

# 基于 DBSCAN-GRU 算法的终端区 4D 航迹预测

韩萍 张启 石庆研 张泽中

(中国民航大学智能信号与图像处理天津市重点实验室, 天津 300300)

**摘要:** 终端区空域环境复杂、航班密集,精确的航迹预测能极大地提高空中交通服务水平,保障航班飞行安全。针对终端区的高精度多航班 4D 航迹预测问题,本文提出了一种基于密度的带噪声空间聚类算法(Density-Based Spatial Clustering of Applications with Noise, DBSCAN)和门控循环单元(Gated Recurrent Unit, GRU)相结合的航迹预测方法,通过 DBSCAN 聚类,将终端区中航迹相近的航班聚类到一簇中,对每一簇航班建立基于 GRU 神经网络的航迹预测模型,对终端区航班进行预测时,先判断该航班属于哪一簇,然后采用与该簇对应的航迹预测模型,进行 4D 航迹预测。与仅研究单一航班的传统预测方法相比,本算法有效地利用了终端区的航迹数据,所建模型可以针对多架航班进行航迹预测,扩大了模型的适用范围,提高了航迹预测的预测精度。

**关键词:** 4D 航迹预测; 终端区; 基于密度的带噪声空间聚类; 门控循环单元

**中图分类号:** TP391.9 **文献标识码:** A **DOI:** 10.16798/j.issn.1003-0530.2023.03.007

**引用格式:** 韩萍,张启,石庆研,等. 基于 DBSCAN-GRU 算法的终端区 4D 航迹预测[J]. 信号处理,2023,39(3): 439-449. DOI: 10.16798/j.issn.1003-0530.2023.03.007.

**Reference format:** HAN Ping, ZHANG Qi, SHI Qingyan, et al. 4D trajectory prediction of terminal area based on DBSCAN-GRU algorithm[J]. Journal of Signal Processing, 2023, 39(3): 439-449. DOI: 10.16798/j.issn.1003-0530.2023.03.007.

## 4D Trajectory Prediction of Terminal Area Based on DBSCAN-GRU Algorithm

HAN Ping ZHANG Qi SHI Qingyan ZHANG Zezhong

(Tianjin Key Lab for Advanced Signal Processing, Civil Aviation University of China, Tianjin 300300, China)

**Abstract:** In the terminal area, the airspace environment is complex and the flights are dense. Accurate trajectory prediction can greatly improve air traffic service level, and ensure aviation safety. To solve the problem of multi flight and high-precision 4D trajectory prediction required by the terminal area, an algorithm that combines Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Gated Recurrent Unit (GRU) is proposed. Through DBSCAN, the flights with similar trajectory in the terminal area are clustered into a cluster, and then the GRU is used to train the trajectory prediction model for the trajectories of different clusters. When a flight enters the terminal area and needs to be predicted, first, the flight is judged which cluster belongs to, and then the trajectory prediction model corresponding to this cluster is used for 4D trajectory prediction. Compared with the traditional prediction method that only studies a single flight, this algorithm effectively uses the trajectory data in the terminal area. The built model can predict the trajectory of multiple flights, expand the scope of application of the model, and improve the prediction accuracy of the trajectory prediction.

**Key words:** 4D trajectory prediction; terminal area; density-based spatial clustering of applications with noise; gated recurrent unit

## 1 引言

智慧空管是我国智慧民航的四大模式形态(智慧出行、智慧空管、智慧机场、智慧监管)之一<sup>[1]</sup>。建立完善的四维航迹管理系统是建设智慧空管的工作重点之一,四维航迹运行是目前国际民航界普遍公认的下一代航行系统的核心运行理念,其中四维是指航空器的空间信息及时间信息。而四维航迹运行的前提是精确的航迹预测,精确的航迹预测在提高空中交通服务水平的同时,也可以保障航班飞行安全。根据航空先驱报(The Aviation Herald)数据统计,2020年民航业总共发生82起Accident级别事故,其中终端区范围内,即起飞、爬升、进近、降落及着陆阶段总计发生72起。根据上海国际机场股份有限公司2021年年度报告数据显示,浦东国际机场全年飞机起降架次约为34.9万架次,平均至每日约为958架航班,每小时40架次航班。针对单一航班进行研究的传统航迹预测无法满足高频次的航班起降要求,终端区需要更加精确、更加高效的多航班航迹预测方案。目前针对航迹预测的算法大致分为三类,分别是采用动力学模型进行航迹预测,利用状态估计方法建立模型进行航迹预测,和利用机器学习方法对航迹数据进行深度挖掘,并进行航迹预测。

采用动力学模型进行航迹预测,是根据航空器的飞行意图、航空器当前的运行状态以及自身的性能参数和需要预测的目标航迹的空域环境,对航空器进行动力学建模,完成对目标航空器的航迹预测。Zammit和Mangion将全能量模型用于在慢车推力下降过程中的快速航迹预测<sup>[2]</sup>。王超等人按飞行阶段分析航班的水平航迹与垂直剖面,然后基于全能量方程进行航迹预测<sup>[3]</sup>。

利用状态估计方法建立航迹预测模型,是将解决航班飞行过程这个物理系统的运行问题转换为数学模型中的状态转移问题。Mondoloni等人研究了改进卡尔曼滤波算法和自适应卡尔曼滤波算法在航迹预测问题上的应用<sup>[4-8]</sup>。Hwang等提出交互多模型(Interacting Multiple Model, IMM)算法和基于新似然函数的残差均值交互多模型(Residual-Mean Interacting Multiple Models, RMIMM)算法,将航迹预测问题视作随机线性混杂系统的估计问题<sup>[9]</sup>。

利用机器学习算法进行航迹预测,是将航迹预测问题转化为有序时间序列预测问题,通过挖掘航迹训练集中数据间关系来进行建模。吴志军等人

提出了一种基于误差反向传播(Back Propagation, BP)神经网络的4D精确航迹预测算法,解决了传统预测方法不能进行多维度的实时航迹预测的问题<sup>[10]</sup>。石庆研等人提出采用长短期记忆递归神经网络(Long Short-Term Memory, LSTM)模型进行航迹预测,考虑航迹数据实时性,引入在线更新概念,使用被预测航班的实时数据进行模型微调,使得模型对目标航迹具有更高的准确性<sup>[11]</sup>。陈玉立等人引入注意力机制对航迹赋予不同权重,利用生成对抗网络进行航迹预测<sup>[12]</sup>。翟俐民等人通过利用航迹聚类将大量整条航迹进行聚类后利用CNN LSTM和ConvLSTM算法进行航迹预测<sup>[13]</sup>。

虽然终端区范围内存在进离场程序,但是真实飞行时需要根据实际情况及管制员指挥进行飞行,部分航班无法严格按照进离场程序飞行。针对终端区航迹预测问题,李海峰通过对飞机进行运动分析和受力分析,建立了基于波音爬升程序的终端区4D航迹预测模型<sup>[14]</sup>。Shane T. Barratt等人利用K-means++聚类算法和高斯混合模型算法(Gaussian Mixture Model, GMM)对起飞阶段和降落阶段的航班进行分析及航迹预测<sup>[15]</sup>。在此基础上,Mayara Condé Rocha Murça等人利用DBSCAN算法聚类出目标终端区范围内进近及降落的主要航迹簇,然后利用GMM算法对降落航班进行航迹预测<sup>[16]</sup>。Hong-Cheol Choi等人利用GMM算法挖掘当前时刻的历史航迹数据,并结合动力学模型利用RMIMM算法预测未来航迹<sup>[17]</sup>。但上述论文所提的模型所需信息较多,复杂度较大。李楠使用基于遗传算法的反向传播神经网络模型,通过遗传算法训练反向传播神经网络中的最优参数,然后利用最优参数下的反向传播神经网络进行终端区目标航迹的预测<sup>[18]</sup>,但是所建模型仍然是针对单一航班,无法对终端区内多个航班进行航迹预测。

针对上述问题,本文利用DBSCAN算法对目标区域大量航班进行聚类,建立多个GRU模型,充分利用空管大数据,挖掘目标区域所有航迹数据特点,每个GRU模型均可解决该簇航班的航迹预测问题,不局限于单一航班,可以针对不同航班实现精准的航迹预测。

## 2 DBSCAN聚类算法

### 2.1 算法原理

目前聚类方法主要分为三类,分别是DBSCAN

为代表的基于密度的聚类算法<sup>[19]</sup>、以 K-means 为代表的划分式聚类方法,以及以 Agglomerative 为代表的层次化聚类方法。而针对于终端区航迹数据的高密度性和高复杂性,DBSCAN 可以很好地完成不同类别航迹的聚类,将终端区内的不同进离场程序的所有航班进行聚类,使聚类后的数据便于完成航迹预测任务。

DBSCAN 为在含有噪声的空间样本中,根据样本分布密度进行分类的一种聚类算法。DBSCAN 规定了两个参数,半径  $\epsilon$  和最小点数 MinPts,以此来确定聚类时密度的阈值。根据这两个参数,可以将整个含有噪声的空间样本点分为三类,分别是核心点、边界点和噪声点。核心点满足以该点为圆心,以  $\epsilon$  为半径画圆,在该圆范围内的样本点总数大于最小点数 MinPts。边界点则是包含在某一核心点半径范围内,但是自身并不能满足核心点要求的点。既不是核心点又不是边界点的样本点被定义为噪声点。DBSCAN 的算法流程如图 1 所示。

DBSCAN 算法步骤如下:

① 先从样本集  $X = \{x_1, x_2, \dots, x_n\}$  中取出一点  $x_i$ ,判断  $x_i$  是否能成为核心点,若  $x_i$  不满足核心点定义,则继续从  $X$  中取点。若  $x_i$  满足核心点定义,则可以以  $x_i$  为核心点形成新的簇  $C_m$ 。

② 从  $x_i$  的邻域内选取其他点  $x_j$ ,判断  $x_j$  是否能作为该簇新的核心点继续扩张该簇。若  $x_j$  不满足核

心点定义,则  $x_j$  为边界点,不再进一步处理。若  $x_j$  满足核心点定义,则将  $x_j$  的邻域点加入  $x_i$  所形成的簇  $C_m$ ,同时再从  $C_m$  继续选点进行核心点判断和簇的扩张,直至簇  $C_m$  中的核心点已经全部遍历完成,无法继续扩张为止。

③ 将簇  $C_m$  中的所有点从样本集  $X$  去除,此时完成簇  $C_m$  的聚类。然后继续从样本集  $X$  中取出新点进行新簇的聚类,直至样本集  $X$  中所有点全部经过判断与聚类,剩下的点既不满足核心点条件,也未被某一簇包含,定义其为噪声点。

### 2.2 评价指标

目前聚类效果优劣判定主要采用的标准为轮廓系数,通过样本的簇内距离和簇间距离的比值确定聚类效果,样本  $x_i$  的轮廓系数公式为:

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (1)$$

其中,  $a(i)$  为样本  $x_i$  与其所属簇内其他样本的平均距离,  $b(i)$  为样本  $x_i$  与其他非同一簇的样本的距离的最小值。  $S$  为轮廓系数均值,即将所有样本点的轮廓系数取均值,定义为该聚类结果的轮廓系数。  $S$  的值域为  $(-1, 1)$ ,当  $S$  越趋近于 1,当前簇划分情况越优。

由于 DBSCAN 算法是针对含有噪声的空间样本数据,当 DBSCAN 算法调整参数时,会导致轮廓系数变化的同时,被判定为噪声的样本数量发生变化,故单纯的轮廓系数并不能完全代表分类结果的

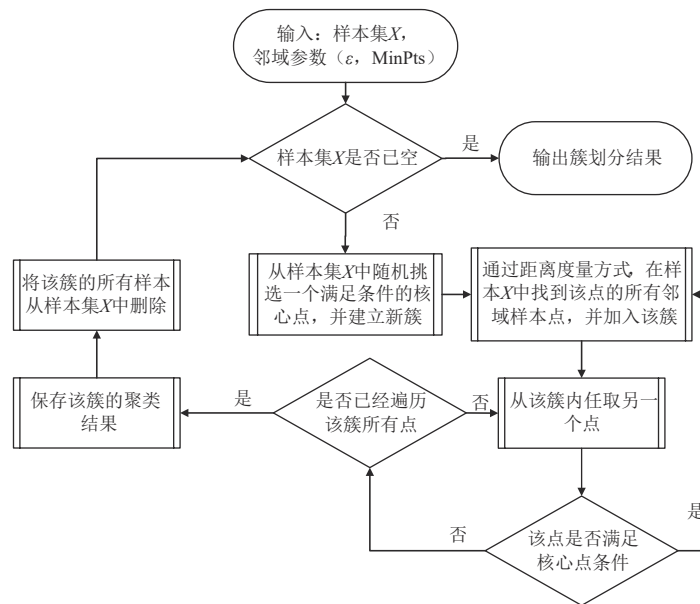


图 1 DBSCAN 算法流程

Fig. 1 DBSCAN algorithm flow

优劣。针对本文所研究的终端区航迹聚类问题,引入非噪声所占比重,即非噪声数据量与总数据量的比值,作为另一个评判DBSCAN聚类情况优劣的参数。考虑实际应用情况,所建立的航迹预测模型需适用于尽可能多的航班,故优先考虑非噪声所占比重,即应当在非噪声所占比重尽可能接近于1的前提下,轮廓系数越大,当前簇划分情况越优。

### 3 GRU神经网络

由于GRU可以很好地解决RNN中长期记忆和反向传播中的梯度消失和梯度爆炸问题。同时相较于LSTM的三个输入及三个输出,GRU只有两个输入和两个输出,这使得在能达到与LSTM同等性能的情况下,GRU比LSTM的参数量减少30%,较大的提高训练的效率。

GRU的隐藏单元结构如图2所示。其中, $\odot$ 表示哈达玛积(Hadamard product),即将矩阵中对应位置的元素进行乘积; $+$ 表示矩阵加法;“1-”表示用1减去当前时刻更新门控状态 $z_t$ ,根据更新门控状态对上一时刻的隐藏状态进行选择记忆与遗忘; $\sigma$ 表示sigmoid函数,通过这个函数可以将数据映射到 $[0,1]$ 范围内,以作为门控信号; $U$ 、 $W$ 均为权重矩阵; $\tanh$ 表示激活函数,以此增加神经网络模型的非线性,使得神经网络可以在一定范围内拟合任意函数。

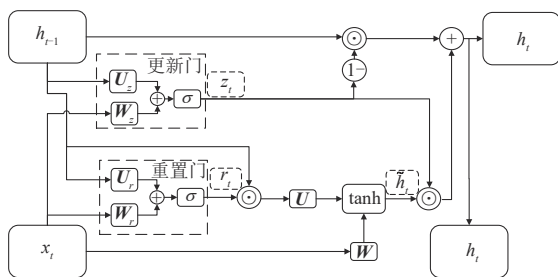


图2 GRU的隐藏单元结构

Fig. 2 Hidden unit structure of GRU

以 $t$ 时刻隐藏单元为例,首先,GRU利用神经元传输过来的上一时间步的隐藏状态 $h_{t-1}$ 和当前时刻的输入 $x_t$ 获得更新门的门控状态 $z_t$ 和重置门的门控状态 $r_t$ 。然后通过重置门的状态 $r_t$ 确定接收到的上一时间步的隐藏状态被当前步所遗弃的量,同时通过更新门的状态 $z_t$ 确定将上一时间步的隐藏状态传输到下一时间步的量,最终得到当前时刻隐藏单元的输出和传递给下一时刻的隐藏状态 $h_t$ 。式(2)

至(5)为 $t$ 时刻隐藏神经元状态更新的计算公式。

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (2)$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (3)$$

$$\tilde{h}_t = \tanh(W x_t + U(r_t \odot h_{t-1})) \quad (4)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (5)$$

### 4 终端区航迹预测方案

由于机场附近区域航班密度大、机型复杂,故设立终端管制区以更好地管理机场附近空域。终端管制区属于C、D类空域,根据《民用航空空中交通管理规则》定义为在一个或几个主要机场附近的空中交通服务航路汇合处的管制区,其垂直范围通常在6000 m(含)以下最低高度层以上;水平范围通常为半径50 km或走廊进出口以内的空间<sup>[20]</sup>。

由于航班在进入终端区时需确定进场程序,实际按照进场程序进行飞行时存在一定飞行范围,只要航班沿规定进场程序飞行即可,故本文所设计的终端区航迹预测方案如图3所示,主要分为三个阶段:

聚类阶段,先针对航迹数据集按照航班号进行划分,得到终端区航迹数据集,然后利用DBSCAN聚类算法分出不同航班簇。

训练阶段,将每一簇的航迹数据分别输入不同的GRU模型中进行模型训练,得到该簇的航迹预测模型。

预测阶段,当有需要预测的航班进入终端区时,采集其进入终端区的航迹点,进行判断其所属航班簇,根据判断结果,使用对应的航迹预测模型进行航迹预测。

#### 4.1 航迹聚类

训练航迹预测模型之前,首先需要确定模型的训练集,故先进行航迹聚类。此时将终端区内航班的数据定义为:

$$M = [m^1, m^2, \dots, m^e] \quad (6)$$

其中 $m^k = [f_m^1, f_m^2, \dots, f_m^o]$ 为终端区内的某一航班的所有航迹点,同时定义该航班的第 $i$ 个航迹点特征 $f_m^i$ 为:

$$f_m^i = [\text{FlagCode}, \text{lng}_i, \text{lat}_i, \text{height}_i, t_i] \quad (7)$$

其中FlagCode为航班号,lng<sub>*i*</sub>,lat<sub>*i*</sub>,height<sub>*i*</sub>,*t<sub>i</sub>*分别表示该航迹点的经度、纬度、高度和时间。然后利用半正矢公式计算航迹点与机场的距离,将不属于终端区范围内的航迹点删除,同时以航班为单位,将航迹点缺失严重的航班数据删除,半正矢公式为:

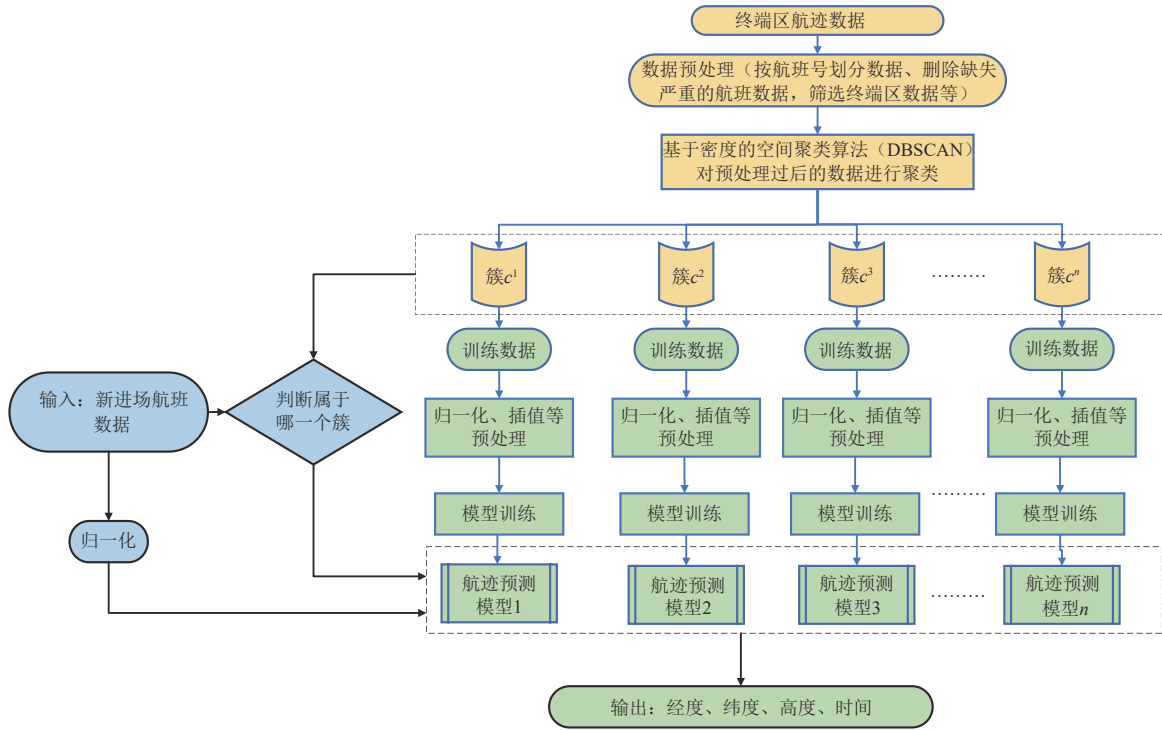


图 3 终端区航迹预测流程图

Fig. 3 Flow chart of trajectory prediction in terminal area

$$d = 2R \arcsin \left( \sqrt{\sin^2 \left( \frac{\text{lat}_2 - \text{lat}_1}{2} \right) + \cos(\text{lat}_2) \cos(\text{lat}_1) \sin^2 \left( \frac{\text{lng}_2 - \text{lng}_1}{2} \right)} \right) \quad (8)$$

其中,  $R$  为地球半径,  $\text{lng}_1, \text{lng}_2, \text{lat}_1, \text{lat}_2$  分别为两个点的经度和纬度。

进行 DBSCAN 聚类时, 要计算不同航班之间的空间距离, 此时考虑按照进入终端区范围或起飞时的时间顺序, 将航迹点坐标转化为 WGS-84 空间坐标系, 计算航班航迹点间空间距离的平均值。

然后以轮廓系数和非异常航迹所占比重为评价指标, 通过调整 DBSCAN 聚类算法的两个参数半径  $\epsilon$  和最小点数  $\text{MinPts}$  将终端区内的航班数据集  $M$  分成合适的簇  $C = [c^1, c^2, \dots, c^n]$ , 其中  $c^i = [f_i^1, f_i^2, \dots, f_i^{c_i}]$ , 由于每簇航班的航迹点数不同, 故用上角标  $c_i$  代表该簇航迹点的数量。根据聚类结果建立对应的航迹预测模型  $G = [g_1, g_2, \dots, g_n]$ 。

#### 4.2 航迹预测模型的训练及预测

完成航迹聚类后, 开始进行航迹预测模型训练及航迹预测。针对第  $i$  簇数据  $c^i$  建立航迹预测模型  $g_i$ , 首先将航迹数据  $c^i = [f_i^1, f_i^2, \dots, f_i^{c_i}]$  中的航班号删除, 避免进行训练时将航班号作为特征, 影响对经度、纬度、高度及时间四个特征的数据分析及预

测。同时定义航班进入终端区范围内的时刻为 0, 将每个航班的所有航迹点的时间特征值减去该航班的第一个航迹点的时间特征值, 然后将第一个航迹点的时间特征值置零, 以此消除不同日期、不同时间的航迹间的时间差异。

然后对航迹数据中的经度、纬度和高度进行归一化处理, 以消除不同特征间的因量纲和类别不同而产生对数据分析结果的影响。归一化方法采用线性函数归一化, 该归一化的优点为比较适用于数值集中的数据, 转化公式为:

$$\tilde{x}_i = \frac{x_i - \min(X)}{\max(X) - \min(X)} \quad (9)$$

其中  $\max(X)$  为样本数据集中的最大值,  $\min(X)$  为样本数据集中的最小值,  $x_i$  为原始样本数据集  $X$  中的某一样本数据,  $\tilde{x}_i$  为归一化后的数据。

下一步是对该簇航迹数据集进行数据分割, 即利用前  $L$  个航迹点预测后  $L'$  个航迹点, 设置一个长度为  $L$  的窗, 利用滑动窗对该簇航迹数据进行截取, 得到输入航迹预测模型的数据形式为:

$$F_i = [F_i^1, F_i^2, \dots, F_i^k] \quad (10)$$

其中  $F_i^a = [f_i^a, f_i^{a+1}, \dots, f_i^{a+L-1}]$ ,  $a \in [1, k]$ ,  $k = c^i - L + 1$ ;  $a, k, L \in \mathbf{N}$ 。

该输入数据所对应的理论输出应为:

$$F'_i = [F_i'^1, F_i'^2, \dots, F_i'^k] \quad (11)$$

其中  $F_i'^a = [f_i'^a, f_i'^{a+1}, \dots, f_i'^{a+L'-1}]$ ,  $a \in [1, k]$ ,  $k = c_i - L - L' + 1$ ;  $a, k, L' \in \mathbf{N}$ 。

训练神经网络模型时,将经过归一化及分割处理后的航迹数据输入到神经网络中,隐藏层为  $b$  个相同结构的 GRU 细胞,同时利用 dropout,在网络训练时随机地将部分隐藏单元暂时丢弃,从而缓解 GRU 模型的过拟合问题。模型根据输入  $F_i = [F_i^1, F_i^2, \dots, F_i^k]$  得到实际输出  $D_i = [D_i^1, D_i^2, \dots, D_i^k]$ ,然后根据理论输出  $F'_i$  与实际输出  $D_i$  计算损失值,利用 Adam 优化器优化模型,最终得到该簇的航迹预测模型。

进行航迹预测时,先将航迹点数据转化到 WGS-84 坐标系,然后计算目标航班前  $L$  个点与簇  $C = [c^1, c^2, \dots, c^n]$  中每一簇航班的平均距离并判断其所属航迹簇,然后利用对应的航迹预测模型进行

4D航迹预测,其训练及预测过程如图4所示。

## 5 实验结果及分析

### 5.1 实验数据与参数

实验设备采用实验室台式机,处理器为 Intel(R) Core(TM) i7-10700 CPU @ 2.90GHz, 2.90 GHz, 机带 RAM 为 16 GB。由于上海浦东国际机场位置的特殊性,即上海浦东国际机场和上海虹桥国际机场距离较近,上海浦东国际机场终端区范围内存在“一区两场”现象,因此,本文所研究的终端区范围为:以上海浦东国际机场为中心,半径为 100 km 的空间。结合终端区设定和上海空域情况,在覆盖上海浦东国际机场的进离场程序的情况下,进行数据选择。本文所采用的实验数据为该终端区范围内,以上海浦东国际机场为起落机场的共计 453 条航班。

#### 5.1.1 DBSCAN 参数优选

首先进行 DBSCAN 聚类算法的最优参数选择对比实验,评判标准同时采用了轮廓系数和非异常航迹所占比重。如图 5 所示,该图横坐标为非异常航迹所占比重,纵坐标为轮廓系数,每条不同颜色

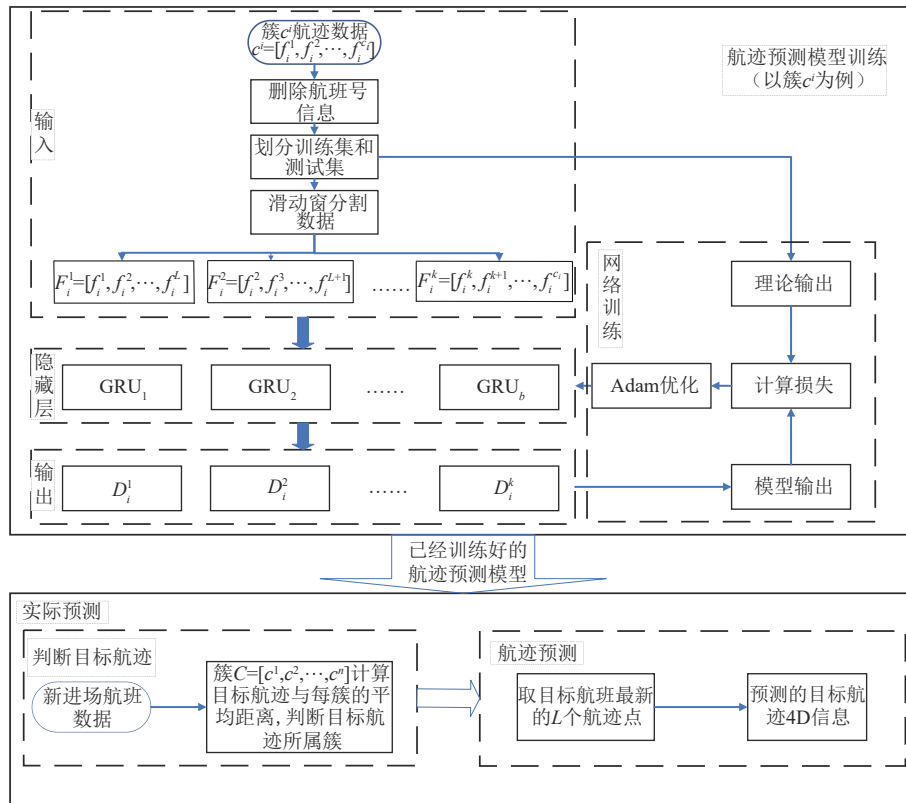


图4 航迹预测模型训练及航迹预测流程

Fig. 4 Trajectory prediction model training and trajectory prediction process

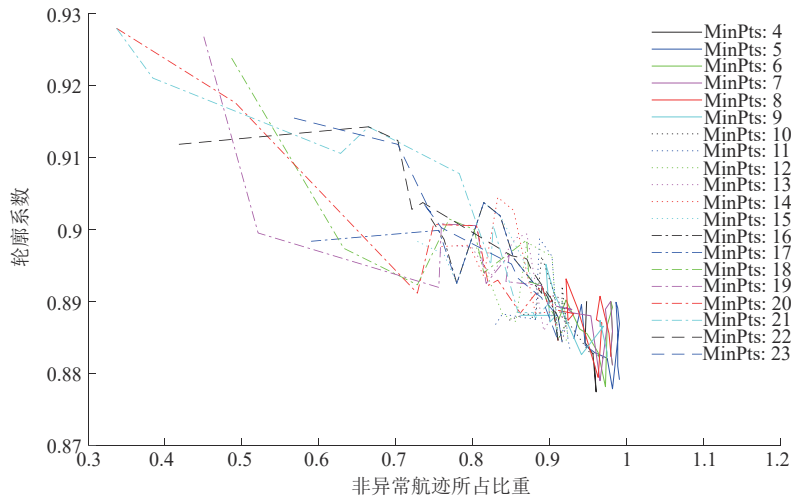


图 5 DBSCAN 不同参数下的非异常航迹所占比重与轮廓系数

Fig. 5 The proportion of non-anomalous trajectory and silhouette coefficient of DBSCAN under different parameters

的线代表了参数最小点数 MinPts 的不同取值,而每条线从左至右为半径  $\epsilon$  的不同取值,其范围为 0.13 至 0.24。由于在非异常航迹所占比重越接近于 1 的前提下,轮廓系数越接近于 1 时聚类效果最优,即图 5 中右上角的位置,结合表 1,在参数选择时应当采用 MinPts=5,  $\epsilon=0.21$ ,此时轮廓系数为 0.890,非异常航迹所占比重为 0.987。

表 1 DBSCAN 参数优选部分结果

Tab. 1 DBSCAN parameter selection results

MinPts	$\epsilon$	轮廓系数	非噪声所占比重
4	0.19	0.883	0.964
4	0.20	0.887	0.964
4	0.21	0.882	0.964
5	0.19	0.886	0.992
5	0.20	0.889	0.989
<b>5</b>	<b>0.21</b>	<b>0.890</b>	<b>0.987</b>
6	0.19	0.886	0.980
6	0.20	0.889	0.983
6	0.21	0.890	0.980

聚类效果如图 6 所示。由于终端区中航班既包括起飞航班,也包含降落航班,所以从图 6 中可以看出部分航迹即使重合,也未被归为一簇。

### 5.1.2 GRU 参数选取

采用 GRU 模型进行模型训练时,将实验数据按 8:2 的比例划分为训练集和测试集,根据航迹预测结果的均方根误差(Root Mean Square Error, RMSE)

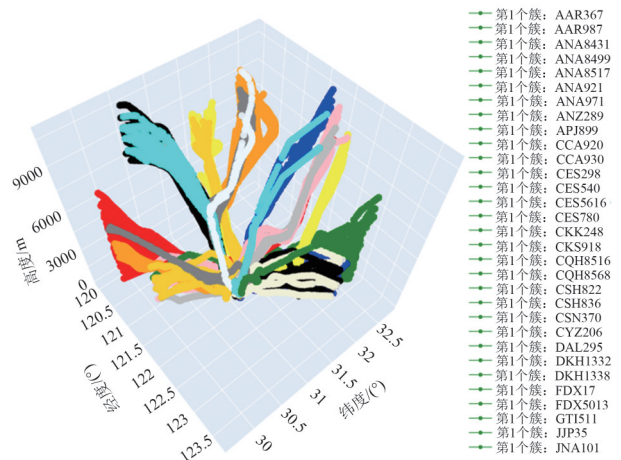


图 6 终端区航迹聚类结果

Fig. 6 Trajectory clustering result in terminal area

调整每层隐藏单元个数、训练窗长、层数等模型参数进行网格化参数优选实验,即在指定的参数空间内,依次调整参数进行训练,取训练结果最优时的参数的实验。RMSE 的计算公式为:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (a_i - b_i)^2}{k}} \quad (12)$$

其中  $a_i$  为每个点的预测值,  $b_i$  为每个点的真实值,  $k$  为测量次数。表 2 给出了部分网格化参数优选实验中参数采用不同值时的预测误差,根据预测误差结果选取最优参数。层数、每层隐藏单元个数等参数优选实验结果如表 3 所示。

表2 采用不同参数时的预测误差

Tab. 2 Prediction errors with different parameters

层数	窗长	隐藏单元个数	RMSE			
			经度/(°)	纬度/(°)	高度/m	时间/s
1	5	32	0.019	0.003	57.876	6.851
<b>1</b>	<b>5</b>	<b>64</b>	<b>0.014</b>	<b>0.003</b>	<b>56.249</b>	<b>2.813</b>
1	5	128	0.018	0.003	55.138	7.913
1	8	64	0.016	0.003	52.295	14.100
1	3	64	0.028	0.004	88.516	28.832
2	5	128	0.014	0.003	60.562	18.566
3	5	192	0.022	0.003	63.950	24.144
4	5	256	0.018	0.003	69.260	26.481

表3 部分参数优选结果

Tab. 3 Partial experimental results of parameter optimization

参数	参数值
滑动窗长	5
层数	1
每层隐藏单元个数	64
每层暂时丢弃数量比重(dropout)	0.02
学习率	0.001
一次训练所选取样本数	32

5.2 预测结果及性能分析

如图7、图8所示,其中图7为本文所提DBSCAN-GRU算法针对目标终端区某簇中两条不同航班航迹的预测值与真实值的比较结果。由图7可以看出目标航迹1与目标航迹2的预测值与真实值之间整

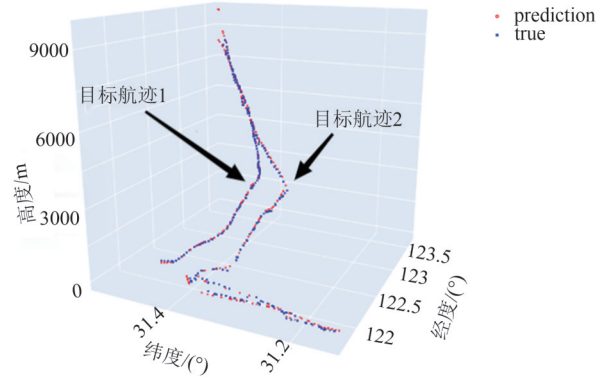


图7 基于DBSCAN-GRU算法的终端区4D航迹预测结果  
Fig. 7 4D trajectory prediction results in terminal area based on DBSCAN-GRU algorithm

体重合度均较高,同时,图7中的航班2与航班1属于同一簇,但航班号不同,证明本算法可以同时满足多航班进近的航迹预测任务,且每次单步预测耗时为0.035 s,远小于相邻航迹点间的时间间隔,满足航迹预测的时间要求。图8为目标航迹1中每个航迹点的经度、纬度、高度及时间的预测误差,而目标航迹1中存在部分缺失航迹点,图8中误差较大的航迹点与这部分缺失航迹点相对应。

为验证本算法的泛化性,故从终端区数据中挑选其他任一簇某航迹进行实验,结果如图9所示,其航迹整体重合度较高,表明该方法针对不同的终端区航迹簇均可较好地完成航迹预测任务,具有较

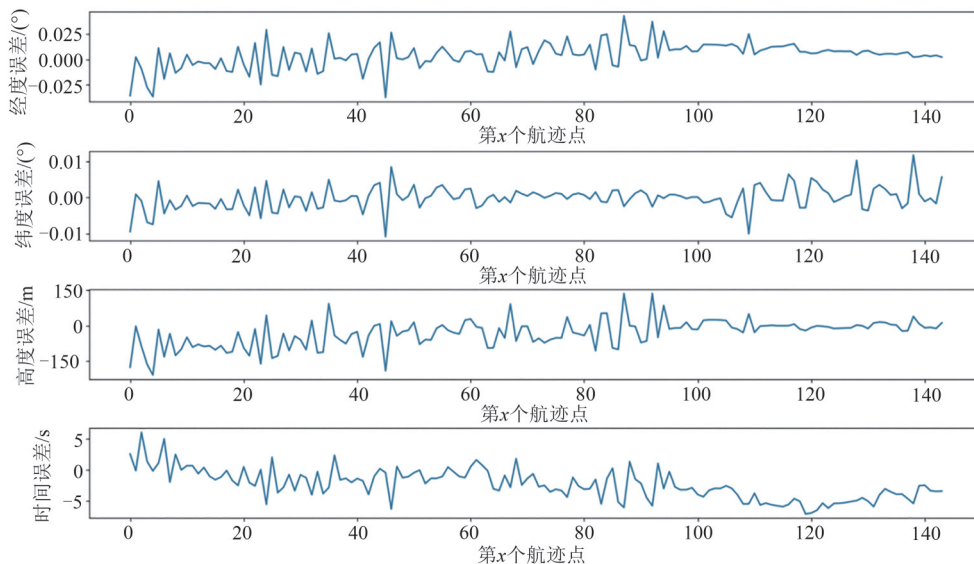


图8 每个航迹点的经度、纬度、高度及时间的预测误差

Fig. 8 Prediction errors of longitude, latitude, altitude and time for each trajectory point



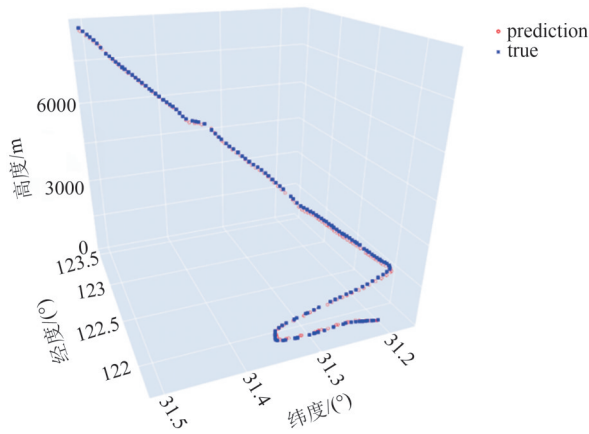


图 9 另一不同簇的航班采用本算法的预测结果  
Fig. 9 Prediction result of another different cluster of flights using this algorithm

好的泛化性。

为了进一步研究天气对终端区航迹预测性能的影响,故任取不同天气情况下的其他航班进行实验,其中某条航迹预测结果如图 10 所示,可以看出本算法在天气发生变化的情况下也有较好的预测效果。

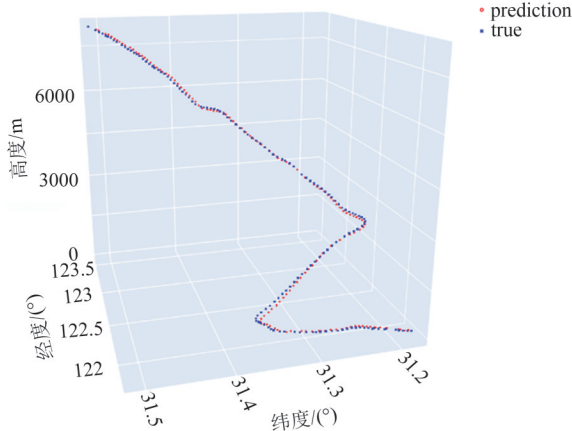


图 10 不同天气情况下另一航班的航迹预测结果  
Fig. 10 Trajectory prediction for another flight in different weather conditions

上述四条航班的航迹预测的 RMSE 如表 4 所示,可以看出,图 9、图 10 中航班的预测误差相对较小,其原因为图 9、图 10 中的航迹数据更为完整。

传统航迹预测进行模型训练时采用的是目标航班过去的航迹数据,而本算法采用的则是同一簇不同航班的历史航迹数据进行模型训练,故在实验时进行了同一航班两种模型的航迹预测结果对比。如表 5 所示,发现针对终端区内航迹,在相同训练数

表 4 不同训练数据的预测结果

Tab. 4 RMSE of prediction results of different training data

	RMSE			
	图 7 中所采 用的航班 1	图 7 中所采 用的航班 2	图 9 中所采 用的航班	图 10 中所采 用的航班
经度/(°)	0.014	0.011	0.007	0.005
纬度/(°)	0.003	0.005	0.002	0.004
高度/m	56.249	60.065	30.773	40.988
时间/s	2.813	2.836	4.237	4.256

表 5 不同训练数据的预测误差

Tab. 5 Prediction errors of different training data

	RMSE		
	同一航班 (x 条航迹)	同一簇航班 (所有航迹)	同一簇航班 (任取 x 条航迹)
经度/(°)	0.013	<b>0.014</b>	0.013
纬度/(°)	0.005	<b>0.003</b>	0.003
高度/m	58.324	<b>56.249</b>	57.385
时间/s	7.849	<b>2.813</b>	5.242

据量的情况下,本算法比传统航迹预测算法的精度高。由于本算法可以使得训练模型时所能利用数据量增加,使得预测精度进一步提高,同时训练所得的模型不只针对某一具体航班,可以对某一簇航班进行航迹预测,也体现出本文所提算法的优势。

本文所提出的基于 DBSCAN-GRU 的终端区航迹预测算法与目前已有的终端区航迹预测算法模型 GA-BP<sup>[18]</sup>、LSTM<sup>[13]</sup>进行对比实验,不同模型进行训练时的损失曲线图如图 11 所示,最终比较结果如表 6 所示。

由表 6 可以看出,本文所提出的算法相较于其

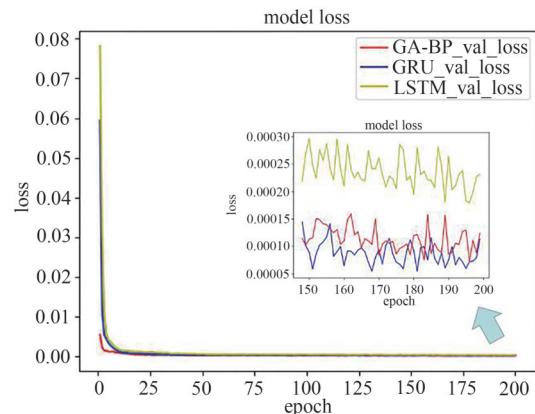


图 11 采用不同模型时的损失曲线图

Fig. 11 Loss curves when using different models

表6 采用不同模型时的预测误差

Tab. 6 Prediction errors when using different models

算法	RMSE			
	纬度/ (°)	经度/ (°)	高度/m	时间/s
<b>DBSCAN+GRU</b>	<b>0.014</b>	<b>0.003</b>	<b>56.249</b>	<b>2.813</b>
DBSCAN+LSTM	0.014	0.003	62.348	2.991
DBSCAN+GA-BP	0.031	0.014	69.242	8.953

他的终端区预测算法的RMSE较小,而且GRU算法模型较为简单,更适合构建较大的网络,处理大数据时效率更高,更适合完成终端区航迹预测任务。

## 6 结论

本文提出将DBSCAN算法与GRU算法相结合进行终端区航迹预测,使得一个航迹预测模型可以预测多架航班,避免每次所建模型仅针对单一航班,更加有效地挖掘终端区的空域数据的内在联系。算法针对复杂多变的终端区空域环境,利用DBSCAN聚类算法将终端区航迹进行分簇,然后针对每簇航迹进行建模,当某航班需要进行航迹预测时,将该航班与DBSCAN算法所分的簇进行匹配,利用同一簇的航迹预测模型去预测属于该簇的航班的航迹,解决了终端区复杂空域的航迹预测问题,为终端区的空域管理提供了一个较好的方案。通过仿真实验结果可以看出,本文所提方法具有较高的准确度和效率,可以较好的处理终端区航迹预测问题。下一步将从其他方面对终端区航迹预测模型进行深入研究,例如采用较新的神经网络模型、将进离场程序作为先验知识、预测时进行在线更新等,以进一步提高终端区范围内的航迹预测精度。

### 参考文献

- [1] 殷时军, 孟令航. 智慧民用航空运输系统思考与实践[J]. 南京航空航天大学学报(社会科学版), 2022, 24(4): 118-126.  
YIN Shijun, MENG Linghang. Thought and practice on the smart civil aviation transportation system of China[J]. Journal of Nanjing University of Aeronautics and Astronautics (Social Sciences), 2022, 24(4): 118-126. (in Chinese)
- [2] ZAMMIT B, ZAMMIT-MANGION D. Accuracy considerations of a simple aircraft trajectory prediction model for idle thrust descents[C]//2013 IEEE/AIAA 32nd Digital Avionics Systems Conference(DASC). East Syracuse, NY, USA. IEEE, 2013: 1E2-1-1E2-16.
- [3] 王超, 郭九霞, 沈志鹏. 基于基本飞行模型的4D航迹预测方法[J]. 西南交通大学学报, 2009, 44(2): 295-300.  
WANG Chao, GUO Jiuxia, SHEN Zhipeng. Prediction of 4D trajectory based on basic flight models[J]. Journal of Southwest Jiaotong University, 2009, 44(2): 295-300. (in Chinese)
- [4] 张萍萍, 孙永侃, 李雪飞. 视景仿真环境下基于卡尔曼滤波的运动目标航迹预测方法[J]. 兵工自动化, 2013, 32(5): 40-42, 45.  
ZHANG Pingping, SUN Yongkan, LI Xuefei. Method of trajectory prediction of moving target in scene simulation system based on Kalman filter[J]. Ordnance Industry Automation, 2013, 32(5): 40-42, 45. (in Chinese)
- [5] 王涛波, 黄宝军. 基于改进卡尔曼滤波的四维飞行航迹预测模型[J]. 计算机应用, 2014, 34(6): 1812-1815.  
WANG Taobo, HUANG Baojun. 4D flight trajectory prediction model based on improved Kalman filter[J]. Journal of Computer Applications, 2014, 34(6): 1812-1815. (in Chinese)
- [6] 杨霄鹏, 欧阳超, 杨朝阳, 等. 基于“当前”统计模型的模糊自适应航迹预测算法[J]. 空军工程大学学报(自然科学版), 2015, 16(2): 1-4.  
YANG Xiaopeng, OUYANG Chao, YANG Zhaoyang, et al. A fuzzy adaptive algorithm based on current statistical model for track prediction[J]. Journal of Air Force Engineering University (Natural Science Edition), 2015, 16(2): 1-4. (in Chinese)
- [7] MONDOLONI S, LIANG D. Improving trajectory forecasting through adaptive filtering techniques[C]//USA/Europe ATM Seminar. 2003.
- [8] LIN Xi, ZHANG Jun, ZHU Yanbo, et al. Simulation study of algorithms for aircraft trajectory prediction based on ADS-B technology[C]//2008 Asia Simulation Conference - 7th International Conference on System Simulation and Scientific Computing. Beijing. IEEE, 2008: 322-327.
- [9] HWANG I, HWANG J, TOMLIN C. Flight-mode-based aircraft conflict detection using a residual-mean interacting multiple model algorithm[C]//AIAA Guidance, Navigation, and Control Conference and Exhibit. Austin, Texas. Reston, Virginia: AIAA, 2003: 1-11.
- [10] WU Zhijun, TIAN Shan, MA Lan. A 4D trajectory prediction model based on the BP neural network[J]. Journal of Intelligent Systems, 2019, 29(1): 1545-1557.
- [11] 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.  
SHI Qingyan, WANG Wenqing, HAN Ping. Short-term 4D trajectory prediction algorithm based on online-updating LSTM network [J]. Journal of Signal Processing, 2021, 37(1): 66-74. (in Chinese)
- [12] 陈玉立, 佟强, 谌彤童, 等. 基于注意力机制和生成对抗网络的飞行器短期航迹预测模型[J]. 计算机应用, 2022, 42(10): 3292-3299.  
CHEN Yuli, TONG Qiang, CHEN Tongtong, et al. Short-term trajectory prediction model of aircraft based on attention mechanism and generative adversarial network [J]. Journal of Computer Applications, 2022, 42(10): 3292-3299. (in Chinese)
- [13] 翟俐民. 基于轨迹聚类 and LSTM 的航迹预测方法研究 [D]. 成都: 四川大学, 2021.  
ZHAI Limin. Research on track prediction method based on trajectory clustering and LSTM [D]. Chengdu: Sichuan University, 2021. (in Chinese)
- [14] 李海峰, 吴喜萍. 基于BCOP的终端区四维航迹预测研究[J]. 武汉理工大学学报(交通科学与工程版), 2012, 36(6): 1209-1211.  
LI Haifeng, WU Xiping. Terminal area 4D trajectory prediction model based on BCOP [J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2012, 36(6): 1209-1211. (in Chinese)
- [15] BARRATT S T, KOCHENDERFER M J, BOYD S P. Learning probabilistic trajectory models of aircraft in terminal airspace from position data [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(9): 3536-3545.
- [16] ROCHA MURÇA M C, DE OLIVEIRA M. A data-driven probabilistic trajectory model for predicting and simulating terminal airspace operations [C]//2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC). San Antonio, TX, USA. IEEE, 2020: 1-7.
- [17] CHOI H C, DENG Chuhao, HWANG I. Hybrid machine learning and estimation-based flight trajectory prediction in terminal airspace [J]. IEEE Access, 2021, 9: 151186-151197.
- [18] 李楠, 孙伯鑫, 焦庆宇, 等. 基于GA-BP神经网络的终端区航迹预测 [C]//世界交通运输工程技术论坛 (WTC2021)论文集(下), 2021: 396-405.  
LI Nan, SUN Boxin, JIAO Qingyu, et al. Trajectory prediction of terminal area based on GA-BP neural network [C]//Proceedings of World Transport Engineering Technology Forum (WTC2021) (Part 2), 2021: 396-405. (in Chinese)
- [19] 毕方明, 王为奎, 陈龙. 基于空间密度的群以噪声发现聚类算法研究[J]. 南京大学学报(自然科学版), 2012, 48(4): 491-498.  
BI Fangming, WANG Weikui, CHEN Long. DBSCAN: Density-based spatial clustering of applications with noise [J]. Journal of Nanjing University (Natural Sciences), 2012, 48(4): 491-498. (in Chinese)
- [20] 刘仁禹. 基于深度强化学习的终端区飞行冲突与调配 [D]. 成都: 四川大学, 2021.  
LIU Renyu. Flight conflict and deployment in terminal area based on deep reinforcement learning [D]. Chengdu: Sichuan University, 2021. (in Chinese)

#### 作者简介



韩萍女, 1966年生, 天津人。中国民航大学教授, 博士生导师, 研究方向为图像处理与模式识别、SAR目标检测与识别等。

E-mail: hanpingcauc@163.com



张启男, 1995年生, 河北邢台人。中国民航大学天津市智能信号与图像处理重点实验室硕士研究生, 主要研究方向为航迹预测。

E-mail: zq\_zhangqi\_1001@163.com



石庆研女, 1977年生, 山东济宁人。中国民航大学讲师, 研究方向为自适应信号处理、盲信号处理。

E-mail: qyshi@cauc.edu.cn



张泽中男, 1999年生, 河北承德人。中国民航大学天津市智能信号与图像处理重点实验室硕士研究生, 主要研究方向为航迹预测。

E-mail: zzz19931952527@163.com