立方体技术，将轨迹折线扩展成彩色条带, 并在高度方向将不同轨迹的条带堆叠起来, 以方便轨迹间的比较，不同的条带代表不同的属性，实现了轨迹的多属性三维可视化。Buchin等人利用边捆绑技术，用Flow Map对美国人口移动事件进行时空可视化。Slingsby等人利用多维平行坐标的方法，多具有多指标多维度特性的的人口统计数据进行了可视化，能非常直观的发现和了解人口统计数据中各组指标之间的相互关系。  **国内研究现状：**  随着对大数据研究的深入，国内学者对可视化的研究也越来越多。周霞娟等人开发了一个用户驱动的微博可视化搜索系统，通过对用户、行为特征等建立模型，实现了更加高效的微博搜索。袁晓如团队开发了一套微博事件可视化系统，能够对微博事件的传播链路进行网络可视化。陈思明等人开发了一套交互式微博可视化系统，针对抽样微博用户发布微博的位置移动，挖掘用户的出行模式。赵华等人通过分析用户发表微博特点，利用3D标签云实现用户兴趣可视化。王松等人开发了一套微博可视化系统，能对微博热点、发布数量、地理信息、时序等信息进行可视化。武泽旭利用融合FN的多层级图可视化算法对微博数据进行网络可视化。陈宏飞等人基于微博进行了西安市夜间时空分布可视化、交通拥堵状况时空分布研究。禹文豪等人基于网络核密度法对网络空间POI点进行三维可视化分析。杨敏等人基于服务器端多层次结构化组织与客户端冲突移位相结合的在线式综合策略，实现了POI点的多尺度可视化。袁晓如团队利用OD-Wheel方法对城市内的出租车轨迹数据进行可视化，可以从中发现出行模式和异常交通拥堵。陈为团队开发了一个交互式可视化系统，能基于出租车轨迹数据实时对城市道路交通流量、道路交叉口交通状况进行模拟可视分析。  **（三）主要参考文献**：  [1] Thom D, Bosch H, Koch S, et al. Spatiotemporal anomaly detection through visual analysis of geolocated twitter messages[C]//Pacific visualization symposium (PacificVis), 2012 IEEE. IEEE, 2012: 41-48.  [2] Marcus A, Bernstein M S, Badar O, et al. Twitinfo: aggregating and visualizing microblogs for event exploration[C]//Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 2011: 227-236.  [3] Zhao J, Chevalier F, Collins C, et al. Facilitating discourse analysis with interactive visualization[J]. Visualization and Computer Graphics, IEEE Transactions on, 2012, 18(12): 2639-2648.  [4] Krüger R, Thom D, Wörner M, et al. TrajectoryLenses–A Set‐based Filtering and Exploration Technique for Long‐term Trajectory Data[C]//Computer Graphics Forum. Blackwell Publishing Ltd, 2013, 32(3pt4): 451-460.  [5] Zeng W, Fu C W, Arisona S M, et al. Visualizing mobility of public transportation system[J]. Visualization and Computer Graphics, IEEE Transactions on, 2014, 20(12): 1833-1842.  [6] Tominski C, Schumann H, Andrienko G, et al. Stacking-based visualization of trajectory attribute data[J]. Visualization and Computer Graphics, IEEE Transactions on, 2012, 18(12): 2565-2574.  [7] Buchin K, Speckmann B, Verbeek K. Flow map layout via spiral trees[J]. Visualization and Computer Graphics, IEEE Transactions on, 2011, 17(12): 2536-2544.  [8] Slingsby A, Dykes J, Wood J. Exploring uncertainty in geodemographics with interactive graphics[J]. Visualization and Computer Graphics, IEEE Transactions on, 2011, 17(12): 2545-2554.  [9] Donghao Ren, Xin Zhang, Zhenhuang Wang, Jing Li, and Xiaoru Yuan. *Proceedings of IEEE Pacific Visualization Symposium (PacificVis 2014),*pages 330-334, (Notes Paper), Yokohama, Japan, Mar. 4-7, 2014.  [10] Ren D, Zhang X, Wang Z, et al. WeiboEvents: a crowd sourcing weibo visual analytic system[C]//Pacific Visualization Symposium (PacificVis), 2014 IEEE. IEEE, 2014: 330-334.  [11] Lu M, Wang Z, Liang J, et al. OD-Wheel: Visual Design to Explore OD Patterns of a Central Region[J].  [12] Chen S, Yuan X, Wang Z, et al. Interactive Visual Discovering of Movement Patterns from Sparsely Sampled Geo-tagged Social Media Data[J]. Visualization and Computer Graphics, IEEE Transactions on, 2016, 22(1): 270-279.  [13] Wang F, Chen W, Wu F, et al. A visual reasoning approach for data-driven transport assessment on urban roads[C]//Visual Analytics Science and Technology (VAST), 2014 IEEE Conference on. IEEE, 2014: 103-112.  [14] 张昕,袁晓如. 树图可视化[J]. 计算机辅助设计与图形学学报,2012,09:1113-1124.  [15]袁晓如,张昕,肖何,郭翰琦. 可视化研究前沿及展望[J]. 科研信息化技术与应用,2011,04:3-13.  [16]袁晓如. 可视化研究与应用[J]. 国际学术动态,2010,04:40-42.  [17]王祖超,袁晓如.轨迹数据可视分析研究[J].计算机辅助设计与图形学学报,2015,01:9-25.  [18]原达,刘日晨,袁晓如. 探地数据可视化研究[J].计算机辅助设计与图形学学报,2015,01:36-45.  [19]周琳,孔雷,赵方庆. 生物大数据可视化的现状及挑战[J]. 科学通报,2015,Z1:547-557.  [20]戴国忠,陈为,洪文学,刘世霞,屈华民,袁晓如,张加万,张康. 信息可视化和可视分析:挑战与机遇——北戴河信息可视化战略研讨会总结报告[J]. 中国科学:信息科学,2013,01:178-184.  [21]彭丰林,陈会忠,张健,沈萍,邹自明,薛国强,袁晓如,马麦宁. 地球物理信息学概念的发展[A]. 中国地球物理学会信息技术专业委员会.大数据、云计算与地球物理应用研讨活动论文摘要集[C].中国地球物理学会信息技术专业委员会:,2014:2.  [22]袁晓如. 大数据可视分析——挑战和机遇[A]. 中国地球物理学会信息技术专业委员会.大数据、云计算与地球物理应用研讨活动论文摘要集[C].中国地球物理学会信息技术专业委员会:,2014:1.  [23]彭丰林,陈会忠,张健,沈萍,邹自明,薛国强,袁晓如,马麦宁. 地球物理信息学概念的发展[A]. 中国地球物理学会、全国岩石学与地球动力学研讨会组委会、中国地质学会构造地质学与地球动力学专业委员会、中国地质学会区域地质与成矿专业委员会.2014年中国地球科学联合学术年会——专题16：地球物理信息学的发展及应用论文集[C].中国地球物理学会、全国岩石学与地球动力学研讨会组委会、中国地质学会构造地质学与地球动力学专业委员会、中国地质学会区域地质与成矿专业委员会:,2014:1.  [24]王丹,彭丰林,袁晓如,邢利军,沈晓阳. 利用Matlab与Google Earth的地磁场模型网络可视化[A]. 中国地球物理学会.中国地球物理学会第二十四届年会论文集[C].中国地球物理学会:,2008:1.  [25]禹文豪,艾廷华. 核密度估计法支持下的网络空间POI点可视化与分析[J]. 测绘学报,2015,01:82-90.  [26]王爽,李炯. 基于城市网络空间的POI分布密度分析及可视化[J]. 城市勘测,2015,01:21-25.  [27]陈宏飞,李君轶,秦超,刘广,孙九林. 基于微博的西安市居民夜间活动时空分布研究[J]. 人文地理,2015,03:57-63.  [28]赵华,纪晓文,曾庆田,郝春燕. 基于话题相关空间的微博用户兴趣识别及可视化方法[J]. 计算机科学,2015,S1:500-502+509.  [29]周霞娟,汪飞,金玲,陈为,王章野. 用户驱动的微博可视化搜索[J]. 中国图象图形学报,2015,05:715-723.  [30]陈宏飞,张心萍,赵艳慧,刘广,孙九林. 基于微博的西安市交通拥堵状况时空分布研究[J]. 陕西师范大学学报(自然科学版),2015,06:83-88.  [31]王松,吴亚东,李秋生,蒋宏宇,邹勇刚. 基于时空分析的微博演化可视化[J]. 西南科技大学学报,2014,03:68-75.  [32]刘经南,方媛,郭迟,高柯夫. 位置大数据的分析处理研究进展[J]. 武汉大学学报(信息科学版),2014,04:379-385.  [33]李清泉,李德仁. 大数据GIS[J]. 武汉大学学报(信息科学版),2014,06:641-644+666.  [34]朱欣焰,周成虎,呙维,胡涛,刘洪强,高文秀. 全息位置地图概念内涵及其关键技术初探[J]. 武汉大学学报(信息科学版),2015,03:285-295.  [35]张莉婷,陈云波,左小清,王伟鹏. 出租车轨迹数据快速提取道路骨架线方法[J]. 测绘科学,2015,01:110-112+125.  [36]周成虎,朱欣焰,王蒙,施闯,欧阳. 全息位置地图研究[J]. 地理科学进展,2011,11:1331-1335.  [37]李德仁,王树良,李德毅,王新洲. 论空间数据挖掘和知识发现的理论与方法[J]. 武汉大学学报(信息科学版),2002,03:221-233.  [38]李德仁,王树良,史文中,王新洲. 论空间数据挖掘和知识发现[J]. 武汉大学学报(信息科学版),2001,06:491-499.  [39]王新洲. 论空间数据处理与空间数据挖掘[J]. 武汉大学学报(信息科学版),2006,01:1-4+8.  [40]郑宇. 城市计算概述[J]. 武汉大学学报(信息科学版),2015,01:1-13.  [41]王铭军,潘巧明,刘真,陈为. 可视数据清洗综述[J]. 中国图象图形学报,2015,04:468-482.  **著作类**  1、《数据可视化的基本原理与方法》陈为等[M]．北京：科学出版社，2013.  2、《空间数据挖掘理论与应用》李德仁等[M] 北京：科学出版社，2013.  3、涂子沛，大数据： 正在到来的数据革命[M]．桂 林：广西师范大学出版社，2013  4、维克托·迈尔－舍恩伯格，肯尼思·库克耶．大数据时代[M] ．盛杨燕，周涛．杭州： 浙江人民出版社，2013  5、《大数据时代遇见未来的新思维》 艾伯特等[M],北京：中国人民大学出版社，2012 |
|  |

**二、研究方案**

|  |
| --- |
| **1．研究目标、研究内容和拟解决的关键问题**  **研究目标：**  本文预期在现有学者研究基础上，针对泛在网络空间数据中的微博数据、POI数据、出租车轨迹数据提出几种具有创新性的的可视化方法、挖掘出这些数据背后的价值。并且能将这些方法实现并且集成为一个在线可视化系统，能提供给普通用户及研究者使用。  **研究内容：**  本文的研究内容主要包括以下几个方面：第一、针对微博数据，将标签云与和聚类可视化两种方法想结合，实现一种新的微博可视化方法，能对微博文本话题事件时空可视化，挖掘出用户的兴趣及话题的生命周期与地理空间的关系；第二、针对微博数据，针对单一话题下，不同区域的热度三维可视化；第三、针对POI数据，将单调的点标注、聚类、热图等方法相结合、实现优势互补，利用交互技术，实现POI的多尺度可视化；第四、针对出租车轨迹数据，对估计进行聚合、分类分级。实现顾及底图的自适应多尺度可视化。  **拟解决的关键问题：**   1. 微博数据的持续爬取与高速读取技术 2. 针对微博文本数据的文本分词、关键字提取方法 3. 话题热度的三维可视化技术研究 4. 针对POI点数据，便捷的能融合其他三种可视化技术的交互技术研究 5. 针对出租车轨迹数据，数据的融合分类分级方法设计 6. 轨迹数据的动态可视化技术研究   **2．拟采取的研究方法、技术路线、实验方案及可行性分析**  **本文所采用的研究方法如下：**  数据采集  数据处理  数据可视化  可视分析  得出结论  **技术路线：**  本文拟采取的技术路线是:   1. 开发高效的网络爬虫，持续的获取泛在网络空间数据比如微博数据、POI数据、出租车轨迹数据等。 2. 对爬虫获取的数据进行清洗、预处理，拟采用非关系型数据库MongoDB对这些数据进行存储。 3. 拟采用Nodejs语言开发便捷高效的的网络数据接口层，满足在进行可视化时对定制化数据的需求。 4. 采用openlayer3、angularjs、echarts等JavaScript语言库，对数据进行可视化。 5. 在可视化的过程中、完成时对可视化的成果进行分析、得出结论。 |
| **3．本论文的特色与创新之处**  第一，本论文的特色在于将在针对大数据时代下，研究泛在网络空间数据的可视化方法，将针对大规模复杂异构的数据进行可视化表达，这其中将要克服复杂空间数据的获取、清洗、存储等技术难题，同时针对空间数据，需要设计出合适的表达方法，对空间数据进行可视分析，进行空间数据挖掘，得到数据背后的价值。  第二，本论文的创新之处在于将创新性的将微博文本信息和地理位置信息相结合进行可视化，在地理空间的纬度下观察微博事件的生命周期及其与地理空间的关系。针对微博事件，将对不同地理区域内的热度进行三维可视化。针对POI数据，将打破传统的可视化手段，利用互联网技术，实现高交互性的可视化作品，利用交互来实现多尺度、多维度下的POI点可视化。针对出租车轨迹数据，也将创新性的进行动态多尺度可视化。  **4．预期的论文进展和成果**  2015年6月——11月：相关资料的收集整理和阅读阶段  2015年12月——2016年1月：撰写论文的初稿  2016年2月——3月中旬：在导师的指导下完成论文初稿的修订和完善  2016年4月初——2016年4月底：论文的再次修改和完善  2016年5月：论文定稿 |

**三、论文大纲**

|  |
| --- |
| **泛在网络空间数据可视化方法研究**  **第一章 绪论**  **1.1． 研究目的与意义**  **1.2. 国内外研究现状**  **1.3. 本论文主要研究内容和方法**  **第二章 泛在网络、可视化技术、理论**  **2.1. 泛在网络**  **2.2. 可视分析**  **2.3. 当前主流可视化技术**  **第三章 泛在网络空间数据抓取、清洗**  **3.1. 泛在网络空间数据的特定**  **3.2. 泛在网络空间数据的抓取技术**  **3.3. 当前主流大数据清洗技术**  **3.4. 本论文采用的数据清洗技术**  **第四章 微博语义时空可视化**  **4.1. 微博语义数据的特点**  **4.2. 自然语言处理技术**  **4.3．事件时空可视化**  **4.4. 事件热度三维可视化**  **第五章 POI自适应可视化**  **5.1 POI点数据特征、可视化需求**  **5.3 自适应可视化理论与技术**  **5.2 POI点自适应可视化**  **第六章 出租车轨迹数据**  **6.1 出租车轨迹数据的预处理**  **6.2 出租车轨迹综合分类、分级**  **6.3 出租车轨迹与底图相适应**  **6.3 出租车轨迹动态可视化**  **第七章 总结与展望**  **参考文献**  **致谢** |

**四、研究基础**

|  |
| --- |
|  |

**五、导师或指导小组意见**

|  |
| --- |
| **导师签名：**  **年 月 日** |