



中国计算机学会会刊
Publication of China Computer Federation

ISSN 1000-7024
CODEN JIGYEP

计算机工程与设计(月刊)

计算机工程与设计

Computer Engineering and Design

2022年4月 第43卷 第4期

Vol. 43 No. 4 Apr. 2022

第四十三卷 第四期

4

2022

二〇二二年四月

中国航天科工集团第二研究院 706 所主办

计算机工程与设计

(JISUANJI GONGCHENG YU SHEJI)

2022 年 4 月 第 43 卷 第 4 期

目 次

信息安全技术

基于 Q 学习与贝叶斯博弈的物联网安全	李 洋, 刘天莺, 朱建明, 王秀利(901)
基于属性分类建模的入侵检测方法	王加梁(907)

网络与通信技术

NFV 系统跨数据中心网络连接管理机制	石 峰, 蔺婧娜(914)
基于麻雀算法的 Web 服务组合	杨宏伟, 薛富城, 李 莉(922)
基于段路由的 IPv6 网络优化算法	刘 威, 黄 萍, 孙凤杰(930)

软件与算法

基于 SURF 的改进 FLANN 匹配算法	张志敏, 李 彬, 田联房, 丁焕文(941)
改进相位补偿结合谐波重构的语音增强方法	崔 磊, 马建芬, 张朝霞(949)
基于集成学习的风险预测模型研究与应用	彭 岩, 马 铃, 张文静, 李 晓, 郭莹莹(956)
改进的蚁群优化算法求解旅行商问题	赵 鑫, 杨雄飞, 钱育蓉(962)
基于特征正则稀疏关联的无监督特征选择方法	白圣子, 降爱莲(969)
采用模糊 k 核多粒度分解机制的高效社区发现	李宏平, 刘 群(977)
改进的社交网络谣言传播模型	吕心怡, 黄贤英(986)
基于混合深度神经网络的就业推荐方法	张 娅, 彭海英(995)

基于 W 距离自编码器半监督生成模型.....	王江晴, 何开杰, 孙 翀, 帖 军, 尹 帆(1002)
基于自校准卷积网络的行人检测方法	强 华, 李琦铭, 周勇军, 高 骁, 李 波, 李 俊(1008)
基于单目视觉的车辆前方障碍物测距方法	高维岳, 陈宇拓, 刘 洋, 陈 标(1016)
基于改进 Faster RCNN 的小尺度铁路侵限算法.....	余志强, 张 明(1023)
面向众测的 GUI 软件操作序列记录方法	黄 松, 王廷永, 孙金磊, 阳 真, 洪 宇(1032)
基于 BP 神经网络的离港航班滑出时间预测	黄龙杨, 夏正洪(1039)
基于改进高分辨表征的人体姿态估计算法	宋玉琴, 曾贺东, 高师杰, 熊高强(1045)

多媒体技术

基于多任务学习的蜡染纹样图案检索方法	邹 悦, 潘伟杰, 吕 健, 方年丽, 岳 迪, 朱姝蔓(1052)
基于编解码器生成对抗网络的 CT 去噪	周广宇, 张鹏程, 刘生富, 刘 玮, 桂志国(1059)
基于光照影响因子的动态 Niblack 算法研究及应用	魏兴凯, 蒋 峥, 傅呈勋, 刘 斌(1066)
基于多尺度条件生成对抗网络的图像去模糊	王晨卿, 荆 涛, 刘云鹏, 陈 颂, 蒋晓瑜, 闫兴鹏(1074)
基于自注意力的多模态 LSTM 的动作预测	莫 晨, 邵 洁(1083)

智能技术

基于 Reformer 模型的文本情感分析	王 珊, 黄海燕, 乔伟涛(1089)
基于孪生网络的车牌检测与跟踪算法	邹承明, 纵耘博(1096)
基于改进果蝇优化算法的投诉举报文本分类.....	范青武, 陈 光, 杨 凯(1103)
自适应重加权池化深度多任务学习的表情识别	王晓峰, 王 昆, 刘 轩, 郝 潇(1111)
基于分层式 CNN 的长文本情感分类模型	徐逸舟, 林 晓, 陆黎明(1121)
基于卷积 LSTM 模型的航空器轨迹预测.....	刘龙庚, 翟俐民, 韩云祥(1127)
利用改进型 VGG 标签学习的表情识别方法	程学军, 邢萧飞(1134)
基于集成学习的离港航班延误预测方法	罗 杰, 侯 霞, 杨鸿波, 刘 林, 谢 丰(1145)
龙格库塔程序的似然蜕变关系识别方法	文双红, 阳小华, 闫仕宇, 刘 杰, 李 萌, 冯晋涛(1152)

融合交叉熵损失的 3DCNN 探水作业动作识别 刘春霞, 高 强, 潘理虎, 龚大立(1160)

基于骨骼时空图的人体行为识别 谷学静, 沈 攀, 刘海望, 郭 俊, 位占锋(1166)

基于局部特征融合的细粒度车辆识别 张晶晶, 雷景生(1173)

基于深度学习的矩阵分解推荐模型 纪佳琪, 蔡永华, 郭景峰(1179)

开发与应用

可实时心率检测的远程心音采集系统设计 熊 艾, 赵征鹏, 杨宏波, 潘家华, 孙 柯, 王威廉(1186)

基于重构设计的 Inception 网络 谢晓燕, 杜卓林, 胡传瞻, 杨 坤, 王安琪(1195)

编 辑 部
主 编 汤铭端
编 审 何文思
责任编辑 韩 庭

本刊为：

★中国计算机学会会刊	★《万方数据—数字化期刊群》收录期刊
★航天科工集团期刊集群成员期刊	★《中国学术期刊文摘(中文版)》收录期刊
★全国中文核心期刊(中文核心期刊要目总览 2020 版)	★《电子科技文摘数据库》收录期刊
★中国科技核心期刊(中国科技论文统计源期刊)	★英国《科学文摘》(SA, INSPEC)源期刊
★世界期刊影响力指数(WJCI)报告收录期刊	★俄罗斯《文摘杂志》(AJ of VINITI)源期刊
★《中国学术期刊综合评价数据库》收录期刊	★美国《乌利希期刊指南》(UPD)源期刊
★《中国学术期刊网络出版总库(CNKI)》收录期刊	★波兰《哥白尼索引》(IC)源期刊

Computer Engineering and Design

Vol. 43 No. 4 Apr. 2022

CONTENTS

Information Security Technology

- Security of internet of things based on Q-learning and Bayesian game LI Yang, LIU Tian-ying, ZHU Jian-ming, et al(901)
- Intrusion detection method based on attribute classification modeling WANG Jia-liang(907)

Network and Communications Technology

- NFV system cross-data center network connection management mechanism SHI Feng, LIN Jing-na(914)
- Web service composition based on sparrow search algorithm YANG Hong-wei, XUE Fu-cheng, LI Li(922)
- IPv6 network optimization algorithm based on segment routing LIU Wei, HUANG Ping, SUN Feng-jie(930)

Software and Arithmetic

- Improved FLANN matching algorithm based on SURF ZHANG Zhi-min, LI Bin, TIAN Lian-fang, et al(941)
- Speech enhancement method for improving phase compensation combined with harmonic reconstruction
..... CUI Lei, MA Jian-fen, ZHANG Zhao-xia(949)
- Research and application of risk forecast model based on ensemble learning
..... PENG Yan, MA Ling, ZHANG Wen-jing, et al(956)
- Improved ant colony optimization algorithm for TSP ZHAO Xin, YANG Xiong-fei, QIAN Yu-rong(962)
- Unsupervised feature selection method based on feature regularized sparse association BAI Sheng-zi, JIANG Ai-lian(969)
- Efficient community detection using fuzzy k-core multi-granularity decomposition mechanism LI Hong-ping, LIU Qun(977)
- Improved social network rumor spreading model LYU Xin-yi, HUANG Xian-ying(986)
- Job recommendation method based on hybrid deep neural network ZHANG Hua, PENG Hai-ying(995)

Semi-supervised generation model of autoencoder with Wasserstein distance	WANG Jiang-qing, HE Kai-jie, SUN Chong, et al(1002)
Pedestrian detection based on self-calibration convolution network	QIANG Hua, LI Qi-ming, ZHOU Yong-jun, et al(1008)
Distance measurement method based on monocular vision for obstacles in front of vehicles	GAO Wei-yue, CHEN Yu-tuo, LIU Yang, et al(1016)
Improved Faster RCNN based railway intrusion detection algorithm for small objects	YU Zhi-qiang, ZHANG Ming(1023)
Crowdsourced GUI software operation sequence recording method	HUANG Song, WANG Ting-yong, SUN Jin-lei, et al(1032)
Taxi-out time prediction of departure flight based on BP neural network	HUANG Long-yang, XIA Zheng-hong(1039)
Human pose estimation algorithm based on improved high resolution representation	SONG Yu-qin, ZENG He-dong, GAO Shi-jie, et al(1045)

Multimedia Technology

Batik pattern retrieval method based on multi-task learning	ZOU Yue, PAN Wei-jie, LYU Jian, et al(1052)
CT image denoising using generative adversarial network with residual encoder-decoder generator	ZHOU Guang-yu, ZHANG Peng-cheng, LIU Sheng-fu, et al(1059)
Research and application of dynamic Niblack algorithm based on illumination impact factor	WEI Xing-kai, JIANG Zheng, FU Cheng-xun, et al(1066)
Image deblurring based on multi-scale conditional generative adversarial network	WANG Chen-qing, JING Tao, LIU Yun-peng, et al(1074)
Self-attention based multi-modal LSTM for action prediction	MO Chen, SHAO Jie(1083)

Intelligent Technology

Chinese sentiment analysis based on Reformer model	WANG Shan, HUANG Hai-yan, QIAO Wei-tao(1089)
License plate detection and tracking algorithm based on siamese network	ZOU Cheng-ming, ZONG Yun-bo(1096)
Complaints and reports text classification based on improved fruit fly optimization algorithm	FAN Qing-wu, CHEN Guang, YANG Kai(1103)

Expression recognition based on adaptive reweighting pooling deep multi task learning	WANG Xiao-feng, WANG Kun, LIU Xuan, et al(1111)
Long text sentiment classification model based on hierarchical CNN	XU Yi-zhou, LIN Xiao, LU Li-ming(1121)
Aircraft trajectory prediction based on convLSTM	LIU Long-geng, ZHAI Li-min, HAN Yun-xiang(1127)
Expression recognition method using improved VGG tag learning	CHENG Xue-jun, XING Xiao-fei(1134)
Prediction of departure flight delay based on ensemble learning	LUO Jie, HOU Xia, YANG Hong-bo, et al(1145)
Identification method for likely metamorphic relations of Runge-Kutta program	WEN Shuang-hong, YANG Xiao-hua, YAN Shi-yu, et al(1152)
Action recognition for water exploration based on 3DCNN combined with cross-entropy error	LIU Chun-xia, GAO Qiang, PAN Li-hu, et al(1160)
Human behavior recognition based on skeletal spatiotemporal map	GU Xue-jing, SHEN Pan, LIU Hai-wang, et al(1166)
Fine-grained vehicle recognition based on fusion of part feature	ZHANG Jing-jing, LEI Jing-sheng(1173)
Matrix factorization recommender system based on deep learning	JI Jia-qi, CAI Yong-hua, GUO Jing-feng(1179)

Developments and Appications

Remote heart sound acquisition system with real-time heart rate detection	XIONG Ai, ZHAO Zheng-peng, YANG Hong-bo, et al(1186)
Reconfigurable design of Inception network	XIE Xiao-yan, DU Zhuo-lin, HU Chuan-zhan, et al(1195)

基于卷积 LSTM 模型的航空器轨迹预测

刘龙庚¹, 翟俐民², 韩云祥²⁺

(1. 中国软件评测中心, 北京 100089;

2. 四川大学 视觉合成图形图像技术国防重点学科实验室, 四川 成都 610041)

摘要: 采集空管大数据, 根据空管大数据的特点, 对数据进行数据融合, 利用改进的聚类算法处理航迹数据, 对得到的各类航迹数据分别构建模型, 提高模型的预测精度。分别构建 Stack LSTM 和基于卷积 LSTM 的航空器轨迹预测模型, 以真实雷达数据为例进行仿真实验, 对仿真结果进行对比, 其结果表明, 基于卷积 LSTM 的航空器轨迹预测模型可以将预测的均方根误差控制在 400 s 内, 验证了预测模型可以实现航空器轨迹的精确预测。

关键词: 空管大数据; 智能交通; 航迹聚类; 长短期记忆网络; 航迹预测; 时间序列; 深度学习

中图法分类号: TP311.13 **文献标识号:** A **文章编号:** 1000-7024 (2022) 04-1127-07

doi: 10.16208/j.issn1000-7024.2022.04.031

Aircraft trajectory prediction based on convLSTM

LIU Long-geng¹, ZHAI Li-min², HAN Yun-xiang²⁺

(1. China Software Test Center, Beijing 100089, China;

2. National Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu 610041, China)

Abstract: The air traffic control big data were collected. According to the characteristics of the air traffic control big data, data fusion was performed, and the improved clustering algorithm was used to process the track data. Models were constructed for the obtained various track data respectively, which improved the prediction accuracy of the model. Stack LSTM and convolution LSTM-based aircraft trajectory prediction models were built respectively, real radar data were taken as an example to conduct simulation experiments, and the simulation results were compared. The results show that the aircraft trajectory prediction model based on convolution LSTM can reduce the predicted root mean square error, which is controlled within 400 s, verifying that the prediction model can achieve accurate prediction of the aircraft trajectory.

Key words: air traffic management big data; intelligent transportation; track clustering; LSTM; track prediction; time series; deep learning

0 引言

随着计算机技术不断发展, 各行业的信息量急剧增长, 一般将规模巨大且无法通过常规软件工具在合理时间内管理和处理的信息称为大数据, 目前业界普遍认为大数据的特征可以归纳为 4 个“V”-Volume (大量)、Velocity (高速)、Variety (多样)、Value (价值低)^[1]。传感器数据是大数据的主要来源之一, 对传感器网络的大数据分析和应用, 是信息技术的发展趋势和重点研究方向^[2], Liu L 等深入研究了大数据环境下传感器网络基础技术^[3-5]。空中交通

管理过程中需要不断收集航空器相关传感器数据, 天然具有大数据的特点, 除了传统的大数据特征之外, 还具有以下特点^[6]: ①体量巨大, 空管一年产生的数据在 PB 级别以上; ②种类繁多, 在空管过程中涉及到多种数据, 包含静态数据、动态数据、结构化数据、非结构化数据; ③来源真实, 空管数据直接采集自空管一线, 来源真实可靠。

在空管大数据开发方面, 美国处于领先地位^[7]。国内在空管大数据开发方面也做出了一些探索。空管大数据可实现空管运行过程智能化精细化, 空管大数据可用于空管流程优化和空管风险管理与评估, 空管大数据可直接用于

收稿日期: 2021-12-15; **修订日期:** 2022-02-25

基金项目: 四川大学专职博士后研发基金项目 (2021SCU12048); 四川省科技计划基金项目 (2022YFG0180)

作者简介: 刘龙庚 (1975-), 男, 四川成都人, 博士, 高级工程师, 研究方向为产业研究、信创规划、信息安全、架构设计与测试评估等; 翟俐民 (1996-), 男, 山东淄博人, 硕士, 研究方向为计算机仿真; 通讯作者: 韩云祥 (1985-), 男, 河南驻马店人, 博士, 副研究员, 研究方向为交通系统优化控制。E-mail: hannyuaa@126.com

空管仿真模型的构建,并据此实现航空器运行过程预测,准确、可靠的空中交通轨迹预测模型(TBO)可以有效提高空中交通运行效率并进一步保障运行安全。

目前关于航迹预测的方法主要有基于航空器性能参数、运动学模型的传统方法以及基于数据挖掘和混合机器学习的研究方法。石庆研等^[8]提出了一种在线更新 LSTM 网络的短期航迹预测算法;李旭娟等^[9]提出了一种自动生成的条件变分自动编码器,以编码-解码的形式直接对未来一段时间的航迹进行预测;Xu 等^[10]提出了一种新的 Social-LSTM 模型,为空域内的每个航空器建立 LSTM 网络,并通过一个池化层整合关联航空器的状态,从而编码了航空器之间的相互作用;马兰等^[11]基于 ADS-B 数据挖掘和统计分析,并使用 CURE 聚类分析;Shi 等^[12]提出了一种基于 LSTM 的航迹预测模型,该模型解决长期依赖问题,避免了相邻序列状态的动态依赖性,有助于提高精度;Gallego 等^[13]针对目前的航迹预测方法缺少对于空中交通中不确定因素的理解,提出了融入航空器周围交通状况的航迹预测方法;Liu 等^[14]提出了一种基于树的高效匹配算法从气象数据中构建特征图,并构建了一个端到端的卷积神经网络,该网络包含一个 LSTM 编码网络和混合密度的 LSTM 解码网络,进一步从历史飞行轨迹中学习时空相关性;Shi 等^[15]提出了一种 4D 航迹预测模型,该模型长短期记忆网络结合滑动窗口来维持长期特征并预测轨迹,并将环境信息作为输入特征;Zeng 等^[16]针对目前的空域状态轨迹方法在终端区这样复杂的空域中效果不好的问题,提出了一种 seq2seq 的深度长短期记忆网络预测模型,可以有效地捕获长期和短期的时间依赖;Giuliani 等^[17]将自然语言处理领域的热门模型 Transformer 应用在航迹预测中,并分别实现了原始的 Transformer 网络和 BERT,实验结果表明 Transformer 在数据存在缺失的情况下表现更好;本文基于空管大数据分析挖掘,利用聚类算法萃取航迹数据,并构建 LSTM 网络航空器轨迹,使用航迹信息进行仿真验证,实现对航空器轨迹的准确预测。

1 数据处理

根据数据的来源的组织形式,可将空管大数据初步划分为静态数据、动态数据、结构化数据、非结构化数据^[6],其中静态数据是在一定时间段内保持稳定的数据,动态数据与静态数据相反,主要包含了航空器运行时数据。本文采集了动态的航空器轨迹数据、机场天气数据和静态的进场标准仪表进港程序用于构建航空器轨迹预测模型。通过一定渠道获取双流国际机场 2019 年 6 月至 2020 年 12 月每天进离岗航班相关数据^[18],包含解析标准 ADS-B 航迹数据以及二次雷达数据,编写爬虫收集双流机场对应日期的 METAR 天气数据报。对于从多种渠道获取的原始航迹数据及气象数据不能直接作为训练数据集,需经过以下处理

流程:数据清洗、多源数据融合、离散特征编码、模型数据集构建。数据清洗是剔除航迹点缺失过多的航迹。多源数据融合是指将来自不同数据源的轨迹数据及气象数据进行整合,使得航迹数据特征更加丰富。数据编码是指将特征中离散数据通过特征规则化,使其满足神经网络模型输入格式,让机器读懂该特征的含义。数据预处理过程如图 1 所示。

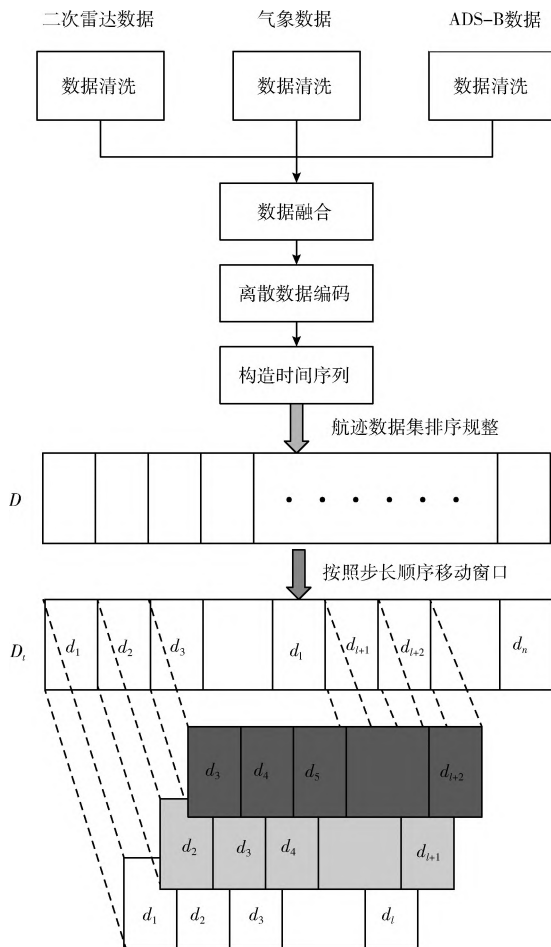


图 1 数据预处理过程^[19]

ADS-B 以及二次雷达中的坐标数据为经纬度,为减少后续计算量,将航迹中的经纬度(B,L)通过墨卡托投影,转换为以双流国际机场为坐标原点的坐标系(X,Y)。坐标系转换关系如式(1)~式(3)所示

$$X = K \times \ln \left[\operatorname{tg} \left(\frac{\pi}{4} + \frac{B}{2} \right) \times \left(\frac{1 - e \sin B}{1 + e \sin B} \right)^{\frac{e'}{2}} \right] \quad (1)$$

$$Y = K \times (L - L_0) \quad (2)$$

$$K = \frac{\frac{a^2}{b}}{\sqrt{1 + e'^2 \cdot \sin^2(B_0)}} \cdot \cos(B_0) \quad (3)$$

其中,标准纬度 B_0 ,标准经度 L_0 , e 为第一偏心率, e' 为第二偏心率, a 为长半轴长, b 为短半轴长。使用一定对应规则将采集的 ADS-B 航迹数据、二次雷达数据、机场 METAR

数据进行融合, 作为完整的训练集。ADS-B 航迹数据与二次雷达数据通过飞行日期 date 和航空器呼号 ID 匹配, 在确定为同一航班之后, 使用时间戳插值的方法构造数据集, 对于过程中差异较大的数据使用均值操作, 如式 (4) 所示

$$\begin{cases} x_{i,j} = \frac{x_{ndar} + x_{ds-b}}{2} \\ y_{i,j} = \frac{y_{ndar} + y_{ds-b}}{2} \end{cases} \quad (4)$$

上式中, (x_{ndar}, y_{ndar}) 为航迹点 $p_{(i,j)}$ 记录的二次雷达位置坐标, (x_{ds-b}, y_{ds-b}) 为航迹点 $p_{(i,j)}$ 记录的 ADS-B 坐标, $(x_{i,j}, y_{i,j})$ 为融合后的航迹坐标。最后为了消除不同特征之间量纲不同导致模型训练困难的问题, 使用归一化和标准化对数值型特征进行处理, 方法如式 (5) ~ 式 (6) 所示

$$x' = \frac{x - \min(X)}{\max(X) - \min(X)} \quad (5)$$

$$\hat{x} = \frac{x - \mu(X)}{\sigma(X)} \quad (6)$$

其中, X 为特征数据集, x 为原始特征数据, x' 为归一化后的特征数据, $\max(X)$ 、 $\min(X)$ 分别为该特征中的最大特征值与最小特征值, $\mu(X)$ 、 $\sigma(X)$ 分别为该特征中的期望与方差。

2 航迹聚类

作为一种无监督的机器学习方法, 聚类在诸多领域应用广泛。聚类将具有类似特征的对象使用划分标准分为不同类别, 目前常用的划分标准有欧式距离、马氏距离、曼哈顿距离等。在航迹聚类中, 需要将类似的飞行轨迹划分为同一类, 从最初选定的航迹簇开始, 通过距离度量选定相近的航迹簇, 不断合并航迹簇后重新计算航迹簇间距离, 最终获得聚类结果。在计算的过程中, 需要使用航迹间距离衡量航迹相似性, 以判断是否将航迹簇合并。由于设备误差等现实原因, 每条航迹长度通常不等, 故无法使用常规的距离度量算法。动态时间规整 (dynamic time warping, DTW) 通过计算航迹点之间的最短距离, 能够计算出航迹间的距离。针对两条航迹, 其计算过程是将其航迹点分别作为行列, 排列为一个二维表, 并使用 DTW 公式计算航迹点间距离, 从二维表的左上角到右下角寻找一条距离最短的路径, 将这条路径上的距离和作为两条航迹间的相似度量。将训练样本中的所有航迹, 两两计算航迹间的相似性, 得到一个二维航迹矩阵, 其大小为航迹数量, 对于航迹集合 T , 通过 DTW 算法计算得到的航迹距离矩阵如式 (7) 所示

$$\begin{bmatrix} d(\tau_1, \tau_1) & \cdots & d(\tau_1, \tau_j) & \cdots & d(\tau_1, \tau_J) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d(\tau_j, \tau_1) & \cdots & d(\tau_j, \tau_j) & \cdots & d(\tau_j, \tau_J) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ d(\tau_J, \tau_1) & \cdots & d(\tau_J, \tau_j) & \cdots & d(\tau_J, \tau_J) \end{bmatrix} \quad (7)$$

其中, $d(\tau_j, \tau_J)$ 为第 j 条航迹与第 J 条航迹之间的相似度。在计算距离矩阵时, 使用快速 DTW, 其使用约束搜索空间的策略, 可以降低时间复杂度, 加速计算过程。

使用航迹聚类, 将航迹分类并在此基础上建立不同预测模型, 可以有效提高航迹预测的精度。本文构建基于快速动态时间规整距离度量的并行剪枝层次聚类算法。在计算航迹相似度中, 使用 DTW 算法; 使用批次运算更新距离矩阵; 及时剪枝远离聚类中心的离群航迹。具体流程: ①计算航迹距离矩阵。对于训练集中的所有航迹, 数量为 n_t , 计算两两之间的相似度, 得到 $n_t \times n_t$ 的距离矩阵, 相似度计算使用 DTW 算法; ②计算航迹簇之间的相似度; ③遍历相似性矩阵中的航迹, 获取其中距离最近的航迹, 形成航迹簇对; ④计算航迹簇对之间的簇间距离, 比较与设定的簇间距离, 小于预设则剔除, 大于预设则为一个类别并重新计算所有航迹簇的距离矩阵, 进行下一步迭代。使用预先设定的航迹数目阈值来判定航迹簇是否为离群航迹簇, 当其中航迹数量大于阈值时, 当前航迹簇成为一个航迹类别, 否则该航迹簇为离群航迹簇, 将被剔除。上诉过程为一次聚类算法的迭代过程, 重新计算航迹簇的距离矩阵进入下一轮迭代。计算过程前需要设定两个关键参数, 分别是离群航迹距离和离群航迹数目。

3 模型构建

航迹序列是一种时间序列, 目前常用 LSTM 处理时间序列问题, 相对于 RNN 而言, LSTM 解决了其存在的长期依赖问题, 能够学习长期依赖关系。LSTM 的网络单元中每个重复模块由 3 个门组成, 分别是遗忘门、输入门和输出门, 每个门可以选择性地决定前一层的信息以什么样的方式传递给当前层, LSTM 网络单元结构如图 2 所示。

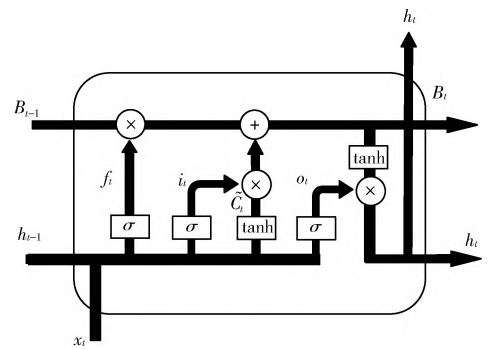


图 2 LSTM 网络单元结构

其中, σ 和 $\tanh$ 分别是 sigmoid 激活函数和 $\tanh$ 激活函数, x_t 为 t 时刻的输入, h_t 为 t 时刻的状态, i_t 为 t 时刻输入门输出, f_t 为 t 时刻的遗忘门输出, o_t 为 t 时刻输出门输出, 结合图 2, 对于上述定义如式

$$f_t = \text{sigmoid}(\alpha_f \cdot [h_{t-1}, x_t] + \beta_f) \quad (8)$$

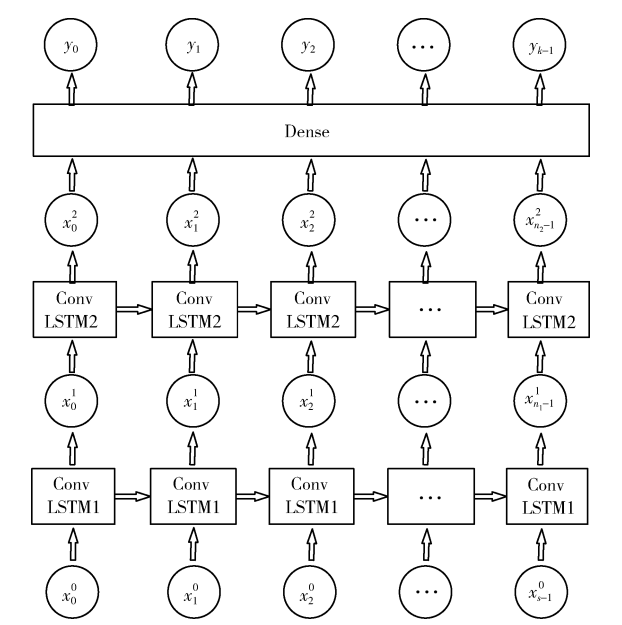


图 5 基于 ConvLSTM 航迹预测模型架构

方法剔除不合理数据后得到航迹数据集。再对航迹进行航迹聚类, 将数据集转换为聚类后数据。最后使用归一化和标准化将数据转换为标准模型输入。使用航迹中的三维位置坐标进行轨迹聚类, 将航迹数据中不同航路的航迹分类。使用墨卡托投影将原始航迹中的经纬度坐标转换为以机场坐标为中心的二维平面坐标值以便后续计算, 并对数据进行归一化处理以避免特征间量纲不同带来的影响。最终获得维度为 9405×3 的数据集。对于样本集中机型和分类结果是非数值类型, 对其使用 One-Hot 编码。训练样本的特征包含时间戳、飞行器三维坐标、航向、速度、垂直速度、机型和聚类结果。训练时使用滑动窗口移动数据, 将初始滑动窗口大小设置为 9, 即将航迹中前 9 个航迹点作为训练样本, 模型输出下一个航迹点后窗口向后滑动, 不断输出航迹点直到输出最后一个航迹点。使用水平误差、垂直误差和时间误差评估模型, 水平误差为预测航迹点 C 和真实航迹点 T 在二维平面下的欧式距离, 垂直误差是垂直方向上航空器的实际飞行高度和预测高度之前的距离差, 时间误差是指两条航迹中对应航迹点的时间差值, 计算公式如下所示

$$e_{horiz} = \sqrt{(x_C - x_T)^2 + (y_C - y_T)^2} \tag{19}$$

$$e_{vert} = |h_T - h_C| \tag{20}$$

$$e_{time} = |t_T - t_C| \tag{21}$$

其中, 水平误差以海里为单位 (nm), 垂直误差以英尺为单位 (ft)。使用平均绝对误差 (mean absolute error, MAE) 和均方根误差 (root mean squared error, RMSE) 衡量预测航迹和实际航迹之间的误差大小, 其计算公式如式 (22) ~ 式 (23) 所示

$$MAE = \frac{1}{m} \sum_{i=1}^m |P_i - R_i| \tag{22}$$

$$RMSE = \left[\frac{1}{m} \sum_{i=1}^m (|P_i - R_i|)^2 \right]^{\frac{1}{2}} \tag{23}$$

其中, P_i 为第 i 个航迹点某个特征的预测值, R_i 为第 i 个航迹点某个特征实际值。使用双流机场航班进港航迹作为测试数据集, 分析了 DTW、快速 DTW (FastDTW)、传统层次聚类 (HC), 以及加入剪枝的改进层次聚类 (IHC) 在不同航迹点数下的实验结果, 在时间开销上进行评价, 结果见表 1, 聚类结果可视图如图 6、图 7 所示。

表 1 聚类算法时间对比

方法	航迹点数	
	N=11540	N=23080
DTW+HC	833.31 s	3592.25 s
DTW+IHC	805.12 s	3531.25 s
FastDTW+HC	183.13 s	782.18 s
FastDTW+IHC	158.36 s	700.14 s

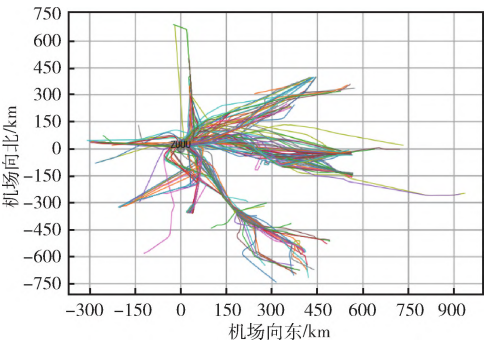


图 6 原始航迹二维平面

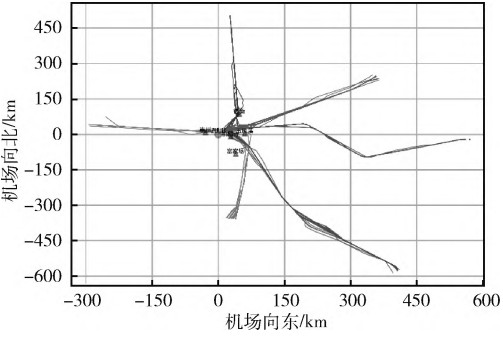


图 7 聚类后二维平面

通过实际结果可知, 该聚类方法将原始航迹分为 6 类, 可以有效分离数据, 将各类数据分别作为测试样本对航迹预测模型进行对比实验。在基于 LSTM 的预测模型中, 包含两层 LSTM 隐藏层, 单元个数均为 300, 为防止模型过拟合, 在 LSTM 隐藏层后使用比例大小为 0.2 的 Dropout,

滑动窗口大小设置为 9。基于 ConvLSTM 的预测模型中，将滑动窗口大小拆分为 3、1、3，在输入层之后，包含两个 ConvLSTM 的隐藏层，每一层设置卷积核大小为 (1,3)，通道数个数为 256。随机选取 60 条航迹作为测试集，数据集剩余部分作为训练集。以测试集中一条航迹为例，预测模型的评价指标计算过程如下：①通过模型不断迭代计算预测点，获得一条预测航迹；②通过航迹点之间最短欧式距离的方式匹配样本标签航迹中的航迹点，计算预测点与实际点之间的各项误差值；③得到单条航迹的各项误差后，使用 MAE 和 RMSE 计算所有测试集误差。实验结果见表 2。

表 2 误差分析对比

方法	平均绝对误差			均方根误差		
	e_{time}	e_{horiz}	e_{vert}	e_{time}	e_{horiz}	e_{vert}
LSTM	424	86.553	8099.64	497	99.743	9976.87
ConvLST	337	65.149	4842.74	395	73.164	6956.37

可以看出 ConvLSTM 预测模型的各项误差值均低于 LSTM 模型，其中，时间误差减少了 87 s，垂直误差减少了 21.404，水平误差减少了 3256.9，由此看出 ConvLSTM 模型预测效果更佳。

5 结束语

本文从空管大数据出发，使用数据挖掘聚类算法构建数据集，在聚类算法中使用 DTW 算法和剪枝加速聚类，实验结果显示聚类时间开销从 3592.25 s 下降到 700.14 s，构建 LSTM 预测模型和 ConvLSTM 预测模型，通过计算模型的 MAE 和 RMSE，结果显示 ConvLSTM 预测模型的各项误差值均低于 LSTM 模型，时间误差减少了 87 s，垂直误差减少了 21.404，水平误差减少了 3256.9。本文借助数据挖掘和神经网络等技术对 4D 航迹预测开展了研究工作，但仍然存在较多有待改进的不足之处，其主要包含：①目前数据源较为单一，随着民航大数据技术不断发展，可以从不同设备中获取航迹数据以供后续研究，后续工作可以获取多源航迹，利用大数据技术对数据进行融合，提供更加完备的数据集；②在预测中仅考虑历史航迹，并未将气象因素和空域态势融入模型，下一步可以将空域气象和空域态势输入模型，提高预测模型的鲁棒性和精确性。

参考文献：

[1] Wikipedia. The Free Encyclopedia. Big data [EB/OL]. [2022-01-20]. https://en.wikipedia.org/wiki/Big_data.
[2] LIU Longgeng. Research on key technologies of wireless sensor networks in big data environment [D]. Chengdu: University of Electronic Science and Technology of China, 2017: 1-136

(in Chinese). [刘龙庚. 大数据环境下无线传感器网络关键技术研究 [D]. 成都: 电子科技大学, 2017: 1-136.]
[3] Liu L, Luo G. Routing optimization in networks based on traffic gravitational field model [J]. International Journal of Modern Physics B, 2017, 31 (11): 1750074.
[4] Liu L, Luo G, Qin K, et al. An algorithm based on logistic regression with data fusion in wireless sensor networks [J]. EURASIP Journal on Wireless Communications and Networking, 2017 (1): 1-9.
[5] Liu L, Luo G, Qin K, et al. An on-demand global time synchronization based on data analysis for wireless sensor networks [J]. Procedia Computer Science, 2018, 129: 503-510.
[6] ZHANG Zhanbo. The concept, characteristics and application of air traffic control big data [J]. Journal of Civil Aviation Flight University of China, 2015 (6): 18-21 (in Chinese). [张战波. 空管大数据的概念, 特征和应用 [J]. 中国民航飞行学院学报, 2015 (6): 18-21.]
[7] ZHANG Chen, LING Fan. Research on the application of american air traffic control big data [J]. Science New Ground, 2017 (6): 94-96 (in Chinese). [张晨, 凌帆. 美国空管大数据应用研究 [J]. 创新科技, 2017 (6): 94-96.]
[8] SHI Qingyan, WANG Wenqing, HAN Ping. Short-term 4D track Prediction algorithm based on online updating LSTM network [J]. Journal of Signal Processing, 2021, 37 (1): 66-74 (in Chinese). [石庆研, 王文青, 韩萍. 基于在线更新 LSTM 网络的短期 4D 航迹预测算法 [J]. 信号处理, 2021, 37 (1): 66-74.]
[9] LI Xujuan, PI Jianyong, HUANG Feixiang, et al. 4D track prediction based on self-generated depth neural network [J]. Journal of Computer Applications, 2021, 41 (5): 1492-1499 (in Chinese). [李旭娟, 皮建勇, 黄飞翔, 等. 基于自生成深度神经网络的 4D 航迹预测 [J]. 计算机应用, 2021, 41 (5): 1492-1499.]
[10] Xu Z, Zeng W, Chu X, et al. Multi-aircraft trajectory collaborative prediction based on social long short-term memory network [J]. Aerospace, 2021, 8 (4): 115.
[11] MA Lan, GAO Yongsheng. 4D track prediction method based on ADS-B data mining [J]. Journal of Civil Aviation University of China, 2019, 37 (4): 1-4 (in Chinese). [马兰, 高永胜. 基于 ADS-B 数据挖掘的 4D 航迹预测方法 [J]. 中国民航大学学报, 2019, 37 (4): 1-4.]
[12] Shi Z, Xu M, Pan Q, et al. LSTM-based flight trajectory prediction [C] //International Joint Conference on Neural Networks, 2018: 1-8.
[13] Gallego C E V, Comendador V F G, Carmona M A A, et al. A machine learning approach to air traffic interdependency modelling and its application to trajectory prediction [J]. Transportation Research Part C: Emerging Technologies, 2019, 107: 356-386.

- [14] Liu Y, Hansen M. Predicting aircraft trajectories: A deep generative convolutional recurrent neural networks approach [J]. arXiv Preprint arXiv: 1812.11670, 2018.
- [15] Shi Z, Pan Q, Xu M. LSTM-Cubic A* -based auxiliary decision support system in air traffic management [J]. Neurocomputing, 2020, 391: 167-176.
- [16] Zeng W, Quan Z, Zhao Z, et al. A deep learning approach for aircraft trajectory prediction in terminal airspace [J]. IEEE Access, 2020, 8: 151250-151266.
- [17] Giuliani F, Hasan I, Cristani M, et al. Transformer networks for trajectory forecasting [C] //25th International Conference on Pattern Recognition, 2021: 10335-10342.
- [18] ZHAI Limin. Research on track prediction method based on 3D trajectory clustering and LSTM [D]. Chengdu: Sichuan University, 2021: 1-67 (in Chinese). [翟俐民. 基于三维轨迹聚类 and LSTM 的航迹预测方法研究 [D]. 成都: 四川大学, 2021: 1-67.]
- [19] ZHANG Yong. Research on predicting aircraft arrival time based on hybrid machine learning [D]. Chengdu: Sichuan University, 2020: 1-82 (in Chinese). [张勇. 基于混合机器学习预测航空器到达时间的研究 [D]. 成都: 四川大学, 2020: 1-82.]