

多模态交通大数据融合与可视化

董晔卉, 涂倩怡*, 罗建书

湖南交通工程学院电气与信息工程学院, 湖南 衡阳

收稿日期: 2024年10月25日; 录用日期: 2024年11月21日; 发布日期: 2024年11月29日

摘要

为缓解当前城市交通数据的复杂性及利用率低下的问题, 本文设计一套适用于中国现代化城市的交通大数据可视化系统。本文遵循前后端分层设计的原则, 前端设计了直观易用的可视化界面, 包括交通大数据中心、交通事件管理等模块的交互展示; 后端则运用Python技术, 实现对路网结构、交通设施、交通事件等多元数据的精细化管理和高效处理, 从而实现了时空相关性理论研究成果的直观呈现。此系统的开发, 将原本繁杂的交通数据以清晰直观的数据统计图表形式展示给决策者, 为城市交通拥堵治理和短时流量预测提供了有力的数据支撑, 极大地提高了城市管理者的决策效率和准确性。

关键词

数据融合, 交通大数据, 数据可视化, Python可视化

Multimodal Traffic Big Data Fusion and Visualization Technology

Yehui Dong, Qianyi Tu*, Jianshu Luo

School of Electrical and Information Engineering, Hunan Institute of Transportation Engineering, Hengyang Hunan

Received: Oct. 25th, 2024; accepted: Nov. 21st, 2024; published: Nov. 29th, 2024

Abstract

To alleviate the complexity of the current urban traffic data and low utilization rate of the problem, this paper designs a set of applicable in China's modern city traffic data visualization system. In this paper, follow the principle of front and back side hierarchical design, front-end design the visual interface intuitive and easy to use, including traffic large data centers, traffic incident management

*通讯作者。

文章引用: 董晔卉, 涂倩怡, 罗建书. 多模态交通大数据融合与可视化[J]. 计算机科学与应用, 2024, 14(11): 182-190.
DOI: 10.12677/csa.2024.1411227

modules, such as interactive display; At the back end, Python technology is used to realize the fine management and efficient processing of multivariate data such as road network structure, traffic facilities, and traffic events, so as to realize the intuitive presentation of the theoretical research results of spatio-temporal correlation. The development of this system presents the originally complicated traffic data to the decision-makers in the form of clear and intuitive data statistics charts, which provides powerful data support for urban traffic congestion management and short-term flow prediction, and greatly improves the decision-making efficiency and accuracy of urban managers.

Keywords

Data Fusion, Traffic Big Data, Data Visualization, Python Visualization

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着城市化进程的加速和城市交通网络的不断拓展,交通拥堵、道路容量受限、交通设施不完善以及缺乏整体规划等挑战日益凸显。为了精准且实时地获取并分析交通大数据,并直观展现其时空动态变化,已成为解决问题的关键所在。针对当前城市交通数据的复杂性及利用率低下的问题,本文设计了一套适用于中国现代化城市的交通大数据可视化系统。

为了缓解道路交通压力,许多学者对如何缓解交通拥堵和设计交通大数据可视化系统开展了研究。毛晨希[1]等人设计基于大数据的智能交通疏导控制系统,设置数据总览、区域路段信息、历史信息概览、应急指挥四个模块。龚晓岚[2]等人利用大数据地理信息系统并借助机器学习模型 Spark MLlib 等设计了实时交通大数据分析和可视化平台。向红艳[3]等人注意到了人流量对道路容量的影响。刘新华[3]关注总结了客流量数据的特征,并构建了客流预测模型,合理规划出行路线,减少拥堵。朱军和吴思豪[4]等人在研究可视化系统过程中提出了在网络环境下使用流光线表达交通态势的轻量化可视化方法。张昀昊和朱军[5]等人提出了面向多样化终端的自适应网络三维可视化方法,为不同设备可视化提供了方法。赵丽[6]总结关注到了车辆轨迹对道路拥堵的影响。刘文杰[7]总结提出了交通的数据可视化的基本框架,进而探讨了可视化系统的实现效果。陈为和朱标[8]等人通过构建贝叶斯网络模型并设计了一个可交互式的可视化分析界面,将多个分布关系可视化视图进行结合。李一凡和杨帆[9]等人通过对城市交通时空数据的可视化研究,利用多层次的可视化方法实现交通路线网络数据地图的 PC 端与移动端显示,缓解城市交通压力,缩短通行时间。

我国智慧交通相较于其他国家发展时间较短,在软件和硬件开发利用中仍然存在许多问题,目前所有的交通系统操作繁琐、信息统筹能力不足等,且交通大数据来源广、模态多样,无法快速处理以及存取海量的数据,为数据挖掘分析带来了不便。

本文设计的交通大数据可视化系统将利用大数据分类存储技术、多模态数据快速检索技术、前端可视化技术并借助大数据地理信息系统技术设计了一个能够缓解城市交通压力、减少交通堵塞的交通大数据可视化系统,通过对车辆检测数据、交通设施数据等各类数据进行处理、融合、分析并进行数据分析后输出所需的结果。系统主要数据为车辆车速、道路车流量、道路事故、道路容量、信号灯以及监控分

布等；首先以道路容量为主要因子，车速、道路车流量、交通事故、监控分布以及道路养护频率、驾驶员状态等为辅助因子，通过对主要因子和不同辅助因子的结合进行计算得出各路段在不同时间之内的拥堵指数，来呈现当前城市交通状态。

2. 交通大数据可视化关键技术及其创新

2.1. 基本架构

系统架构的设计采用层次化原则，具体分为用户层、业务层、应用层、服务层以及数据层。用户层作为整个系统的服务终端，直接面向用户群体；业务层则专注于将用户需求转化为具体的业务模块设计；应用层为业务模块的开发提供必要的组件和技术支持；服务层则是连接业务需求与数据处理的桥梁，通过数据的协调和处理来满足业务需求；而数据层则是整个系统架构的基石，承载着系统的底层数据源，见图 1。



Figure 1. System architecture diagram
图 1. 系统架构图

2.2. 关键技术

2.2.1. 多模态数据融合算法

本文利用基于自监督学习的多模态融合算法将交通大数据可视化系统所需的数据进行融合处理。通过自监督学习捕捉不同模态数据之间的关联性，增强数据的通用特征，并融合各类数据中的相同性质，使得算法能够更加精准地抓取各类数据之间的联系。通过实验数据表明，与现有算法中效果最佳的多模态融合算法相比，该算法在多个数据性能上取得了显著提升[10]。

2.2.2. 道路基本数据分析

交通数据主要包括车流量、车速、交通事故等基本算数据。其中道路容量在计算道路拥堵以及道路疏通缓解占据主要地位。本文通过利用相关性计算车速、道路车流量、监控分布(信号灯分布)、交通事故、道路养护频率与道路容量之间的关系。使用皮尔逊相关系数分析车速、道路车流量、监控分布(信号灯分布)、交通事故、道路养护频率与道路容量之间的相关性。本文通过分别各类别数据与道路容量的相关系数：

$$r_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^n \frac{(X_i - E(X))}{\sigma_X} \frac{(Y_i - E(Y))}{\sigma_Y}}{n} \quad (1)$$

式中 r_{XY} ——皮尔逊相关系数； $X_i Y_i$ ——总体数据中的第 i 个数据。

而皮尔逊相关系数的变化范围为-1 到 1。若系数的值为 1，则意味着 X 和 Y 可以很好的由直线方程来描述，所有的数据点都很好的落在一条直线上，且 Y 随着 X 的增加而增加；若系数的值为-1 意味着所有的数据点都落在直线上，且 Y 随着 X 的增加而减少；其中若系数的值为 0 意味着两个变量之间没有线性关系。

为了将各类别数据更好的利用，计算完相关系数之后，通过聚类算法对数据进行相关性分析，最后进行图形化展示，以便直观地发现道路拥堵的最大因素，缓解道路压力，减少道路堵塞。

2.2.3. 时空相关性

在城市交通规划中，路段之间的空间拓扑关系呈现出一种独特且复杂的分布模式。交通流作为车辆在城市路网中的动态表现，其特性不仅受到时间变化的深刻影响，更与道路网络本身的空间拓扑结构紧密相连。这种时间与空间因素的交织作用，使得道路交通流数据呈现出显著的时空特性。为了深入理解和分析这些特性，时空相关性分析成为了一种有效的手段。通过这种方法，能够度量交通流量数据在时空维度上的聚合性和相似程度，进而揭示交通流在时间和空间上的动态变化状态。

(一) 全局时空相关性度量指数

为了评估路网中全局时空相关性的系数，本文采用了 Pferifer 等人定义的时空自相关函数(Spatio-Temporal Autocorrelation Function，简称 ST-ACF)。这一函数在研究全局时空自回归移动平均模型(STARMA)的过程中，特别是在计算其移动平均(MA)部分时，已展现出广泛的适用性。

在实际应用中，由于路网中路段间复杂的连接关系，一个路段往往与多个其他路段存在直接的相邻关系。因此，在计算全局时空自相关度量指数之前，本文引入了空间延迟算子的概念。该算子定义如下：

$$\begin{aligned} L^{(0)}x_i(t) &= x_i(t) \\ L^i x_i(t) &= \sum_{j=1}^N W_{ij}^{(i)} x_j(t) \end{aligned} \quad (2)$$

其中

$L^{(i)}$ ——空间延迟算子，它描述了路网中路段之间的空间和时间延迟关系；

$x_i(t)$ ——在特定时刻 t ，路段 i 的交通流观测值；

$W_{ij}^{(i)}$ ——权重分配机制，用于量化路段 i 与路段 j 之间的关联程度。

通过这一框架，本文能够更系统地分析路网中全局时空相关性的动态变化，从而为交通流和路网规划提供有力的支持。

(二) 局部时空相关性度量指数

在处理交通流问题时，由于各种因素的影响，上下游的交通流量往往会呈现出增加或减少的趋势。为了深入研究上下游交通流之间的关联，本文需要解决如下两个关键问题：

- (1) 确定上游交通流对下游产生影响所需的时间跨度；
- (2) 量化上游交通流对下游的具体影响程度。

尽管有学者，如 Vlahogianni 等人，尝试采用遗传算法(GA)来检测交通流量的关系，但这种纯粹的数学分析方法仅能通过数据变化识别出“相关”位置，却未能提供基于相关函数(Cross Correlation Function, CCF)所测量的交通流量固有的空间关系信息。这意味着，虽然遗传算法能在一定程度上揭示交通流量之间的关系，但它对于深入理解这种关系的空间特性却显得力不从心。

Box 等人提出的相关函数(Cross-Correlation Function, CCF)是一种有效工具，用于评估两个空间对象在特定时间延迟范围内的相关性。当考虑任意两个时间序列 X 和 Y 时，其中 X 经过一个或多个时间段后可能导致 Y 的产生，本文可以通过以下公式计算这两个序列在给定时间延迟 s 下的时空相关系数：

$$\widehat{\rho}_{xy}(s) = \frac{\gamma_{xy}(s)}{\sqrt{\sigma_x^2} \sqrt{\sigma_y^2}} \quad (3)$$

其中

$\widehat{\rho}_{xy}(s)$ ——在时间延迟 s 下时间序列 X 和 Y 的相关函数；
 $\gamma_{xy}(s)$ ——在时间延迟 s 下时间序列 X 和 Y 的胡协方差系数；
 σ_x^2, σ_y^2 ——分别是时间序列 X 和 Y 的方差。

$$\gamma_{xy}(s) = E[(x(t) - \mu_x)(y(t+s) - \mu_y)]$$
$$\sigma_x = [\sum (x(t) - \mu_x)^2]^{1/2}; \sigma_y = [\sum (x(t+s) - \mu_y)^2]^{1/2} \quad (4)$$

其中

μ_x, μ_y ——分别是时间序列 X 和 Y 的均值；
 $x(t)$ ——一定时间间隔内，上游车流量观测值；
 $x(t+s)$ ——在时间延迟 s 后，上游车流量观测值；
 $y(t+s)$ ——在时间延迟 s 后，下游车流量观测值。

从这些变量可以看出，当存在多个影响因素时，上述公式可以视为互相关函数；而在单一因素作用下，它同样可以作为自相关函数来使用。这为本文分析交通流量在上下游之间的动态关系提供了有力的数学工具。

3. 交通大数据可视化系统功能模块

3.1. 系统登录分发界面

在系统登录界面，用户输入正确的账号和密码后，系统会快速响应并自动跳转至主界面。在主界面上，用户可以根据自身兴趣选择并访问不同的功能模块，提高系统各项资源的利用率。如图 2。



Figure 2. System login interface
图 2. 系统登录界面

3.2. 交通大数据模块界面

交通大数据中心包含八个小模块，如图 3，如图 4，如图 5 都为拥堵分析页面。



Figure 3. Congestion analysis interface 1
图 3. 拥堵分析界面 1

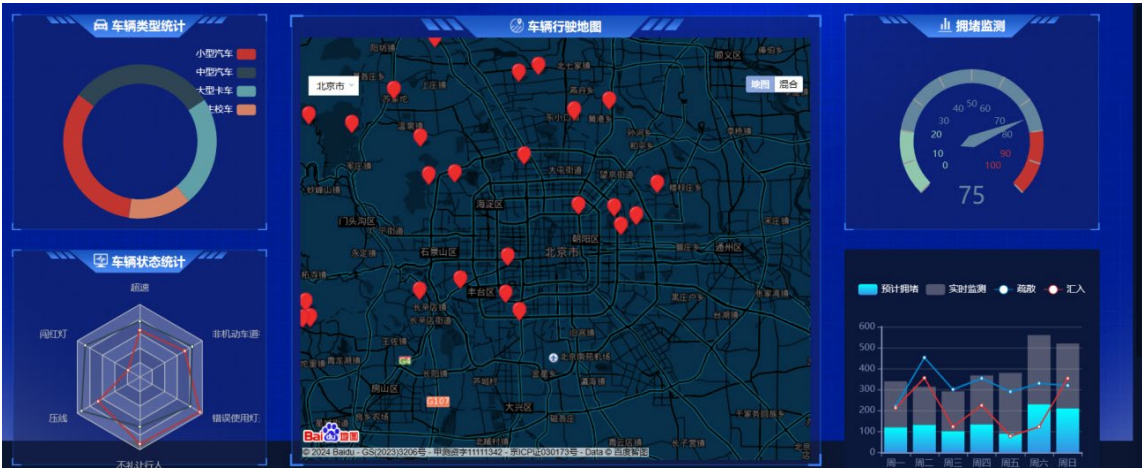


Figure 4. Congestion analysis interface 2
图 4. 拥堵分析界面 2

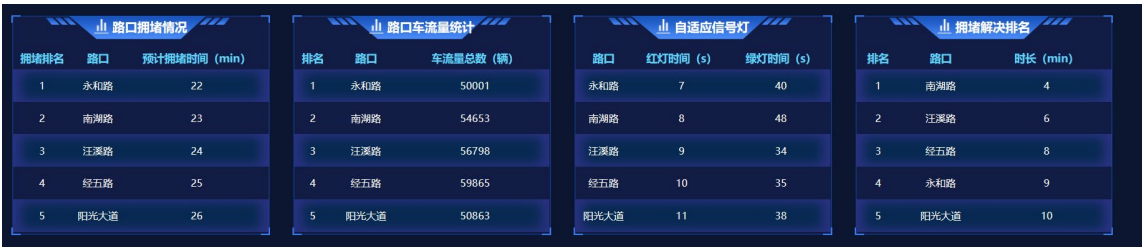


Figure 5. Congestion analysis interface 3
图 5. 拥堵分析界面 3

3.3. 静态物资管理模块界面

静态物资管理由底图、区域图组成。如图 6 为地图主页，如图 7 为静态物资管理的数据页面。

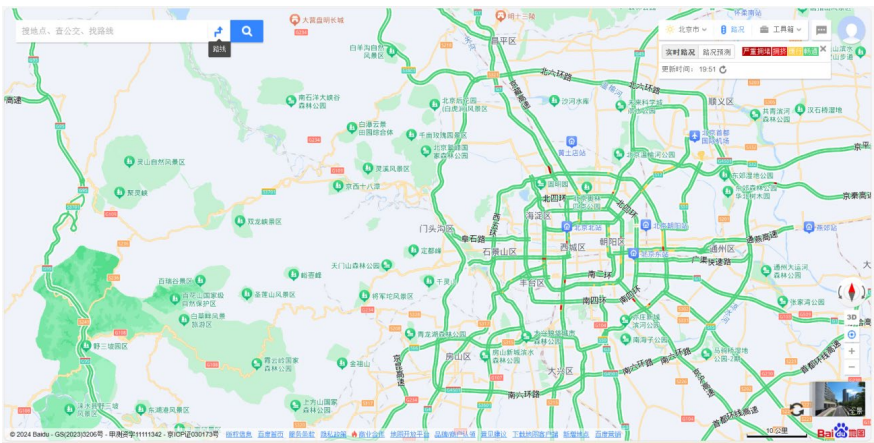


Figure 6. Map home page
图 6. 地图主页

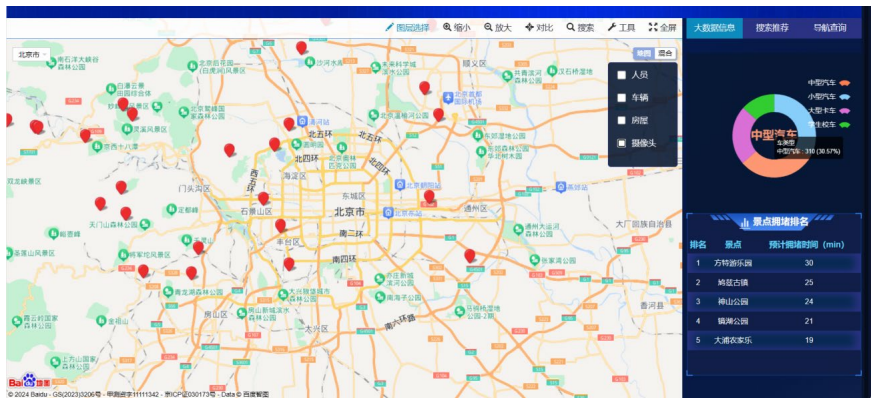


Figure 7. Static material management data page
图 7. 静态物资管理数据页面

3.4. 交通事件管理模块界面

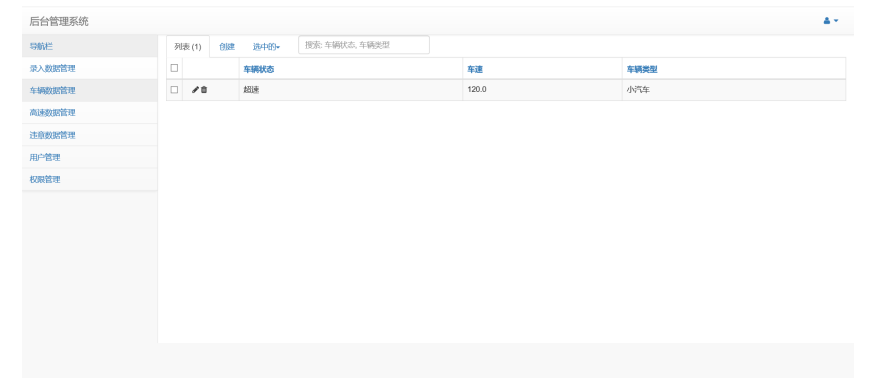


Figure 8. Vehicle data management page
图 8. 车辆数据管理页面

交通事件管理包括主要是对交通事件的记录。其核心功能在于实现对关键交通事件的精准记录 and 有效管理。如图 8 为车辆数据管理页面，如图 9 为新事件添加框。

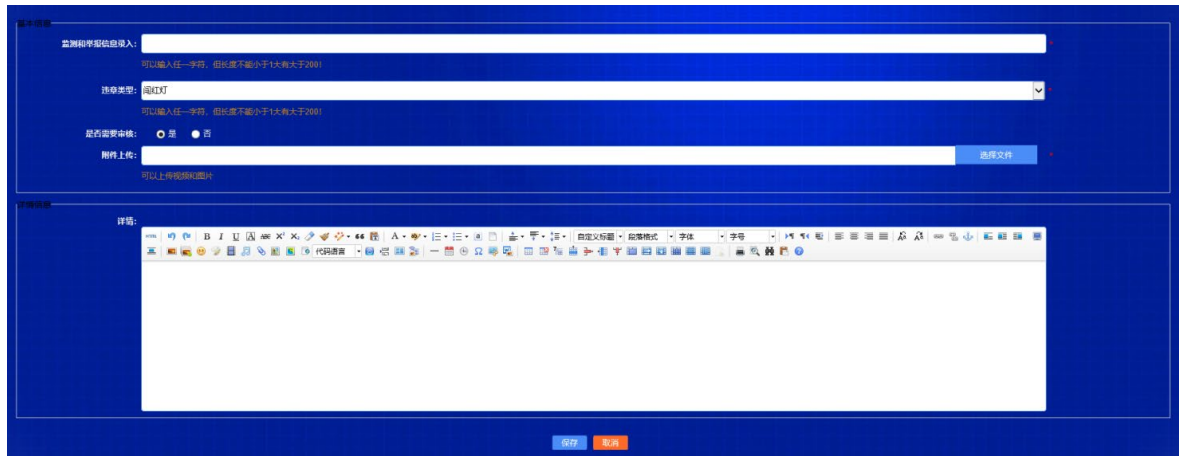


Figure 9. New event add box

图 9. 新事件添加框

3.5. 指挥调度模块界面

对管理模块的细化。包含对相关事件的详细记录。如图 10 为指挥调度界面，左边为交通事件的信息，中间是事件视频，右边为地图。



Figure 10. Command and dispatch event page

图 10. 指挥调度事件页面

4. 分析与总结

4.1. 本平台在实际应用中亦存在不足之处

尽管平台集成了专业的路段相关性分析功能，但在面向公众的解读和展示方面仍有待提升。未来，

将努力打造一個既专业又易于理解的平台，并进一步提升其短期交通流量预测能力，以满足更广泛用户的需求。

4.2. 展望

总体而言，本平台的建立对于城市交通研究具有深远意义，但后续的优化与升级工作仍需持续进行。在科技日新月异的时代背景下，期待通过不断的创新与努力，为城市交通的持续发展贡献更多智慧与力量。

基金项目

衡阳市科技局重点实验室，项目编号 202010041588。

参考文献

- [1] Mao, C.X. and Song, J.Y. (2022) Design of Intelligent Traffic Facilitation Control System Based on Big Data. *Software Engineering and Applications*, **11**, 1383-1393.
- [2] 龚晓岚, 于孝美. 实时交通大数据分析和可视化地理信息平台研究[J]. 贵州警察学院学报, 2023, 35(4): 77-83.
- [3] 陈玉飞. 面向交通大数据的可视化系统关键技术研究[D]: [硕士学位论文]. 唐山: 华北理工大学, 2022.
- [4] 朱军, 吴思豪, 张昀昊, 等. 大规模道路交通数据网络轻量化可视化方法[J]. 西南交通大学学报, 2021, 56(5): 905-912.
- [5] 张昀昊, 朱军, 李维炼, 等. 面向多样化终端的自适应网络三维可视化方法[J]. 西南交通大学学报, 2019, 54(5): 989-996.
- [6] 赵丽. 基于轨迹大数据的交通趋势可视化方法研究[D]: [硕士学位论文]. 沈阳: 沈阳理工大学, 2023.
- [7] 刘文杰. 城市交通大数据可视化框架及实现[J]. 科技创新导报, 2017, 14(36): 121-122.
- [8] 陈为, 朱标, 张宏鑫. BN-Mapping: 基于贝叶斯网络的地理空间数据可视分析[J]. 计算机学报, 2016, 39(7): 1281-1293.
- [9] 李一凡, 杨帆, 余正红. 城市交通时空数据可视化研究[J]. 软件导刊, 2021, 20(5): 139-147.
- [10] 任泽裕, 王振超, 柯尊旺, 等. 多模态数据融合综述[J]. 计算机工程与应用, 2021, 57(18): 49-64.