

多源异构航班航迹数据流实时融合方法研究

张曠熹^{1,2}, 田旺^{1,3}, 朱少川^{1,4}, 刘洪岩^{1,4}, 朱熙^{1,5}

(1. 北京航空航天大学综合交通大数据应用技术国家工程实验室, 北京 100083; 2. 中国人民解放军 32751 单位, 北京 100039;
3. 北京航空航天大学大型飞机高级人才培训班, 北京 100083; 4. 北京航空航天大学电子信息工程学院, 北京 100083;
5. 北京航空航天大学前沿科学技术创新研究院, 北京 100083)

摘要: 二次雷达和广播式自动相关监视 (ADS-B, automatic dependent surveillance-broadcast) 是在空域监视系统中共存的两种主要监视手段, 为了提高监视的精度和稳定性, 实现二次雷达和 ADS-B 航迹实时融合至关重要。针对现有方法难以满足大规模航迹的实时融合需求, 设计了一种使用大数据技术的二次雷达与 ADS-B 数据流实时融合的方法。该方法基于微批处理的大数据处理框架, 遵循 MapReduce 编程模型, 在得到较高质量融合航迹的同时, 保障了系统数据处理的高并发能力与实时性。最后, 基于真实航班数据开展了航迹实时融合仿真实验, 验证了方法的可行性。

关键词: 航迹融合; 多源异构; 微批处理; MapReduce; 流式大数据

中图分类号: V355.1

文献标识码: A

doi: 10.11959/j.issn.2096-3750.2020.00181

Research on real-time fusion method of multi-source heterogeneous flight trajectory data stream

ZHANG Zhuxi^{1,2}, TIAN Wang^{1,3}, ZHU Shaochuan^{1,4}, LIU Hongyan^{1,4}, ZHU Xi^{1,5}

1. National Engineering Laboratory of Big Data Application Technologies of Comprehensive Transportation, Beihang University, Beijing 100083, China
2. Unit 32751, Chinese People's Liberation Army, Beijing 100039, China
3. Large Aircraft Advanced Training Center, Beihang University, Beijing 100083, China
4. School of Electronic and Information Engineering, Beihang University, Beijing 100083, China
5. Research Institute of Frontier Science, Beihang University, Beijing 100083, China

Abstract: Secondary surveillance radar (SSR) and automatic dependent surveillance-broadcast (ADS-B) are the two main surveillance methods coexisting in the airspace surveillance system. In order to improve the accuracy and stability of surveillance, real-time fusion of SSR and ADS-B trajectory is crucial. In view of the fact that the existing methods are difficult to meet the real-time fusion requirements of large-scale trajectories, a real-time fusion method of SSR and ADS-B data streams was designed with big data technology. This method was based on the big data processing framework of micro-batch processing and followed the MapReduce programming model. While obtaining a fusion trajectory of high quality, it ensured high concurrency and real-time data processing capability of the system. Finally, a real-time flight simulation experiment based on real flight data was carried out to verify the feasibility of the method.

Key words: trajectory fusion, multi-source heterogeneous, micro-batch processing, MapReduce, streaming big data

收稿日期: 2020-06-08; 修回日期: 2020-07-31

通信作者: 朱熙, zhuxi@buaa.edu.cn

基金项目: 国家重点研发计划 (No.2019YFF0301400); 国家自然科学基金资助项目 (No.61722102, No.61671031, No.61961146005)

Foundation Items: The National Key R&D Program of China (No.2019YFF0301400), The National Natural Science Foundation of China (No.61722102, No.61671031, No.61961146005)

1 引言

近年来,我国民用航空发展迅猛^[1],航班体量不断增加。截至2019年年底,我国在册运输飞机数量为3 818架,定期航线数量为5 155条,这两个数据在过去5年的平均增长率均达10%^[2]。民用航空流量的迅速增长,使空中交通管制自动化系统面临显著的航班监视压力^[3]。飞机航迹实时监视是支撑空中交通管制的基础能力,管制中心获得的航迹信息的准确性和及时性,直接关系到发出管制指令的准确性与及时性,决定了空中交通安全有序的运行,因而十分重要。

为了提升空域监视能力、维护运行秩序,多种监视手段被应用于空中交通管制系统。早期,空域监视主要依靠二次雷达^[4]完成。2007年,我国引入了第一部S模式雷达,随后S模式雷达在我国得到全面推广。目前,二次雷达监视技术较成熟可靠,但也存在建设维护成本高、易受恶劣天气影响等缺点。近年来,民用航空领域开始推广新型空域监视技术ADS-B^[5],ADS-B主要基于全球导航卫星系统(GNSS, Global Navigation Satellite System),利用空对地、空对空数据链实现监控和信息传递。ADS-B具有建设维护成本低、使用寿命长、定位精度高、更新频率快等特点,但也具有较大的局限性:1) 依赖GNSS对目标进行定位,本身不具备对目标位置的验证;2) 现阶段ADS-B设备的覆盖率远不及二次雷达的覆盖率。因此,二次雷达和ADS-B将在相当长的时间内共存。

为了实现优势互补,更精准地监视与管控空域,同时使用二次雷达和ADS-B作为监视信号源实现航迹的融合是一项非常重要的工作。在已有的研究中,文献[6-8]指出多源航迹融合对提升航迹监视精度、增强空中交通安全的作用,分析了现有空域监视系统中二次雷达与ADS-B的数据融合方案,论述了协同式的空域监视技术框架,研究了基于二次雷达和ADS-B的数据流控体系结构。在算法层面,文献[9]提出了调整二次雷达与ADS-B航迹融合权重因子的可变周期更新算法;文献[10]提出了在传统的地心地固(ECEF, earth-centered, earth-fixed)坐标系下二次雷达与ADS-B的航迹误差校准算法;文献[11]系统处理了航迹融合中的时间对准、空间对准、目标跟踪、数据融合等关键技术。在系统实现架构层面,文献[12]使用串行方式

依次处理多部雷达的数据;文献[13]指出在现有的空中交通管制自动化系统中,针对单个目标航迹融合的流程是先将单雷达航迹融合为多雷达航迹,再将雷达的航迹与ADS-B的航迹融合为系统航迹;文献[14]提到,现有空中交通管制系统中封闭式的网络结构、复杂的通信终端、落后的传输协议给空中交通管制系统的全面信息化、网络化、系统化带来了困难;文献[15]验证了消息通信库不同的通信模式对提高异构信源数据传输与交互能力的有效性。

综上所述,现有的研究主要侧重对单个目标的航迹融合算法的设计与改进,虽然有较好的融合精度,但设计的处理流程没有较好地考虑在实际运行时,系统承受高并发与灵活扩展计算性能的能力。现有空中交通管制系统对多源二次雷达与ADS-B数据的处理融合仍以较低效的串行方式为主,在面对大规模航迹带来的对计算性能的挑战时,低效的串行方式将限制空域的监视能力,难以满足在航班体量不断增长的将来,空中交通管制系统对空域实时监视与反馈的需求。如何在保障航迹融合精度的前提下对大规模二次雷达与ADS-B数据实现并行化快速处理、提高航迹数据流的实时融合效率,是一个具有挑战性的课题。

针对这一问题,本文基于微批处理的大数据处理框架,遵循MapReduce编程模型,设计了二次雷达与ADS-B数据流实时融合方法。在多数数据源、多目标、多维度的航迹数据之间采用并行处理模式,在提高航迹融合质量的同时,保障数据处理的实时性与高并发能力。在处理算法上,融合航迹由原始记录经过最小二乘法建模和拟合,再经过考虑决定系数的动态加权算法获得。在处理流程上,首先由分布式消息队列接收多源航迹数据,再按一定大小的滑动窗口划分批次,每一个批次数据的处理流程则包含多个Map和Reduce过程。最后,基于真实航班数据进行航迹实时融合仿真实验,对方法进行验证。

2 航迹实时融合关键技术

空中交通管制自动化系统复杂度、对系统稳定性和安全性要求以及部署成本均较高,这些特点使其在设备和技术更新换代方面较保守,导致很多数据处理新技术还未在该领域应用。而在互联网领域,为了应对大数据带来的分析处理方面的挑战,已经有很多成熟的技术被应用于实际生产和生活

中。在众多技术中,针对实时航迹融合问题,分布式消息队列可以有效地应对各类复杂数据来源,微批处理可以在保障实时性的同时,利用数据的时序信息,MapReduce 编程模型可以通过并行处理提高数据的处理效率。接下来,对上述技术基础进行简要地介绍。

2.1 消息队列

消息队列中间件是分布式系统中的重要组件,在面临大规模、多数据源的航迹数据时,使用分布式的消息队列有助于将数据处理过程与数据源解耦^[16],屏蔽各数据源差异带来的影响^[17]。通常来说,消息队列包含两个主要对象,即生产者和消费者。其中,生产者负责将数据放入消息队列,消费者负责从消息队列中获取数据并进行后续处理。常见的消息队列主要包含两种模式,即点对点模式及发布/订阅模式^[18]。前者用于生产者和消费者一对一的情况,后者用于生产者和消费者一对多的情况。

在处理多源异构航迹数据时,生产者可以是任意二次雷达或者 ADS-B 接收机,消费者可以是监控平台、存储中心或任意将航迹作为输入的计算过程。在这种情况下,采用发布/订阅模式更符合实际需求。在分布式场景下,消息队列中的数据会被分散至不同的节点,消费者的消费可能不会严格按照数据到达的时间顺序进行。而在处理航迹问题时,航迹点的时间顺序是十分重要的信息。在实际应用中,可以通过必要的配置确保局部消息的有序性。以 Kafka 为例,通过设置固定分区,将某个主题下的数据全部发往同一个分区,保证了这部分数据的有序性。在实际操作时,可以为每一个数据源注册一个主题,数据源的每一架航班划分到一个分区,实现每架航班航迹的有序性。

消息队列的引入可以屏蔽不同数据源之间的差异,提高多数据源之间航迹数据传输的效率。通过将每一个数据源抽象为一个主题 (Topic),在应对大规模的航迹数据时,可以方便地进行扩容,提升了系统的吞吐量。

2.2 微批处理

根据不同的应用场景和处理流程,大数据处理框架一般有批处理和流处理两种类型^[19]。其中,批处理主要用于处理大量的静态数据,这些数据往往经过了持久化存储,拥有明确的大小和边界。批处理适合进行数据挖掘的任务,如根据用户的浏览数据分析用户喜好^[20]、根据交通流量信息分析拥塞^[21]等。

与批处理相对应的是注重数据实时性的流处理^[22],流处理主要用于处理动态数据流,它往往与各种消息队列结合使用。流处理的数据源大多具有长度未知、时间边界不定、数据上下文关系弱、一旦丢失则不易寻回的特点。流处理能够对每一条数据进行快速的处理,常被用于运营监控领域,如网站运行日志的收集、服务器的状态监视等。

然而,在处理多源实时航迹融合的问题时,单一的批处理或者流处理大数据框架便无法满足业务的需求。批处理的模型对历史航迹进行分析,无法满足实时性的要求。流处理只能针对每一条位置记录进行处理,无法兼顾使用飞机在过去一段时间内的航迹所蕴含的信息。鉴于此,一种混合了批处理和流处理特性的模式——微批处理则更合理。

微批处理属于流处理的范畴,其核心思想是将连续的数据流按照一定的大小进行离散化^[23]。离散化后的每一个数据片段,可以看作是在该时间段内一部分数据的集合。在此数据集合的基础上,可以完成交互式查询、机器学习等任务。微批处理牺牲了少量的实时性能,但是能够更快地获取动态数据的中间结果,而不需要等待所有的数据传输完成。与批处理依赖磁盘存储相比,微批处理每次处理的数据量不大,因此可以将计算的中间过程保存在内存中,只需要最后将结果进行持久化存储即可,从而避免反复读取磁盘带来的输入/输出开销,加快了数据处理的速度。

多源航迹的融合算法正是基于离散化的连续航迹数据流进行的,微批处理可以提供航迹建模拟合等过程所需的历史航迹信息,同时也能确保融合航迹生成的实时性。

2.3 MapReduce 编程模型

MapReduce 是 Google 公司提出的一种并行计算框架^[24],它的设计目标是用于大规模数据集的并行计算。MapReduce 包含“Map(映射)”以及“Reduce(归约)”两部分操作。其中,Map 操作可以通过一个函数将列表里的每一个独立元素映射至新的元素进而组成一个新的列表,Reduce 操作则是对列表中的元素进行适当的合并求得计算结果。MapReduce 将输入抽象为 key/value 对的形式,通过 key 对各元素进行标识和追踪。因为元素相互独立,因此 Map 操作具有很好的并行性能。在分布式集群中,任务调度程序可以将 Map 操作的各个元素分配到不同的计算节点分别运行,或者在单机中使

用多线程的方式，达到并行处理的效果。在需要进行 Reduce 操作时，任务调度程序再将数据以及计算集中到某一个节点或者主线程获得结果。

针对航迹融合问题，通过合理地将数据处理任务划分为不同的 Map 过程和 Reduce 过程，可以充分发挥 MapReduce 编程模型的并行优势，有效地提高航迹融合的效率。同时，MapReduce 编程模型使得计算集群规模化效应优势更突出，通过简单地增加计算集群的硬件资源，可以方便地扩充集群的计算能力^[25]。

3 航迹实时融合方案设计

3.1 整体架构

根据现有技术基础，结合问题本身的特性与需求，航迹实时融合方案框架如图 1 所示。

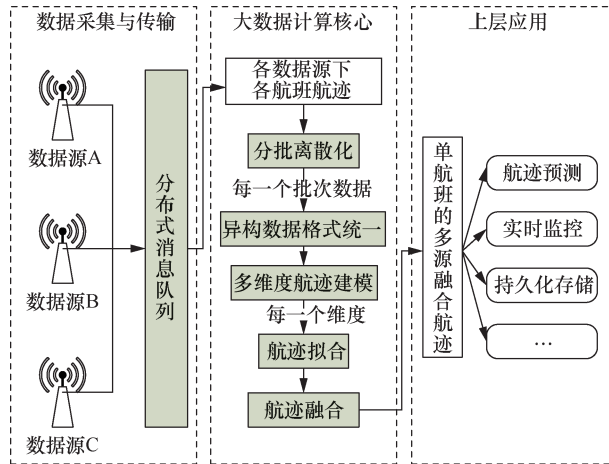


图 1 航迹实时融合方案框架

航迹实时融合方案框架分为 3 部分，底层是数据采集与传输，涉及多源数据的接收与发送。在大数据计算核心中，使用微批处理对连续数据进行离散化，对每一个批次数据进行异构数据格式转换、航迹建模等一系列计算。最后，融合后的航迹可供实时监视，持久化存储后可用作后续的科研支撑以

及其他航迹相关的研究。其中，对各个数据源的每一架航班的各个维度的航迹均使用了并行处理，提升了系统整体的处理效率。

3.2 数据采集与传输

监视数据源提供最原始的数据来源。在航空领域，常见的监视方式主要包括二次雷达和 ADS-B 两种。另外，飞机通信寻址与报告系统（ACARS, aircraft communications addressing and reporting system）也可以获取飞机的实时飞行状态。这些数据源具有不同的报告频率和格式，需要在后续流程中进行进一步处理。

数据传输层面的主要组成是分布式消息队列。分布式消息队列可以以多种形式（包括但不限于文件、TCP 长连接、Socket 通信等）接收数据，可以充分考虑不同数据源的差异。通常来说，数据源每一次发送的数据，会包含该时刻多架航班的位置信息，因此需要按照航班的唯一标识符（可以使用呼号、ICAO 地址、机尾号等）将原始记录拆分为多条记录，并依照唯一标识符发送至各自的消息队列分区。消息队列处理内容示意图如图 2 所示，展示了一个数据源的情况。一个数据源被注册为一个 Topic，供生产者和消费者在发布/订阅模式下使用，多个数据源被注册为多个 Topic，处理流程同理。分布式消息队列并行操作的特性可以使得多个数据源、多架航班之间的数据相互独立，提高系统的处理速度。

为了满足大数据计算核心部分使用微批处理模式的要求，需要对图 2 中每一架航班的信息进行分批离散化。为了在利用航班历史航迹信息的情况下，尽量确保实时性，每一个批次数据的组成由最新的一次航迹记录结合其前面的 $n-1$ 次航迹记录共同组成一个长度为 n 的数据集合，可以理解为一个长度为 n 的滑动窗口。在每次监听到新的记录产生时，以步长为 1 进行移动，便可以获得一个新批次的数据。微批处理示意图如图 3 所示，

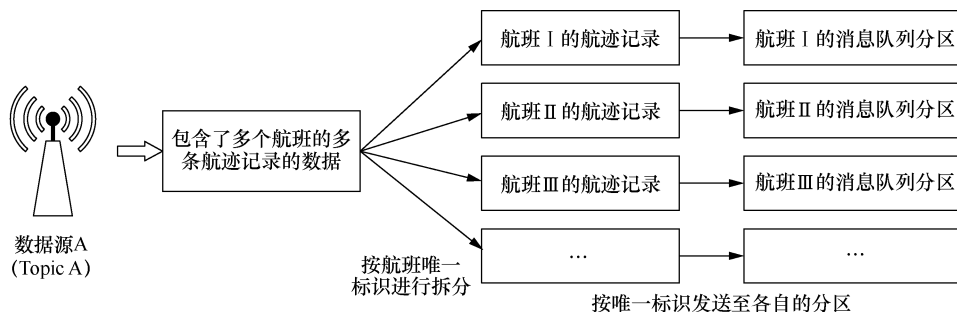


图 2 消息队列处理内容示意图

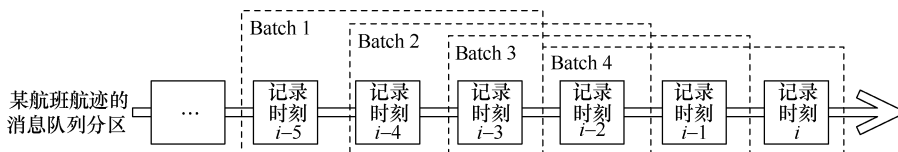


图3 微批处理示意图 (n=3)

给出了一个 $n=3$ 的示例。其中，时刻为 i 的记录为最新的一次记录，这个记录以及时刻为 $i-1$ 和 $i-2$ 的两个记录一起组成了 Batch 4，被送入下一个流程进行计算。在此之前，还有多个批次的数据以同样的方式产生并被处理。

3.3 异构数据处理

由于工作原理不同，二次雷达通常使用以接收站为原点的极坐标系，而 ADS-B 依赖于导航卫星使用的 WGS-84 坐标系，高度也可能使用米或者英尺这样不同的单位。为了便于后续计算，需要先对航迹报告的各数据项进行格式的统一。异构数据格式转换过程示意图如图 4 所示。考虑将一个航迹点的时空信息用带有时间维度的多维向量表示，为了提高数据并行处理的效率，需要先将数据按照维度进行拆分，多条记录、多个维度同时进行格式转换，并携带时间维度，将同一个批次内所有航迹记录相同维度的数据进行聚合。

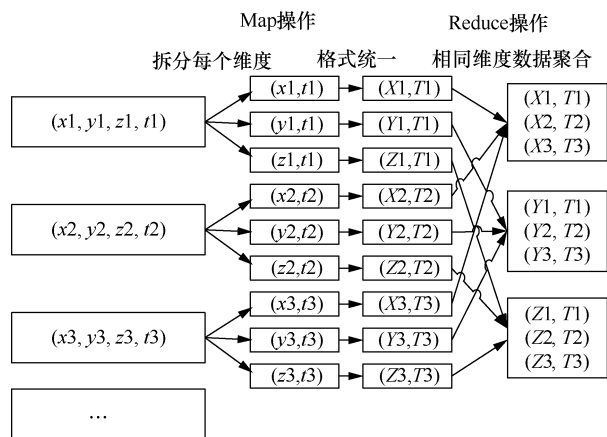


图4 异构数据格式转换过程示意图

3.4 航迹建模

考虑二次雷达和 ADS-B 报告数据的频率并不一致，为了得到某架航班在两个数据源任意一个时刻的航迹位置，需要依据一段时间内的航迹记录进行位置和时间关系的建模，将离散的航迹点用连续的方程进行表示。

考虑飞机在飞行过程中的绝大部分时间均处于直线飞行状态，为了减小计算的复杂度，将每一

个维度和时间组成的二维向量作为一组数据，每一个维度的数据变化与时间构成式(1)所示的一元线性模型。其中， V 为某一维度的值，如经度、纬度或者高度， T 为时间， w 为系数， b 为偏置。

$$V = wT + b \quad (1)$$

使用最小二乘法，基于单个批次内的航迹数据，计算各自的 w 与 b ，对各个维度进行建模。航迹建模过程示意图如图 5 所示。

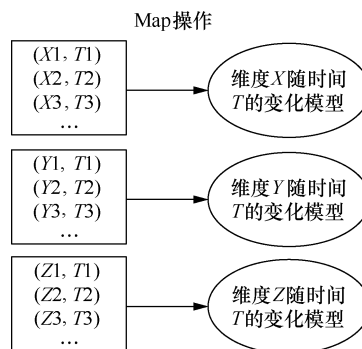


图5 航迹建模过程示意图

3.5 航迹点插值拟合

对于每个数据源，使用最新接收记录的时间戳作为输入，得到该时刻同一架航班的各个数据源的各个维度经由各自的模型拟合计算后的值。航迹点拟合过程示意图如图 6 所示。其中， T_i 表示最新的一个航迹记录的时间戳，该时间戳所属的记录可以来自二次雷达数据源或者 ADS-B 数据源，与图 3 中数据的时间戳标记对应。 X_{i_a} 表示航班位置维度 X 的 ADS-B 数据源经由对应模型计算得到的在 T_i 时刻的拟合值，其余标识同理。

3.6 多源航迹融合

通过对各数据源得到的同一维度的信息按式(2)进行加权求和，可以得到各维度融合后的结果。

$$V = \sum_{j=1}^K (\theta_j v_j) \quad (2)$$

其中， V 表示融合后的值， K 表示数据源总数， v_j 表示某一个维度在第 j 个数据源内的拟合值， θ_j 表示各数据源占有的权重，并且满足 $\sum_{j=1}^K \theta_j = 1$ 。

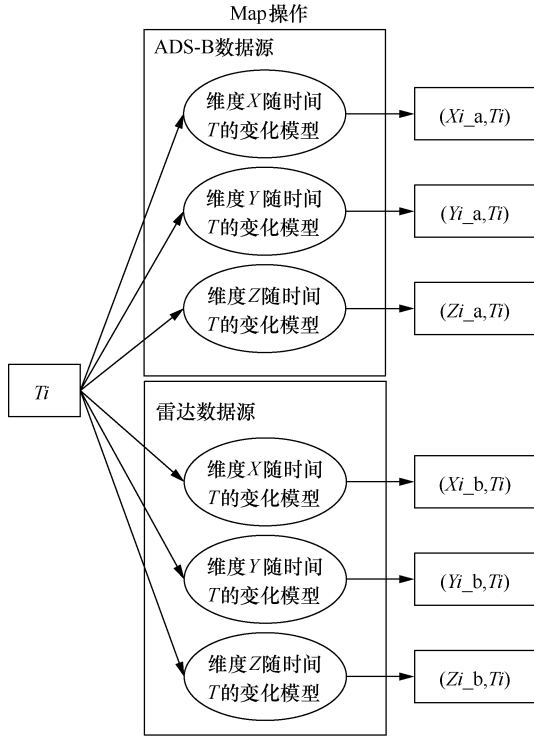


图 6 航迹点拟合过程示意图

θ_j 的取值大小主要由第 j 个数据源在当前批次前期参与建模的航迹点的质量而定。考虑航迹建模是一个以一个批次历史航迹作为基础、使用最小二乘法进行线性回归的过程, 决定系数 R^2 反映原始记录点相较于拟合航迹的吻合度, 其值越接近 1, 表示原始航迹点更平滑稳定, 更有可能接近真实航迹。用 R^2 作为衡量数据源质量的指标, 式(3)和式(4)给出了 R^2 和 θ_j 的计算式。

$$R^2 = 1 - \frac{\sum_i (\hat{y}^{(i)} - y^{(i)})^2}{\sum_i (\bar{y} - y^{(i)})^2} \quad (3)$$

$$\theta_j = \frac{R_j^2}{\sum_{i=1}^K R_i^2} \quad (4)$$

在式(3)中, $y^{(i)}$ 表示参与建模的当前批次内的第 i 个点的真实值, \bar{y} 表示参与建模的当前批次所有点的平均值, \hat{y}_i 表示参与建模的当前批次内的第 i 个点对应的模型预测拟合值。

最后, 将多个维度的记录合并成一条完整的航迹记录, 得到一个融合了多个数据源记录的在 T_i 时刻的航迹信息。多源航迹融合过程示意图如图 7 所示。

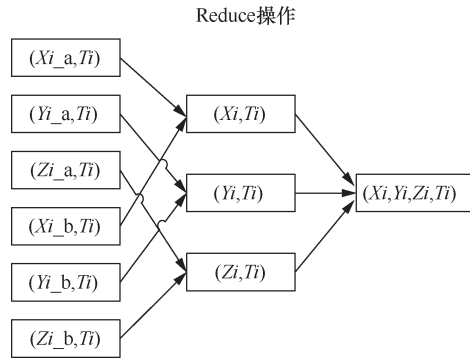


图 7 多源航迹融合过程示意图

4 航迹实时融合仿真

考虑单个二次雷达基站以及 ADS-B 基站的覆盖面积, 选取以上海浦东国际机场为中心, 300 km 为半径范围内的部分航班在 2020 年 1 月 18 日的包含了二次雷达及 ADS-B 的监视记录。对离线文件进行处理, 按照原始记录的时间戳模拟航迹发送过程, 在多 CPU 核心服务器集群上搭建并行实验环境。

以呼号为 ALK454 的航班数据为例, 设置滑动窗口的大小为 10, 各数据源的各维度拟合结果如图 8 所示, 给出了一次建模后二次雷达和 ADS-B 各自经度、纬度与时间关系的线性模型及拟合结果。可以看出, 相较于 ADS-B, 二次雷达的航迹记录偏差更大, 线性建模能够较好地纠正二次雷达的记录偏差。ADS-B 的数据虽然更准确, 线性模型的质量也更高, 但是其数据获取依赖于航班本身, 故航迹报告频率并不稳定。通过线性建模, 使用时间插值拟合, 可以对各自系统内的记录进行补充。

监视区域内航班航迹融合效果如图 9 所示, 给出了监视范围内呼号为 ALK454 和 JJA4304 的两架航班的原始航迹以及融合航迹。可以看出, 融合航迹稳定在原始的二次雷达以及 ADS-B 航迹之间, 符合原始航迹的记录。

JJA4304 航迹融合详细数据如图 10 所示。在“203148”时刻, 二次雷达以及 ADS-B 均汇报位置, 该时刻融合点位于两个原始航迹记录的位置之间。在“203156”时刻, ADS-B 数据源的航迹记录出现了缺失, 对该时刻的 ADS-B 数据源航迹进行拟合, 结合二次雷达报告的原始数据, 得到的融合航迹位置符合预期。由于二次雷达的报告频率更低, 大部分融合航迹点则来自于 ADS-B 数据源的原始数据及对应时刻二次雷达数据源的拟合值。

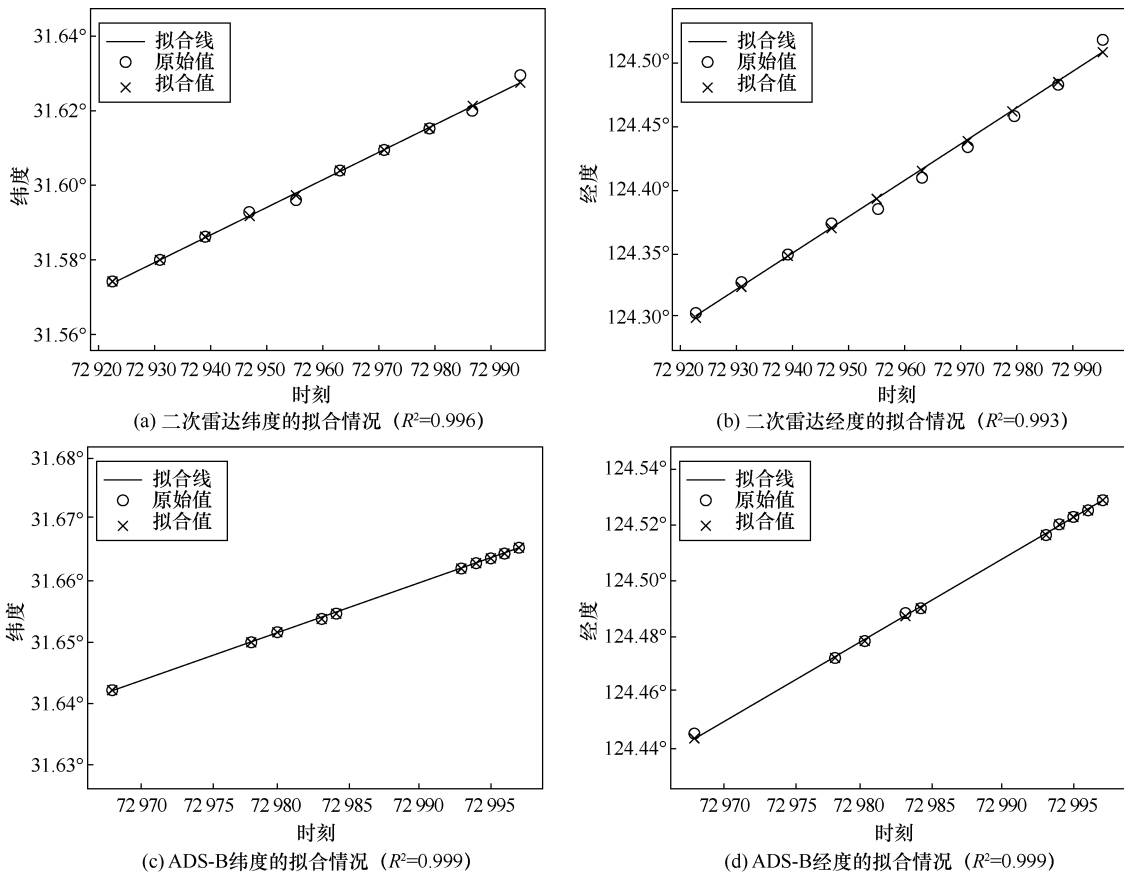


图 8 各数据源的各维度拟合结果

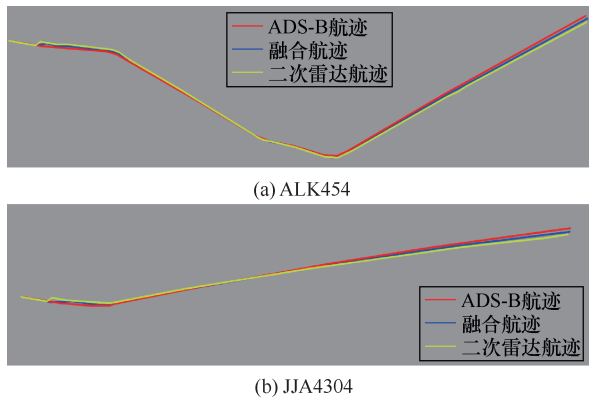


图 9 监视区域内航班航迹融合效果

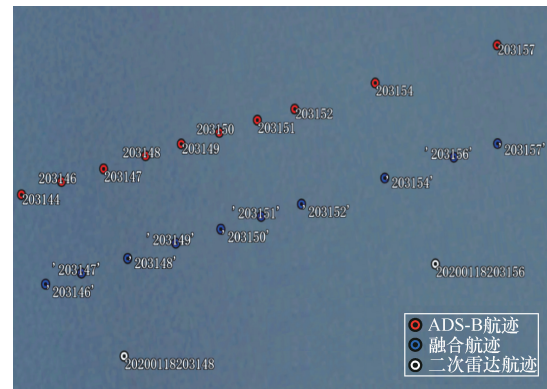


图 10 JJA4304 航迹融合详细数据

单个航班航迹融合效率评估如表 1 所示，给出了部分航班在监视范围内数据处理的耗时统计。在使用本地计算机单独计算每架航班数据的情况下，每条航迹记录从接收到计算出融合航迹，所需时间平均不超过 0.02 s，远低于 ADS-B 系统的报告频率（每 1 s 报告 1 次）。处理速度大于数据报告的速度，实时性能够得到保障。

对呼号为 ALK454 的航迹记录，分别复制 3 次、50 次、100 次模拟同时处理多架航班数据的情况。

在使用计算集群（32 核 CPU）的情况下，得到的不同规模航班航迹融合效率评估如图 11 所示。

表 1 单个航班航迹融合效率评估

呼号	ADS-B 记录数/条	二次雷达记录数/条	总处理耗时/s	平均每条数据处理耗时/s
ALK454	1 433	254	28.13	0.016
JJA4304	1 460	239	30.47	0.017
ANA8490	1 342	247	31.52	0.019

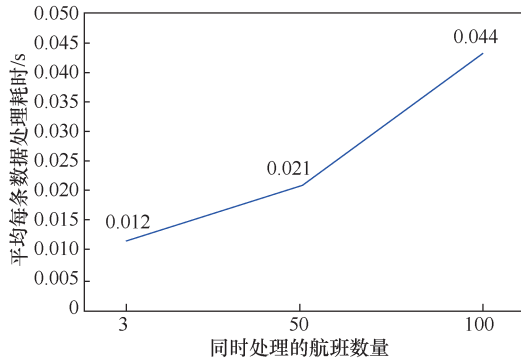


图 11 不同规模航班航迹融合效率评估

可以看出,在计算资源受限的情况下,考虑一个监管中心正常的监视容量,本方法处理每条航迹记录所需的时间依然控制在 0.05 s 以内,满足实时性要求。

本文还对所提出的并行处理方法的效率与传统串行处理方法进行了实验对比。传统串行方法采用了文献[12]中的航迹融合方法,即在多部二次雷达之间依次处理各个雷达的数据,并结合文献[13]介绍的处理流程,即在融合时先进行二次雷达数据融合,再将二次雷达数据与 ADS-B 数据融合,并依次处理多个航班。采用传统串行模式处理相同规模航迹和本文方法效率对比如图 12 所示。可以看出,本文方法采用的并行处理架构能够明显地提升系统整体的吞吐量,在相同的时间内,可以承受更多航班带来的并发。

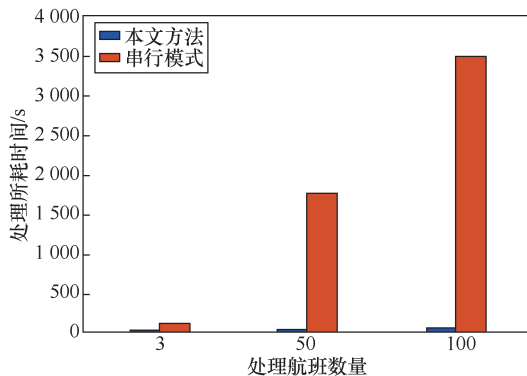


图 12 采用传统串行模式处理相同规模航迹和本文方法效率对比

实验结果表明,本文方法对各自数据源的航迹记录有一定的误差纠正作用,在任意数据源出现数据缺失或者质量不佳的状况时,依然能够很好地融合二次雷达与 ADS-B 的航迹数据,保证输出融合航迹的连续、稳定。航迹的处理流程效率较高,每条航迹处理所耗时间在可接受范围内,能够保障实时性的要求。本文方法的处理流程还

着重考虑了并行性能,有效提升了系统的并发能力,同时具备良好的伸缩性,在处理多架航班时,通过扩展硬件资源,可以方便地应对大规模的处理需求。

5 结束语

本文提出了一种二次雷达与 ADS-B 航迹实时融合的方法,针对现有研究和工程实现中对处理大规模航迹带来的计算压力关注度不足的问题,借鉴了互联网领域中较成熟的大数据技术,综合使用了分布式消息队列、微批处理框架、MapReduce 编程模型,设计了包括航迹建模和拟合、考虑决定系数的加权融合的处理流程,并通过实验验证了方法的可行性与有效性。相比于传统的以串行为主的处理流程,本文方法使用的以并行为主的处理流程能够显著地提升系统的并发能力,显著增强了系统处理大规模航班监视流式数据的效率,是应我国民航行业快速发展、航班监视需求“井喷式”增加的亟需技术。本文方法的整体框架也适用于航班航迹实时预测、流量信息实时统计、航班异常航迹实时判断等基于历史数据进行实时计算的场景,有助于进一步提升空中交通管制自动化、智能化的水平。

参考文献:

- [1] 中国民用航空局. 2018 年民航行业发展统计公报[R]. 2019. CAAC. Civil aviation industry development statistical bulletin 2018[R]. 2019.
- [2] 前瞻产业研究院. 2019 年全年中国民航业市场分析: 运输飞机数量接近 4 000 架, 营业收入突破万亿元[R]. 2020. Forward Business Information Co., Ltd., Shenzhen. Market analysis of China civil aviation industry for 2019: the number of transport aircrafts is close to 4 000, operating revenue exceeds trillion yuan[R]. 2020.
- [3] 王治. 空中交通管理自动化系统的设计与实现[D]. 厦门: 厦门大学, 2015. WANG Y. Design and implementation of air traffic management automation system[D]. Xiamen: Xiamen University, 2015.
- [4] 王洪, 刘昌忠, 汪学刚. 二次雷达 S 模式综述[J]. 电讯技术, 2008, 48(7): 113-118. WANG H, LIU C Z, WANG X G. Mode S for secondary surveillance radar (SSR): an introduction and overview[J]. Telecommunication Engineering, 2008, 48(7): 113-118.
- [5] ZHANG J, LIU W, ZHU Y B. Study of ADS-B data evaluation[J]. Chinese Journal of Aeronautics, 2011, 24(4): 461-466.
- [6] 孙沂, 吴仁彪. 空管自动化系统的多雷达与 ADS-B 数据融合技术综述[C]//第二十五届中国(天津)2011 IT、网络、信息技术、电子、仪器仪表创新学术会议. 2011: 19-23. SUN Y, WU R B. The data fusion of multi-radar and ADS-B in ATC

- automation system[C]//The 25th China (Tianjin) 2011'IT, Network, Information Technology, Electronics, Instrumentation Innovation Conference. 2011: 19-23.
- [7] 张军. 空域监视技术的新进展及应用[J]. 航空学报, 2011, 32(1): 1-14.
ZHANG J. New development and application of airspace surveillance technology[J]. Acta Aeronautica et Astronautica Sinica, 2011, 32(1): 1-14.
- [8] BAUD O, HONORE N, TAUPIN O. Radar/ADS-B data fusion architecture for experimentation purpose[C]//2006 9th International Conference on Information Fusion. IEEE, 2006: 1-6.
- [9] 白松浩. 多雷达与 ADS 数据融合的可变周期更新算法[J]. 交通运输工程学报, 2007, 7(2): 19-23.
BAI S H. Variable period updating algorithm of multiradar and ADS data fusion[J]. Journal of Traffic and Transportation Engineering, 2007, 7(2): 19-23.
- [10] 刘伟, 黄智刚, 张军, 等. 星基 ADS 与雷达误差校准算法的研究[J]. 航空学报, 2006, 27(1): 120-124.
LIU W, HUANG Z G, ZHANG J, et al. Studies on satellite-based ADS and radar registration algorithm[J]. Acta Aeronautica et Astronautica Sinica, 2006, 27(1): 120-124.
- [11] 周雷, 辛晓娜, 陈川波. 结合 ADS-B 的航管监视数据融合关键技术[J]. 计算机工程与应用, 2013, 49(14): 231-235.
ZHOU L, XIN X N, CHEN C B. Key techniques of ATC surveillance data fusion with ADS-B[J]. Computer Engineering and Applications, 2013, 49(14): 231-235.
- [12] 董宇涵, 山秀明, 任勇. 一种高性能航迹融合反馈系统的设计与实现[J]. 微计算机信息, 2005, 21(12S): 145-147.
DONG Y H, SHAN X M, REN Y. A high performance track fusion system with feedback[J]. Microcomputer Information, 2005, 21(12S): 145-147.
- [13] 翟羽佳. 多雷达信号与 ADS-B 数据融合技术浅析[J]. 空中交通, 2017: 27.
ZHAI Y J. A brief discussion on data fusion technology of multi-radar signal and ADS-B signal[J]. Air Traffic, 2017: 27.
- [14] 许文君. 空管自动化系统及数据融合方法研究[D]. 南京: 南京邮电大学, 2020.
XU W J. Research on air traffic control automation system and data fusion method[D]. Nanjing: Nanjing University of Posts and Telecommunication, 2020.
- [15] 张杰, 王丽娜, 赵媛, 等. 基于 ZeroMQ 消息通讯的多源空中目标跟踪处理平台设计[J]. 计算机测量与控制, 2018, 26(9): 219-222.
ZHANG J, WANG L N, ZHAO Y, et al. Design of multi-source tracking platform for air target based on ZeroMQ communication library[J]. Computer Measurement & Control, 2018, 26(9): 219-222.
- [16] 李小智, 陶勇. 基于消息中间件的航班信息显示系统的设计与实现[J]. 计算机系统应用, 2010, 19(10): 7-11.
LI X Z, TAO Y. Design and implementation of flight information display system based on message-oriented middleware[J]. Computer Systems & Applications, 2010, 19(10): 7-11.
- [17] 李文道, 杨小虎. 基于分布式缓存的消息中间件存储模型[J]. 计算机工程, 2010, 36(13): 93-95.
LI W X, YANG X H. Storage model based on distributed cache for message oriented middleware[J]. Computer Engineering, 2010, 36(13): 93-95.
- [18] JOHN V, LIU X. A survey of distributed message broker queues[J]. arXiv: 1704.00411, 2017.
- [19] 陈纯. 流式大数据实时处理技术、平台及应用[J]. 大数据, 2017, 3(4): 1-8.
CHEN C. Real-time processing technology, platform and application of streaming big data[J]. Big Data Research, 2017, 3(4): 1-8.
- [20] CAI Y J, WU B, ZHANG X W, et al. Flow identification and characteristics mining from Internet traffic with Hadoop[C]//2014 International Conference on Computer, Information and Telecommunication Systems (CITS). IEEE, 2014: 1-5.
- [21] VAIDYA M, DESHPANDE S. Study of Hadoop-based traffic management system[J]. IJCA Proceedings of ICRTITCS, 2013, 3: 38-42.
- [22] SAFAEI A A. Real-time processing of streaming big data[J]. Real-Time Systems, 2017, 53(1): 1-44.
- [23] SHAHRIVARI S. Beyond batch processing: towards real-time and streaming big data[J]. Computers, 2014, 3(4): 117-129.
- [24] DEAN J, GHEMAWAT S. MapReduce: simplified data processing on large clusters[J]. Communications of the ACM, 2008, 51(1): 107-113.
- [24] DEAN J, GHEMAWAT S. MapReduce[J]. Communications of the ACM, 2008, 51(1): 107-113.
- [25] 李建江, 崔健, 王聘, 等. MapReduce 并行编程模型研究综述[J]. 电子学报, 2011, 39(11): 2635-2642.
LI J J, CUI J, WANG D, et al. Survey of MapReduce parallel programming model[J]. Acta Electronica Sinica, 2011, 39(11): 2635-2642.

[作者简介]

张曩熹, 男, 博士, 主要研究方向为综合交通信息化及大数据技术、航空交通数据挖掘分析等。

田旺, 男, 北京航空航天大学硕士生, 主要研究方向为交通大数据、航迹融合等。

朱少川, 男, 北京航空航天大学博士生, 主要研究方向为机器学习、航迹模式挖掘、航迹预测等。

刘洪岩, 男, 北京航空航天大学硕士生, 主要研究方向为交通大数据、空域态势计算与调控等。

朱熙, 男, 博士, 北京航空航天大学助理研究员、硕士生导师, 主要研究方向为交通大数据、空域态势计算与调控等。