

基于政务服务热线数据的城市交通治理分析技术研究

马山¹, 郭玉彬², 周佳玮³, 唐立波⁴

天津市城市规划设计研究总院有限公司, 天津 300190

摘要 : 对百姓真实诉求重视不足、认识不清会让城市规划治理工作面临将为民服务理念难以落到实处的困境。基于“12345”市民服务热线收集的政务留言大数据,在时效性、真实性和广泛性等方面,能为城市规划治理的方向聚焦、策略引导、价值判断提供重要依据。本研究针对百姓政务留言大数据进行深度挖掘,通过人工智能大语言模型技术算法,从主题细化分类、空间位置关联、趋势预警研判等方面精准获取百姓对于城市发展的真实诉求,梳理出百姓对于城市交通环境的关注热点和存在痛点,为政府部门决策提供数据支撑和建议参考。

关键词 : 百姓真实诉求; 政务留言大数据; 人工智能; 大语言模型; 城市规划治理

Research on Analysis Technology of Urban Transportation Governance Based on Government Service Hotline Data

Ma Shan¹, Guo Yubin², Zhou Jiawei³, Tang Libo⁴

Tianjin Urban Planning and Design Institute Co., Ltd, Tianjin 300190

Abstract : Insufficient attention to the people's real demands and unclear understanding will make the urban planning and governance work face the dilemma of difficult to put the concept of serving the people into practice. Based on the "12345" Citizen Service Hotline, the big data of political messages can provide an important basis for focusing the direction, guiding the strategy, and judging the value of urban planning and governance in terms of timeliness, authenticity, and extensiveness. In this study, the people's political message big data are deeply mined, and through the technical algorithm of artificial intelligence big language model, the people's real demands for urban development are accurately obtained in terms of theme refinement and classification, spatial location correlation, and trend early warning and research judgment, etc. The hotspots of people's concern for the urban transportation environment and the pain points are sorted out, which can provide data support and suggestion reference for the decision-making of the governmental departments.

Key words : people's real demands; government message big data; artificial intelligence; big language model; urban planning and governance

一、工作背景

对百姓真实诉求重视不足、认识不清会让城市规划治理工作面临将为民服务理念难以落到实处的困境。目前城市规划城市治理已走向对公众负责的市民时代,正确回应社会热点舆情、推动城市交通治理能力和治理体系的现代化发展,已成为社会各阶层的一致共识。交通出行一直以来是城市重要的民生问题,也是百姓的主要关注点和生活痛点,交通问题的出现存在着高度时空不确定性,因此如何快速获取诉求信息、聚焦热点问题、了解问题成因、锁定争议区域成为当今政府以及广大社会研究人员关注的重点。

二、国内外研究现状

早期国外的网络舆论研究源于领导竞选中的民意调查。20世纪90年代,很多针对网络舆情的研究机构,发表了大量与网络舆

情相关的论文和著作。如Pawel Sobkowicz[1]等人提出了社会网络舆情系统模型,详细讲述了网络舆情从出现、传播到发展和消亡的四个过程,并分析了其中的关联性。Feldman[2]于1995年率先在非结构化文本数据上进行数据挖掘,涉及统计分析、自然语言处理、机器学习等多种方法。Nan Li[3]等采用文本挖掘技术对网络中网友探讨的热点问题进行甄别,并服务于预测分析。

由于中文语言文字的复杂性和特殊性,国内对舆情文本分析的研究起步相对较晚,目前主要研究方向包括语言分析、信息抽取、情感感知、文本检索等方面。Minghui Wang[4]等开发了利用学术文献的参考文献信息,通过对中文处理,实现多文档中文自动摘要系统。孙晓宁等[5]通过构造聚类分析、共词矩阵等分析方法,对近十年的学位论文进行文本挖掘,提炼高频词汇形成高校学术研究热点。陈冰洁[6]等基于旅游地网友留言评论语料,梳理历史街区旅游地的发展建议;赵阳[7]等基于高铁信号系统车载设备的故障记录文本,通过文本处理技术,完成了准确率较高的故障诊断。

三、研究目标及技术路线

本项目工作初衷是践行“听民声、察民情、解民忧、惠民生”的发展理念。采用创新的数据采集方式开展以人民为中心的调查研究工作，以“12345”市民服务热线海量工单数据为研究对象，基于人工智能大语言模型算法，自动化精准化完成工单数据主题细化归类、时空关联匹配等信息挖掘提炼，形成针对“12345”热线数据分析处理的完整路径体系。

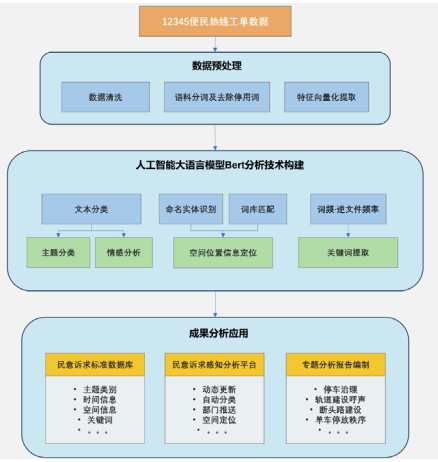


图1 技术路线

四、研究内容

(一) 数据来源

12345政务服务便民热线向企业和群众提供7×24小时全天候人工服务，本次研究获取2020年—2022年“12345”交通类热线工单数据107万条，包含主题类别、时间、留言内容等信息，涵盖交通出行领域所有百姓诉求问题。

表1 便民热线交通类工单数据

Table with 4 columns: 类别 (Category), 问题 (Issue), 时间 (Time), 问题 (Issue). It lists various transportation-related issues such as bus routes, traffic jams, and public transport services.

(二) 算法设计

1. 留言主题细化分类

本课采用人工智能大语言模型预训练方法进行特征提取，再结合实际场景进行微调形成定制化的百姓留言数据分类模型，实现对留言文本数据的精细化主题分类。

(1) 类型划分

目前采集数据或是分类质量较差或是并未包含分类信息，为保证从城市规划城市治理的角度能够直击问题要害，需要对留言类型进行细化，形成交通类分类体系。通过人工梳理留言内容，结合责任部门管理职责和重点需求，构建分类清单，其中共包含大类22项，细分类型407项。

(2) 数据标注

基于类型清单，通过人工识别的方法，对留言数据进行类型多标签人工标注，各类型共标注21万样本，作为模型训练数据。

表2 留言数据标注情况（以出租车类为例）

Table showing annotation details for taxi-related comments, including columns for comment text, category, and label.

(3) 分类器模型选择

本次训练方法基于大语言模型方法进行训练，选择大语言模型 BERT 进行多标签训练。它是一个预训练的語言表征模型，能够生成深度的双向語言表征。

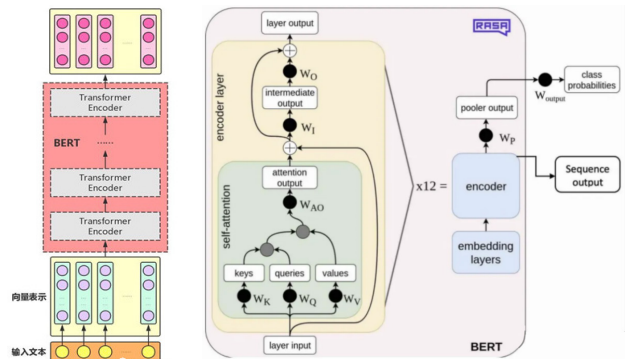


图2 BERT 模型分类任务框架

(4) 结果评估

本次研究针对所有类型进行分类模型训练，并采用多标签分类方法，平均准确率80%，远高于话务员人工分类准确率（33%），且类型划分更加精细，多标签元素体现更为全面。

Table showing classification results with columns for Precision, Recall, F1, and Class. It lists numerous specific transportation issues and their corresponding model performance metrics.

图3 多标签分类任务结果（以常规公交类为例）

表3 各类型留言分类准确率评价

类型	原类型	准确率	细分类型	准确率
出租车	18	32%	43	84%
网约车	13	24%	52	84%
常规公交	33	35%	40	83%
轨道	17	28%	40	75%
共享单车	5	28%	20	78%
停车管理	5	31%	10	83%
道路交通	7	21%	14	83%
慢行交通	3	22%	12	84%
高速公路	63	34%	27	79%
普通公路	30	33%	27	78%
货运交通			18	85%
物流	24	37%	3	85%
长途客运	10	36%	27	79%
航运管理	15	48%	10	77%
小客车指标	8	45%	6	83%
执法处罚	6	38%	3	78%
驾照培训	8	42%	5	79%
车辆维修	7	47%	2	81%
车辆管理	5	22%	8	79%
交通事故	4	28%	9	80%
行车管理	15	31%	20	77%
交通设施管理	5	26%	11	80%
平均准确率	301	33%	407	80%

(6) 模型预测

各大类交通方式训练完成后准确率若满足应用要求，会形成模型参数权重文件进行存储，后续将基于训练生成的模型权重文件进行留言主题多标签分类预测，服务于后续多元分析场景。

2. 空间位置关联匹配

(1) 词库构建

百姓的留言通常会携带关联地址位置信息，如“某社区停车难、某道路缺少路灯、某道路交叉口信号灯配时不合理”等。这些地址位置信息可通过构建自定义的地理空间词库（包含小区名称、道路名称、交叉口名称、公交站和地铁站名称等），进行位置信息匹配提取，结合位置类型颗粒度，对不同位置类型赋予权重，通过取权重最小值，规避掉多位置信息重复提取的问题。

表4 位置信息词库构建

类型	内容信息	数据量（万）	权重
路桥	路名、桥名	1.4	2
小区名称	小区名称及相关变体	1.4	1
地铁站	现状地铁站、在建及规划地铁站	0.9	1.5
公交车站	车站名称及相关变体	2	1.5
交叉口	交叉口名称及相关变体	49	1
公建建筑	公建名称及相关变体	0.5	1
区域范围	行政区、街道、居委会、组团	0.8	3
高速收费站	收费站名称	0.05	1.5
合计		56	

(2) 大语言模型调用

引入大语言模型（ChatGPT3.5）算法，针对不同类型留言场景，设置不同场景的位置信息提取提示词，进一步提升位置信息提取精度。

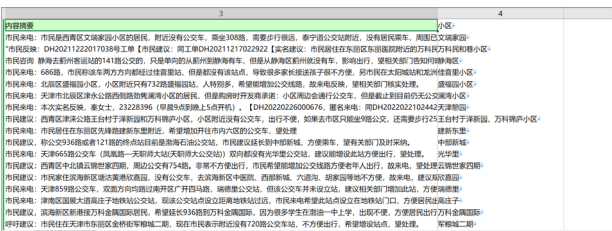


图4 大语言模型批量化结构输出

(3) 地址解析

对提取后的地址位置信息借助互联网地图的地址解析功能进行经纬度坐标的转换，从而将百姓留言与城市空间进行关联，形成对城市规划城市治理来说具有真正价值的数



图5 留言位置信息提取技术路径

五、成果应用—以天津为例

(一) 总体情况

近几年百姓对交通出行环境的关注意识不断加强，本次研究数据分别收集了2020—2023年“12345”便民服务热线工单留言数据约106万条。

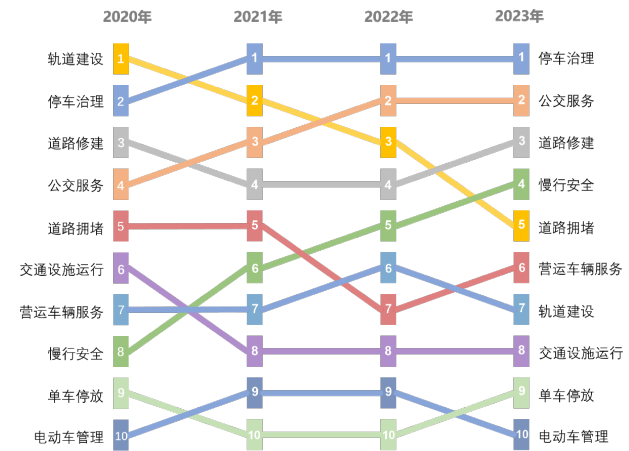


图6 2020—2023年百姓交通关注占比变化趋势

近几年百姓较为关注的交通环境问题中，停车治理、公交服务、道路修建以及慢行安全等方面排名前列。相比2020年，随着轨道交通线路陆续推进建设，百姓关于轨道交通建设呼声的留言占比逐年下降，而停车治理问题已上升为现阶段天津百姓最为关注的交通环境问题，停车难和停车乱现象逐渐加剧，停车环境正在逐步恶化，已成为影响和制约我市发展的重要因素。此外，百姓在公交出行服务、慢行出行环境等方面的关注留言占比均呈现显著上升的态势，均应引起重点关注。

(二) 主要问题

1. 停车设施短缺问题

近几年停车难停车乱现象日趋明显，也引起了百姓的高度关注，关于停车问题涉及的人大建议、政协提案、百姓投诉案件数

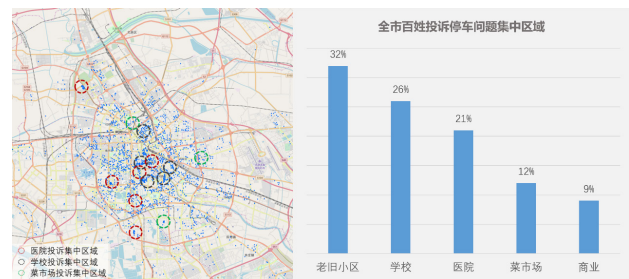


图7 百姓关于停车难的诉求空间分布

量逐年增加，已成为老百姓诟病最多的城市交通之痛。停车难、停车乱问题已成为制约城市发展的重要因素。百姓投诉停车难主要集中在老旧小区、医院、学校、商业集中地区，上述区域体现出明显的停车泊位总量缺口。

2. 公交服务低效问题

百姓对于公交服务的投诉主要为运行时间不准时、候车时间长、服务缺失、出行不便以及驾驶员素质等方面问题。



图8 百姓关于常规公交服务诉求排名

除车队组织管理因素外，主要由于公交场站分布不均、中心区大型换乘枢纽缺乏，造成线路过长，所有线路平均长度24.5km，津城平均线路长度20.5km，线路长度全国排名第一，带来准点率不高，运输效率下降等问题。

3. 慢行环境低质问题

百姓对于慢行出行环境的关注主要在慢行空间被侵占、增设过街天桥以及骑行安全性等方面。

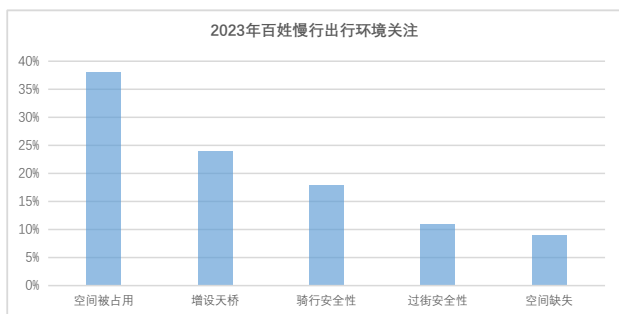


图9 百姓关于慢行出行环境诉求排名

其中津城核心区慢行空间尺度偏低，78%的道路人行道宽度不足3米，步行交通空间较为局促。慢行空间保障不足，非机动车道与车行道之间缺乏良好的物理分隔，82%的道路缺乏独立路权的非机动车道，非机动车交通路权保障有待提升。过街便利性和安全性也是关注重点，天桥的建设呼声较高区域主要集中在翠阜新村地区、一中心地区、华明空港地区等，主要为地铁车站出入口单侧设置及高等级干路切割造成，应重点弥补。

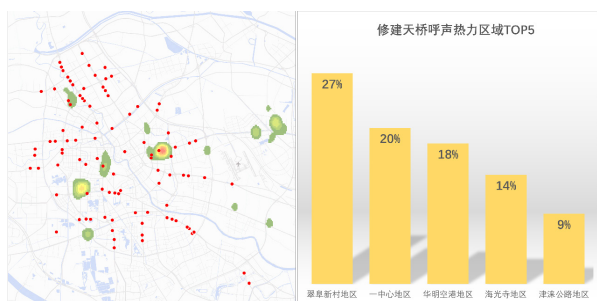


图10 百姓关于修建天桥呼应的空间位置及排名

(三) 建议措施

1. 停车分区差异化供给

(1) 已建成小区：在满足通行、安全等相关规范前提下，结合居民意愿和需求，通过平改立扩容、公共空间挖潜等措施，见缝插针增加停车设施；小区周边贴近需求建设公共停车场，形成“10分钟停车圈”；依靠次、支道路夜间设置临时停车区域，有效补充停车供给。

(2) 新建区域：高标准建设配建停车设施，保证停车供给满足基本停车需求；设置少量公共停车场，兼顾周边其他项目停车需求；设置少量路内停车位，补充特殊情况下配建不足问题。

2. 公交线网多层次一体化服务

(1) 形成骨架线（含BRT）为主体（骨干层）+普线为覆盖（支撑层）+微循环线为加密（接驳层）+定制公交等为补充（辅助层）的多层次、一体化公交线网服务体系。

(2) 通过整合现有资源，形成网络简洁、服务高频、换乘便捷，外环线附近布局公交枢纽站，位于轨道线路末端，截流远郊区-市区的市郊线路；中环线（轨道环）附近布局公交枢纽站，缩短公交运行时间，强化运行的准时性。

3. 慢行空间多功能场所营造

(1) 构建安全顺畅的慢行路网系统：重点考虑轨道接驳和弱势群体通行效率和安全提升，针对地铁2号线西段卞兴站、芥园西道站等，地铁4号线东段成林道站、泰昌路站等相关车站周边慢行接驳网络连通性和过街便利性进行提升。

(2) 构建宜商宜居的慢行空间场景：重点针对历史街区结合慢行空间改造进行慢行空间品质和景观环境提升，同时在居住区周边，选取绿视率较高路段，打造“落叶不扫”景观路，建议针对滨水道、泰安道、五号堤路等路段营造高品质景观路。商业消费提振方面，重点关注南开西湖道地区、红桥西北角地区、河北鸿顺里等区域。

参考文献

[1]Pawel Sobkowicz, Michael Kaschesky, Guillaume Bouchard. Opinion mining in social media: modeling, simulating, and forecasting political opinion in the web[J]. Government Information Quarterly, 2012, 29(4): 470-479.

[2]Feldman R, Dagan I. Knowledge Discovery in Textual Databases(KDT)[C]. International Conference on Knowledge Discovery and Data Mining, 1995: 112-117.

[3]LI N, WU D D. Using text mining and sentiment analysis for online forums hotspot detection and forecast[J].Decision Support Systems, 2010, 48(2): 354-368.

[4] Wang M , Tanaka H. Summarization of Multiple Chinese Technical Articles[J]. INFORMATION-YAMAGUCHI, 2002, 5(2): 221-226.

[5] 孙晓宁, 储节旺. 近十年知识管理领域硕博学位论文研究热点分析——以共词分析为方法[J]. 情报杂志, 2012, 31(6): 85-90.

[6] 陈冰洁, 胡巍. 基于文本挖掘的历史街区游客感知研究——以苏州平江路为例[J]. 云南农业大学学报(社会科学), 2016, 10(4): 48-55.

[7] 赵阳, 徐田华. 基于文本挖掘的高铁信号系统车载设备故障诊断[J]. 铁道学报, 2015, 37(08): 53-59.

[8]何梦娇, 吴戈, 梁华, 唐倩. 基于多源文本挖掘的城市交通舆情分析——以苏州为例, 交通信息与安全, 2018

[9] 赵金旭, 王宁, 孟天广. 链接市民与城市: 超大城市治理中的热线问政与政府回应——基于北京市12345政务热线大数据分析[J], 电子政务, 2021, 218(2): 2-14

[10] 朱智君. 以人民为中心的热线服务平台建设问题研究——基于L市12345市民服务热线的案例分析[D]. 吉林: 吉林大学公共管理, 2021.