面向城市交通场景的大数据并行分析与可视化系统开发

王辉
沈阳大学，辽宁 沈阳 110044
DOI: 10.61369/TACS.2026090002

摘　　要： 随着城市高速发展，面临着诸多复杂的城市交通难题如堵车及交通安全隐患等，传统交通控制手段已不能满足需求。而大数据因能存储大量信息并具备高处理能力，在交通领域有着良好的应用潜力。本文提出了一种面向城市交通环境的大数据并行分析及可视化的方法设计，在分布式计算环境下进行交通大数据并行计算，并采用合适的数据可视化方法将数据分析结果直观地展示出来。关键技术点在于如何利用交通信息特征构建高效率的并行算法以及如何建立多层次的复杂交通信息可视化模型。该系统的开发实现为提升交通管理水平、优化资源配置提供科学依据，有助于推动城市智慧交通建设发展。

关 键 词： 城市交通；大数据；并行分析；可视化系统

Development of a Big data Parallel Analysis and Visualization System for Urban Traffic Scenarios

Wang Hui
Shenyang University, Shenyang, Liaoning 110044

Abstract: With the rapid development of cities, many complex urban traffic problems such as traffic congestion and traffic safety hazards are faced. Traditional traffic control methods can no longer meet the demands. Big data, due to its ability to store a large amount of information and high processing capacity, has good application potential in the transportation field. This paper presents a method design for parallel analysis and visualization of big data in urban traffic environments. It conducts parallel computing of traffic big data in a distributed computing environment and adopts appropriate data visualization methods to visually display the data analysis results. The key technical points lie in how to utilize the characteristics of traffic information to construct highly efficient parallel algorithms and how to establish multi-level complex traffic information visualization models. The development and implementation of this system provide a scientific basis for improving traffic management levels and optimizing resource allocation, which is conducive to promoting the construction and development of smart urban transportation.

Keywords: urban transportation; big data; parallel analysis; visualization system

引言

城市的不断发展使在道路上面临着越来越多的问题如严重的堵车现象与事故等，这些都给的出行带来了不便，导致了更多的能源浪费与环境污染。同时也会威胁到人们正常的日常生活甚至生命安全。然而大数据却能很好地应对大量的数据、多种多样的数据、杂乱无章的数据等等，因此对于大规模的城市交通来说有着很大的发展前景。通过大数据信息的收集并对其进行分析处理之后可以更加清楚地了解到交通状况并且为改善其交通情况提供了依据。

一、面向城市交通场景的大数据并行分析

（一）交通大数据特性分析

交通大数据是智慧城市的有机部分之一，在特点上有着大量的、多样化的、高速度的及有价值的数据等特点[1]。首先是大量性的体现，即在数据来源方面有很多数据产生源头，比如车上的各种传感器产生的数据、摄像头采集到的画面信息、手机信号传输过来的信息甚至是通过社交软件收集到的各种信息等等，而这

作者简介：王辉（1970.03—），男，籍贯：凤城市，沈阳大学就职，职称：高级实验师。学历：本科。毕业单位：辽宁大学，电子学专业，获理学学士学位；研究方向：计算机科学与技术。

Copyright © This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

些数量一般都会达到 PB 级别的大小，并且会随着时间推移不断增多。其次是多样化的特点，也就是类型上的丰富，从种类上看有结构化的数据例如汽车运行轨迹，也有半结构化的数据如 XML 形式的日志文件，还有不规则的数据如视频片段或者图片文件，这就需要去处理并整合。

此外，交通大数据的特点之一是速度快，在实时交通管理过程中，产生的数据数量巨大且变化非常迅速，因此要求系统的高速度来保证不会出现滞后的现象。另外一方面，虽然交通大数据的数量很大但是其中蕴含的信息并不多，必须利用高效的方法对其加以分析，并发现其中所包含的一些规律，才能够帮助管理者做出正确的决定。例如，通过仔细研究过去的交通量，可以估计出高峰期时会遇到什么样的拥挤状况，并且能够更好地分配资源。例如，在不影响准确性的前提下，如何快速地处理大量信息等。还有如何构建一个灵活的并行结构来适应不断增加的信息数量等等。

（二）并行分析框架选择

对于大数据并行分析而言，在面对城市交通问题时需要考虑的是如何找到一个合适的数据分析方式来保证系统的运行效率以及可靠性[2]。目前比较流行的两种数据分析模式为 Hadoop 和 Spark 模型，这两种模式都具有一定的分布式功能，但是由于各自的特点不同所以在实际应用过程中会存在一些差异。第一点就是以 Hadoop 的特点来说的话，它是一个非常优秀的、有较高容错率且可以实现良好拓展的功能，其中主要由 MapReduce 组件构成，这个组件就可以很好地将大量数据进行有效的处理，并且适用于交通行业中的相关历史信息追溯及分析工作当中去。比如通过建立于 Hadoop 平台上的分布式文件系统（HDFS）就能有效地保存大量的交通日志文件，同时利用 MapReduce 模型也可以很好的完成车辆行驶路径的相关并行化操作。

然而 Hadoop 仍然不能够满足实时处理的需求，由于其基于磁盘进行迭代式计算的特点，难以满足对交通控制系统的超低延时要求；而基于内存计算引擎 Spark 框架可以更好地支持实时交通流数据以及机器学习算法分析工作。比如 spark streaming 组件可以实现交通状况的实时监测，mllib 组件可以提供用于预测交通量的各种机器学习方法。根据对交通环境的要求，本文采用 spark 框架进行并行分析，不仅能够实现海量数据分析，而且具备实时性，并支持与 hadoop 生态系统对接，便于现有系统的集成应用。

（三）并行分析算法设计

针对交通流量预测与拥堵分析等具体场景，设计高效的并行分析算法是提升交通管理效率的核心环节。在交通流量预测方面，本研究采用基于 LSTM（长短期记忆网络）模型的并行化算法，该算法能够有效捕捉时间序列数据中的长期依赖关系，从而实现对未来交通流量的高精度预测[3]。具体而言，算法首先通过 Hadoop 平台的 MapReduce 框架对历史交通流量数据进行预处理，将其转换为适合 LSTM 模型输入的格式；使用 Spark 的分布式计算平台来获取和处理的训练数据，然后开启多个线程以优化的模型参数。最后，通过模型的推断过程，可以获得未来的交通流动预测结果。对于拥堵问题的解决，本研究提出了一种新的、创新性的并行方法，用于检测交通网络中关键瓶颈路段及路线。该方法将交通网络建模成图，其中顶点代表街道或交叉口，如果有边相连就表示是相关的，而且顶点负载量（流量）和行驶时间作为输入值，从而快速预测出潜在的拥堵区域。实验结果表明，本方案不仅具备较高的运算准确率，同时具备较高的运算效率，在海量道路交通网络中具有更好的适用性；此外，为进一步提升本算法的应用效果，本文采用分布式索引方式结合 ElasticSearch 和 HBase 实现海量道路交通信息的快速检索和调用。这样可以为交通管理提供更多的信息依据。

二、面向城市交通场景的大数据可视化系统开发

（一）可视化技术选型

对于海量城市交通场景的大数据可视化来说，需选择合适技术实现系统功能并满足用户需求体验。ECharts 是一个采用 JavaScript 开发的开源可视化图表库，在数据可视化领域中应用广泛，具有图表类型多、交互性强等特点。它能够支持多种图表形式包括折线图、柱状图、散点图等，并且可以利用地图模块更好地表达地理信息，特别适合于反映交通流、堵车状况等时空数据。其次，ECharts 被称为另一个著名的可视化库，因高度可配置化和灵活性而著称，能够使用 SVG 和 Canvas 技术开发出较为复杂的数据图表以实现交通场景下精细化显示需求。因此选取 ECharts 主要是考虑其易用性及优化效果。例如 ECharts 提供了丰富的配置项及 API 接口，开发者可以快速构建满足交通管制需求的可视化展示系统，在海量的数据集上有良好的绘制性能，这是实时呈现交通流状况的关键所在。

（二）可视化系统架构设计

通常来说，基于大数据可视化的系统框架是面向于城市交通领域的，并由数据层、处理层以及显示层构成，在这些层次中存在明确的分工及数据流方向。数据层是整个系统的基础部分，用来存储并维护各类交通数据源，包括公交刷卡数据、出租车辆 GPS 数据及路面监控图像等。这种类型的数据一般是不固定或者是半固定的数据格式，需要依靠分布式存储系统 (Hadoop HDFS) 来高效存储，并用 NoSQL 数据库 (HBase) 支持快速访问查询；此外在数据层还需具备数据清洗预处理功能：这样就可以确保输入到处理层中的数据质量能够满足数据分析的要求。

在系统中的数据层与显示层之间就是处理层了，这是整个系统的计算中心点，在这个过程中就需要把收集到来的原始数据经过一定的处理、分析及转换等操作，形成可供显示使用的数据形式。在这个阶段就可以利用像 MapReduce 或者 Spark 这样的分布式的计算方式来进行大规模的交通数据的处理工作，并且大大提高了数据分析挖掘能力。比如可以通过 SparkMLib 这样一个机器学习工具来进行过去的历史交通流量数据建模并且做出相应的预测，以此帮助相关的管理部门作出更好的决定。但是从这里产生的结果一般会通过一种结构化的形式保存在一个缓存数据库（如 Redis）当中，这方便之后能够快速地把这些数据展现出来[4]。

展示层作为用户的入口，主要是把数据通过清晰明了的方式

 Copyright © This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

展现出来供用户查看。一般而言，在这个层面中采用的是 web 技术来完成，并且运用一些前端框架以及 Echarts 等工具来进行页面设计。在设计时要考虑到交通方面的特殊情况，如利用地图来显示汽车行驶路线、用热力图表示交通拥挤的地方等，还可以用一个动态图形来反映交通量的变化情况等等。各个层次之间一般是单向流动的，比如是从数据层流到处理层再流到展现层，在这样的系统中不但使得整个系统更加容易被分解成一个个的小单元来进行管理和控制，并且也方便了对于整个系统的扩充与修改。

（三）交通信息可视化展示

对于大数据可视化的系统来说，在针对城市的交通情况上进行分析时，如何能够把得到的结果呈现给使用者，这就需要借助于一种有效的手段来实现这一目标了。而这种有效的方式就是利用合理的展示方法去呈现出复杂的信息，并且将其转变为人们容易理解的形式，这样才能够让人们对这些重要的信息有更快的理解。例如其中的一种形式就是采用地图的方式来展现，这种方式特别适合用来表示有关交通量以及交通堵塞的情况还有汽车运行路线等等都带有地理位置性质的信息。如在城市中用 Echarts 的地图来表示城市的交通网路，在各个地区上显示交通的压力大小的不同，采用不同颜色的热力图层来进行表现。这样不仅可以看出哪些地方是较为严重的堵车路段，还能发现一些容易出现堵塞的地方。除此之外，用图来表示也是一种很好的方法，尤其是用来研究时间上的变化情况以及统计数据的分布特点等。

将一段时间内的交通量用折线图表示出来可以直观地看出这段时间内车辆数量的变化情况，并且可以通过比较同一时间段下各条道路之间的速度来判断哪条路更适合通行，也可以利用饼形图显示各种交通工具所占的比例等方法都可作为交通部门制定计划的重要参考标准，在实际运用过程中我们还可以根据用户的需要添加一些交互性操作如当鼠标经过时出现相应的信息、以时间为横坐标的时间轴控件等都能够增加其使用的便利程度并让用户可以从各个方向了解其中的信息。比如在选择不同的展示方法上可以根据实际需要来决定。比如说对于大量的汽车行驶路线来说，用地图的方式来展现更加直观明了，但是如果是针对某个时间点上的交通量最大最小的时候就需要采用图表的方式来进行展示了。这样就可以充分利用各种展示的方法达到最佳的效果，并且给相关部门提供最完善的数据信息帮助他们做出更好的决策[5]。

（四）可视化系统的交互设计与用户体验优化

对于基于大数据的城市交通系统的开发来说，在实现功能的同时还需要注重用户的交互体验感。因为交通类的数据都是处于不断变化的状态，并且还存在一定的空间关系，因此除了要展示出这些信息之外还要让用户能够对其做更多的了解与研究。因此，该系统的可视化需要很多的互动功能，例如：可以改变时间轴、选择地区范围、开启关闭图层以及点选数据，这样就可以让使用者按照自己的需要来更改所看到的画面内容及大小了。比如：交通警察可以通过拉动时间轴来看到早中晚高峰期时车辆数量的变化情况，也可以直接点击某条道路以获得此地的具体车速、拥堵程度以及事故发生的情况等等。为了给用户提供更好的体验，在系统的界面上要尽可能做到简单明了，并将地图、图表及控制板等元素巧妙地结合在一起而不显得杂乱无章。此外还要考虑到该软件能在各种大小屏幕上的应用情况，比如电脑、平板以及大屏指挥中心等等。同时，在地图上添加一些视觉引导及过渡效果（例如交通流变化时的颜色渐变热力图或车辆行驶路径上的平滑动画），可以提升数据的易读性和理解度。随着不断获取用户的反馈意见，并加以改进，系统的可视化方式将逐渐形成一种基于用户的数据互动体验形式，从而真正做到由“看数据”转变为“用数据”。

三、结束语

展望未来，本系统有希望同其他新技术如人工智能、物联网相结合，使其得到更广泛的应用。如在系统中加入人工智能，可以让整个系统的运行变得更加聪明，在一定程度上解决交通拥堵的问题以及一些突发事件的发生等等。还可以通过利用物联网技术来提高数据收集的能力，使得汽车、路途设备之间相互联系起来，并且也方便了交通中心的管理，这样就形成了一套更为完整的体系——一个智能化的城市交通管理系统。这样一来，我们的生活将会发生翻天覆地的变化，使城市的交通向着更高的层次去努力和发展，给人们的生活带来了极大的便利。

参考文献

[1] 路守望；王恒 . 大数据分析在交通领域中应用的特点与研究机会 [J]. 新经济 ,2021,(11):49-53.

[2] 洪真；郭宇飞；孔娟娟 . 大数据在城市智慧交通系统中的应用研究 [J]. 通讯世界 ,2021,28(9):125-127.

[3] 傅思萍 . 基于大数据技术的城市智慧交通体系建设 [J]. 湖南邮电职业技术学院学报 ,2023,22(3):59-63.

[4] 石秀君；李鹏浩 . 大数据技术在道路智能出行中的应用 [J]. 中国高新科技 ,2023,(14):56-58.

[5] 张昊；张健钦；王家川；石睿轩；陆浩；张安 . 基于云存储的城市交通大数据可视化系统 [J]. 科学技术创新 ,2021,(1):81-82.

Copyright © This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License