

中文引用格式:刘宏,张鑫迪,卢飞,等. 基于改进 Transformer 模型的四维航迹预测[J]. 中国安全科学学报,2024, 34(12): 74-83.

英文引用格式:LIU Hong, ZHANG Xindi, LU Fei, et al. Research on 4D flight trajectory prediction based on improved Transformer model[J]. China Safety Science Journal, 2024, 34(12): 74-83.

基于改进 Transformer 模型的四维航迹预测*

刘宏¹副教授,张鑫迪¹,卢飞¹副教授,张成裕²

(1 中国民航大学 空中交通管理学院,天津 300300;2 北方自动控制技术研究所,山西 太原 030006)

中图分类号:X949

文献标志码:A

DOI: 10.16265/j.cnki.issn1003-3033.2024.12.0497

基金项目:国家自然科学基金资助(52272356);国家重点研发计划项目(2022YFB4300904);国家空管委项目(ZKG2023-03);中央高校基本业务费自然科学基金重点项目(3122022101)。

【摘要】 针对现有的四维航迹预测未充分考虑序列航迹数据之间存在关联关系等问题,改进 Transformer 模型架构,完善四维航迹预测的编码和解码操作。首先,改进卷积模块,利用普通卷积捕捉相邻时序点的关联关系,通过扩张卷积捕捉邻近时序序列点之间的隐式相关性,从而覆盖更大的序列范围;其次,采用多头自注意力对航迹的时空特征结合注意力分数的重要性进行调参计算,学习历史航迹数据的全局依赖关系;再次,通过引入概率稀疏方法,降低自注意力机制的计算复杂度,提高模型的计算效率;最后,搭建试验平台,预测对比航迹的经度、纬度和高度的时序特征。结果表明:改进 Transformer 模型与传统的 Transformer 模型等 4 种神经网络模型相比,时间性能提高 14.4%;采用均方根误差(RMSE)和平均绝对误差(MAE)作为评价指标,改进 Transformer 模型对经度、纬度和高度等航迹特征预测的偏差的平均值分别为 0.027 和 0.021;改进 Transformer 模型与传统 Transformer 模型的预测平均偏差 0.099 和 0.084 相比,分别减小 0.072 和 0.063。对预测序列长度的敏感性分析得到,改进 Transformer 模型与基准模型相比,预测的稳定性更高。

【关键词】 改进 Transformer 模型; 四维航迹; 航迹预测; 深度学习; 扩张卷积; 注意力机制

Research on 4D flight trajectory prediction based on improved Transformer model

LIU Hong¹, ZHANG Xindi¹, LU Fei¹, ZHANG Chengyu²

(1 School of Air Traffic Management, Civil Aviation University of China, Tianjin 300300, China;

2 North Automatic Control Technology Institute, Taiyuan Shanxi 030006, China)

Abstract: Flight trajectory prediction plays a crucial role in ensuring safe and efficient air traffic operation. In order to consider the implicit correlations between flight trajectory characteristics, the encoding and decoding operations of the prediction framework in the transformer model were enhanced. Firstly, the convolutional block was improved, and ordinary convolutions were applied to capture the correlations between neighboring time series trajectory characteristics, and dilated convolutions were added to capture correlations between non-neighboring time series trajectory characteristics. Secondly, multi-head self-attention was employed to perform calculation based on the spatiotemporal features of the flight trajectory combined with the importance of attention scores. Thirdly, probabilistic sparse method was

designed to reduce the computational complexity of the multi-head self-attention and improve the model's computational efficiency. Finally, an experimental platform was established to verify the flight trajectory prediction framework. The results show that compared to the traditional transformer model and the other three neural network models, the improved transformer model shows a 14.4% improvement in time performance. By using root mean square error (RMSE) and mean absolute error (MAE) as evaluation metrics, the average prediction deviations of the improved transformer model for trajectory features such as longitude, latitude, and altitude are 0.027 and 0.021, respectively. These deviations are reduced by 0.072 and 0.063 compared to the traditional transformer model's average prediction deviations of 0.099 and 0.084. Sensitivity analysis on the lengths of prediction sequences indicates that the improved transformer model is more stable than the baseline models.

Keywords: improved Transformer model; 4D flight trajectory; trajectory prediction; deep learning; dilated convolution; attention mechanism

0 引言

近年来,随着全球经济的发展,航空运输需求显著增加,这导致飞行交通量激增,空域复杂性增加。国际民航组织提出基于航迹运行 (Trajectory Based Operation, TBO) 的空中交通管理模式^[1]。TBO 运行模式以四维航迹为基础,在空管、机场和航空公司等相关方之间实时共享和动态维护航空器的四维航迹信息,进而实现多方协同决策,提高航班的运行效率^[2]。四维航迹预测通常是指利用航空器的当前和历史状态及其相关信息来估计未来的轨迹。提高航迹预测精度能够帮助空中交通管制员更安全、更高效地管理空域,如延误预测、冲突检测等,进而提升整体航空运输系统的安全性和效率。

航迹预测的主要方法有状态估计法、基于动力学建模的方法和基于深度学习的方法等。PREVOST^[3]、吕波^[4]等利用卡尔曼滤波方法等单模型估计法进行航迹预测,提高了航迹预测的稳定性与实时性;ZHANG Junfeng^[5]、FAIRLEY^[6]等考虑航空器在不同飞行阶段的恒速、协调转弯、恒定下降等飞行模式对应的不同状态转移,采用多模型估计法进行航迹预测。然而,航空器飞行运动过程还受到复杂环境、人机交互等不确定因素的影响。FUKUDA^[7]建立了航空器在各个飞行阶段的动力学方程并进行航迹的预测;SCHUSTER^[8]将传统的点质量模型进行扩展,通过动态调整质量大小,减小了预测的误差。张军峰等^[9]考虑了航空器在终端区进近与着陆阶段,不同飞行模式下的受力与约束,分别建立对应的动力学方程,减小了预测的位置误差和时间误差。在实际运行中,受到气象、通信、导航和监视设备等条件的限制,影响航空器位置的因素

及实时气象数据并不能及时、完整地地被地面管制系统获取。基于深度学习的方法在航迹预测中逐渐受到关注,王兴隆^[10]、ZENG Weili^[11]等利用深度学习中的长短时记忆网络 (Long Short-Term Memory, LSTM) 模型进行航迹预测并取得了较好的预测效果;2017年,ASHISH等^[12]将注意力机制 (Attention) 应用到神经网络中,提出一种名叫 Transformer 的模型,该模型既考虑数据的时序关系,又利用 Attention 来捕获数据各个部分的依赖关系。GUO Dongyue等^[13]利用 Transformer 模型预测航迹,有效提高了预测精度。冯霞等^[14]应用 Informer 模型作了关于长时四维航迹预测的研究。

针对现有的四维航迹预测未充分考虑序列航迹数据之间存在关联关系等问题,因此,笔者拟提出改进 Transformer 模型,在解码器和编码器输入的卷积部分加入扩张卷积,扩大航迹信息维度并增加邻近航迹点的相关性;利用多头稀疏自注意力机制减小 Transformer 模型计算复杂度,以期提高经度、维度和高度等航迹特征的预测准确性和计算效率。

1 航迹预测描述

航迹预测本质是多变量时间序列的预测,即利用给定的历史航迹信息预测未来一段时间的航迹点特征。航迹预测本质是预测多变量时间序列,即利用给定的历史航迹信息预测未来一段时间的航迹点特征。历史航迹表示如下式:

$$\mathbf{O}_{m,q}(t_0) = \begin{pmatrix} o_{t_{-m+1},1} & o_{t_{-m+1},2} & \cdots & o_{t_{-m+1},q} \\ o_{t_{-m+2},1} & o_{t_{-m+2},2} & \cdots & o_{t_{-m+2},q} \\ \vdots & \vdots & \ddots & \vdots \\ o_{t_0,1} & o_{t_0,2} & \cdots & o_{t_0,q} \end{pmatrix} \quad (1)$$

式中: t_0 为当前时刻; m 为历史航迹序列长度; q 为历史航迹信息的维度; $o_{i,j} \in \mathbf{O}_{m,q}(t_0)$ ($-m+1 \leq i \leq 0$, $1 \leq j \leq q$) 表示时刻 t_i 的第 j 维历史航迹信息点; 预测航迹表示如下式:

$$\mathbf{P}_{n,r}(t_0) = \begin{pmatrix} p_{t_1,1} & p_{t_1,2} & \cdots & p_{t_1,r} \\ p_{t_2,1} & p_{t_2,2} & \cdots & p_{t_2,r} \\ \vdots & \vdots & \cdots & \vdots \\ p_{t_n,1} & p_{t_n,2} & \cdots & p_{t_n,r} \end{pmatrix} \quad (2)$$

式中: n 为预测航迹序列的长度; r 为预测航迹点特征的维度; $p_{i,j} \in \mathbf{P}_{n,r}(t_0)$ ($1 \leq i \leq n$, $1 \leq j \leq r$) 为时刻 t_i 的第 j 维预测航迹点。特别地, 当不需要强调特征维度时, 将 $\mathbf{O}_{m,q}(t_0)$ 和 $\mathbf{P}_{n,r}(t_0)$ 分别记为 $\mathbf{O}_m(t_0)$ 和 $\mathbf{P}_n(t_0)$ 。航迹预测过程表示如下式:

$$\mathbf{P}_n(t_0) = A(\mathbf{O}_m(t_0)) \quad (3)$$

式中 $A(\cdot)$ 为预测模型。

2 广播式自动相关监视数据预处理

广播式自动相关监视 (Automatic Dependent Surveillance-Broadcast, ADS-B) 格式化数据包包含航空器坐标系位置、目标地址、几何高度、品质因数、侧滚角、气压高度、空速、真实空速、几何垂直速率及磁航向等信息。选取经度、纬度、高度、航向、空速以及航班编号等 5 个维度作为历史航迹信息, 选取经度、纬度和高度作为预测航迹点的特征。

ADS-B 格式化数据是根据欧盟航空交通管制系统监控数据交换标准, 由航空器机载端按照报文协议自动向周围广播的^[15]。地面接收站将处理收到的 ADS-B 信息, 经传输网络送往空管监视应用系统, 生成综合航迹, 并显示在终端上。由于设备、环境等因素的干扰, ADS-B 航迹数据往往不完整, 存在缺失、相邻航迹点之间的采样时间间隔不相同等问题。因此, 需要对采集的 ADS-B 航迹数据进行异常值剔除、插值以及归一化处理。

2.1 异常数据的处理

采用轨迹平滑方法航迹异常值剔除。给定滑动窗口大小 z , 偏差阈值 α , 历史航迹点个数 M , 计算时刻 t_i 的第 j ($1 \leq j \leq q$) 维历史航迹信息 $o_{i,j}$ 的平滑值 $f(o_{i,j})$:

$$f(o_{i,j}) = \frac{\sum_{i=a}^b o_{i,j}}{b-a} \quad (4)$$

式中: a 为 $0 \sim i - \left\lfloor \frac{z}{2} \right\rfloor$ 的最大值, 即 $a = \max$

$\left\{0, i - \left\lfloor \frac{z}{2} \right\rfloor\right\}$; b 为 $M - i + \left\lfloor \frac{z}{2} \right\rfloor + 1$ 的最小值, 即 $b = \min \left\{M, i + \left\lfloor \frac{z}{2} \right\rfloor + 1\right\}$; $\lfloor \cdot \rfloor$ 为向下取整符号; q 为输入序列的长度; 计算平滑后的航迹与原始航迹的绝对值偏差 $\varepsilon_{i,j} = |o_{i,j} - f(o_{i,j})|$, 并与阈值 α 比较, 当 $\varepsilon_{i,j} > \alpha$ 时, 认定 $o_{i,j}$ 为异常值, 删除该值。

2.2 样条插值

针对航迹数据不规整, 时间间隔不相同的问题, 对经度、纬度、高度、方向、速度等 5 维信息特征分别采用 3 次样条插值法^[16]处理。给定时间间隔 Δt , 构建 3 次样条函数:

$$g(x) = ax^3 + bx^2 + cx + d \quad (5)$$

以时刻 t_i 的第 j ($1 \leq j \leq q$) 维历史航迹信息点 $o_{i,j}$ 与相邻点 $o_{i+1,j}$ 作为边界节点带入 $g(x)$ 中, 根据连续性条件, 计算出对应区间的系数 a, b, c 和 d , 得到对应于区间 $[t_i, t_{i+1}]$ 的插值函数并计算插值点数; 按照给定的时间间隔, 计算出对应时间点的数值。更新区间边界为 $[t_{i+k} * \Delta t, t_{i+2}]$, 重复以上步骤, 直到遍历所有历史航迹点。

2.3 数据归一化

为消除量纲差异、避免权重不平衡问题, 对初始数据进行线性变换, 使结果映射到 $[0, 1]$ 的范围。给定 M 个历史航迹点和特征维度 j ($1 \leq j \leq q$), 分别计算时刻 t_i ($1 \leq i \leq M$) 的航迹特征 $o_{i,j}$:

$$\text{norm}(o_{i,j}) = \frac{o_{i,j} - o_{\min}}{o_{\max} - o_{\min}} \quad (6)$$

式中 $o_{\min} = \min\{o_{i,j}\}_{i=1}^M$ 和 $o_{\max} = \max\{o_{i,j}\}_{i=1}^M$ 分别为第 j 维历史航迹信息点中的最大值和最小值。以经度数据为例, 缺失值 (Not Available, NA) 及处理后的信息见表 1。

表 1 经度预处理

Table 1 Preprocessing longitude

时间	原始数据	插值后	归一化后
2023-11-11 T7:51:08	-118.409 8	-118.409 8	0.005 203 874
2023-11-11 T7:51:13	NA	-118.399 1	0.005 440 012
2023-11-11 T7:51:18	NA	-118.407	0.005 265 667
2023-11-11 T7:51:23	NA	-118.421 4	0.004 947 873
2023-11-11 T7:51:24	-118.424	NA	NA
2023-11-11 T7:51:28	NA	-118.431 3	0.004 729 39
2023-11-11 T7:51:33	NA	-118.435 6	0.004 634 493
2023-11-11 T7:51:38	NA	-118.437 5	0.004 592 562
2023-11-11 T7:51:40	-118.4384	NA	NA
2023-11-11 T7:51:43	NA	-118.440 3	0.004 530 769

续表 1

时间	原始数据	插值后	归一化后
2023-11-11 T7:51:48	NA	-118.444 9	0.004 429 251
2023-11-11 T7:51:53	NA	-118.450 6	0.004 303 458
2023-11-11 T7:51:56	-118.454 1	NA	NA
2023-11-11 T7:51:58	NA	-118.456 4	0.004 175 457
2023-11-11 T7:52:03	NA	-118.462 1	0.004 049 664
2023-11-11 T7:52:08	NA	-118.467 5	0.003 930 491

3 改进的 Transformer 模型构建

3.1 系统框架构建

Transformer 模型主要由编码器和解码器 2 部分组成。编码器中的嵌入层、注意力机制层与前馈层将标准化历史航迹数据抽象成高维数据并进行调参计算;再利用解码器输出标准化的预测数据;通过数据反归一化处理输出预测航迹。

改进 Transformer 模型架构中编码操作和解码操作,构建四维航迹预测的系统框架如图 1 所示,其中,黑色为卷积块的改进用于扩大模型对历史航迹点的感知范围;灰色为注意力机制层引入的多头概率稀疏方法。

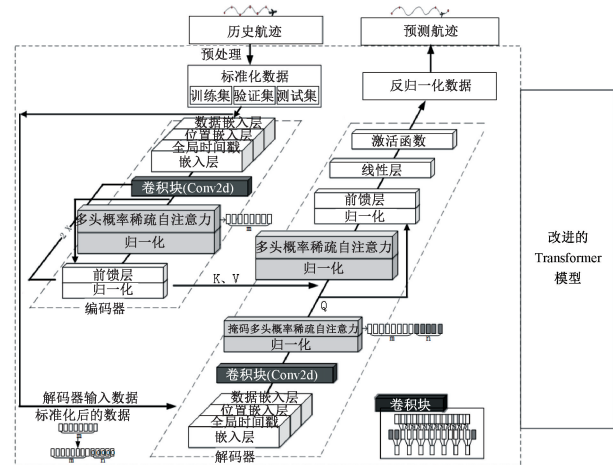


图 1 四维航迹预测系统框架

Fig. 1 Framework of four-dimensional trajectory prediction system

为统计航迹序列的关联关系,在传统 Transformer 模型中,编码器与解码器的嵌入层与注意力机制层之间增加扩张卷积层。记 $\mathbf{O}'_{m,d}(t_0)$ 表示当前时刻 t_0 长度为 m ,历史航迹 $\mathbf{O}_{m,q}(t_0)$ 经过嵌入层升高维度后得到的 d 维数据; $\mathbf{O}''_{m,d}(t_0)$ 表示 $\mathbf{O}'_{m,d}(t_0)$ 经过卷积层后的数据。卷积过程表示如下:

$$\mathbf{O}''_{m,d}(t_0) = \text{Conv2d}(\mathbf{O}'_{m,d}(t_0)) \quad (7)$$

此外,为提高计算效率,引入 Informer 模型中的多头概率稀疏自注意力机制计算序列特征数据的权重见下式:

$$\mathbf{B}''_{m,d}(t_0) = \text{SparAtten}(\mathbf{O}''_{m,d}(t_0)) \quad (8)$$

3.2 卷积块

卷积块包含普通卷积与扩张卷积 2 层,如图 2 所示。用 NormConv2d 表示普通卷积层,ExtenConv2d 表示扩张卷积层,则式(7)表示为:

$$\text{Conv2d} = \text{ExtenConv2d} \cdot \text{NormConv2d} \quad (9)$$

普通卷积层用于学习相邻航迹点的特性。

$\tilde{\mathbf{O}}'_{m,d}(t_0)$ 为 $\mathbf{O}'_{m,d}(t_0)$ 经过普通卷积作用后的数据,见下式:

$$\tilde{\mathbf{O}}'_{m,d}(t_0) = \text{NormConv2d}(\mathbf{O}'_{m,d}(t_0)) \quad (10)$$

给定 $o'_{t_i,j} \in \mathbf{O}'_{m,d}(t_0)$ 和权重矩阵 $\mathbf{W} \in \mathbb{R}^{\kappa \times d}$,首尾各加一行 0 向量:

$$\tilde{o}'_{t_i,j} = \sum_{s=1}^d \sum_{k=1}^{\kappa} o'_{t_{i-k+1},j^{s-1}} * w_{k,s} \quad (11)$$

式中: $\tilde{o}'_{t_i,j}$ 为 $\tilde{\mathbf{O}}'_{m,d}(t_0)$ 每个航迹点的数据; κ 为卷积核的大小, $i \in [-m+1, 0], j \in [1, d]$ 。

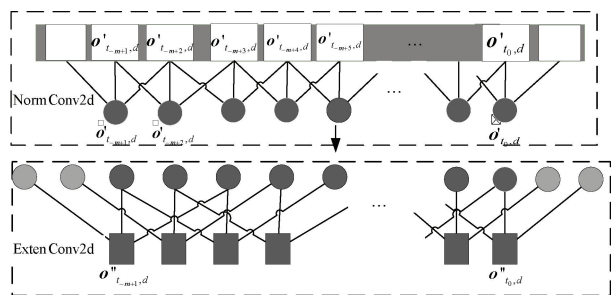


图 2 卷积层作用

Fig. 2 Schematic diagram of convolutional layer

扩张卷积层用于学习间隔一定距离航迹点的特性,见下式:

$$\mathbf{O}''_{m,d}(t_0) = \text{ExtenConv2d}(\tilde{\mathbf{O}}'_{m,d}(t_0), l) \quad (12)$$

具体过程见下式:给定 $\tilde{o}'_{t_i,j} \in \tilde{\mathbf{O}}'_{m,d}(t_0)$,权重矩阵 $\mathbf{W} \in \mathbb{R}^{\kappa \times d}$ 和扩张率 l ,得到:

$$o''_{t_i,j} = \sum_{d=1}^D \sum_{\kappa=1}^{\kappa} \tilde{o}'_{t_{i+\kappa(l-1)},j+d(l-1)} * w_{\kappa,d} \quad (13)$$

3.3 多头概率稀疏自注意力

3.3.1 多头自注意力

Transformer 模型的 Attention 模块由自注意力机制层、归一化层以及残差连接组成。给定 $1 * m$ 的行向量 $\mathbf{v} = (v_1, v_2, \dots, v_m)$, $\mathbf{A} = (a_1, a_2, \dots, a_m) =$

$\text{softmax}(\mathbf{v})$, 元素 $a_i = \frac{e^{v_i}}{\sum_{i=1}^m e^{v_i}}$ 。自注意力机制包

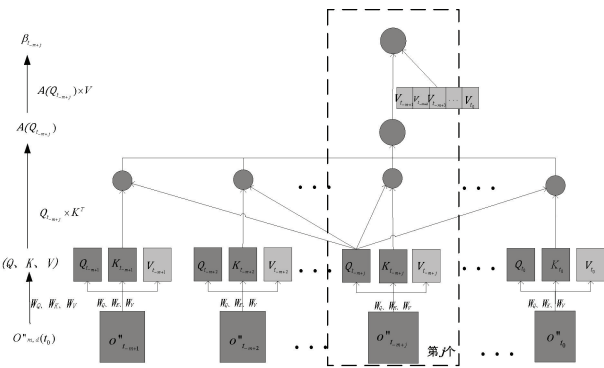


图3 标准自注意力机制

Fig. 3 Diagram of standard self-attention mechanism

含3步。第1步,在时刻 t_0 ,将输入的数据 $\mathbf{O}_{m,d}^n(t_0)$ 分别与实数空间 $\mathbb{R}^{d \times m}$ 中设定的网络参数 $\mathbf{W}_Q, \mathbf{W}_K$ 和 \mathbf{W}_V 加权运算,得到实数空间 $\mathbb{R}^{m \times m}$ 中的查询矩阵 \mathbf{Q} 、键矩阵 \mathbf{K} 、值矩阵 \mathbf{V} ,标准自注意力机制如图3所

示;第2步分别计算查询矩阵 \mathbf{Q} 中的第 i 个 ($i = 0, 1, \dots, -m + 1$) 行向量 \mathbf{Q}_{t_i} 。对应的向量见下式:

$$A(\mathbf{Q}_{t_i}, \mathbf{K}) = \text{softmax}\left(\frac{\mathbf{Q}_{t_i} \mathbf{K}^T}{\sqrt{d_K}}\right) \quad (14)$$

式中 d_K 为矩阵 \mathbf{K} 的维度。第3步,分别计算对应的标准注意力机制权重 \mathbf{B} 所对应的行向量 β_{t_i} ($i = 0, 1, \dots, -m + 1$),见下式:

$$\beta_{t_i} = \text{Atten}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = A(\mathbf{Q}_{t_i}) \mathbf{V} \quad (15)$$

记 h 为自注意力的头数,它是在自注意力机制的第1步中,将 $\mathbb{R}^{d \times m}$ 空间中的网络参数 $\mathbf{W}_{Q_i}, \mathbf{W}_{K_i}$ 和 \mathbf{W}_{V_i} 变换为 h 组 $\mathbf{W}_Q^1, \mathbf{W}_K^1, \mathbf{W}_V^1; \dots; \mathbf{W}_Q^h, \mathbf{W}_K^h, \mathbf{W}_V^h$, 分别与数据 $\mathbf{O}_{m,d}^n(t_0)$ 加权运算,得到 $\mathbb{R}^{m \times 1}$ 空间中的 h 组即:

$$\{\mathbf{Q}_{t_i}^1, \mathbf{K}_{t_i}^1, \mathbf{V}_{t_i}^1\}; \dots; \{\mathbf{Q}_{t_i}^h, \mathbf{K}_{t_i}^h, \mathbf{V}_{t_i}^h\}$$

为方便描述,将第 l ($1 \leq l \leq h$) 头对应的矩阵表示为 $\{\mathbf{Q}_{t_i}^l, \mathbf{K}_{t_i}^l, \mathbf{V}_{t_i}^l\}_{i=-m+1}^0$, 简记为 $\{\mathbf{Q}^l, \mathbf{K}^l, \mathbf{V}^l\}$ 。多头自注意力如图4下半部分所示。

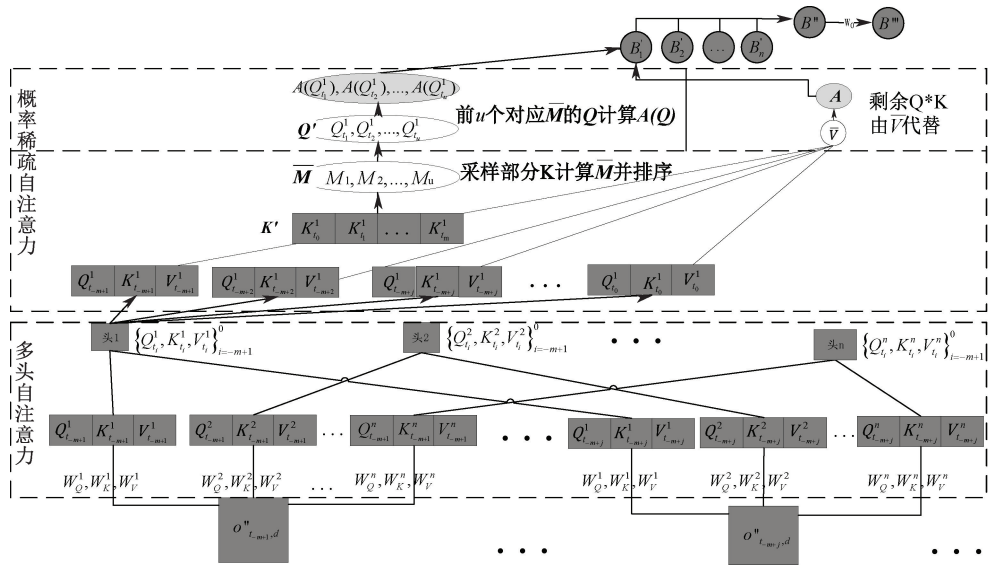


图4 多头概率稀疏自注意力

Fig. 4 Schematic diagram of multi-headed probability sparse self-attention

3.3.2 概率稀疏自注意力

记 $u = c * \ln(m)$, 其中, c 为恒定采样因子。概率稀疏自注意力是从 \mathbf{Q} 中筛选 u 个重要的 \mathbf{Q} , 计算对应的 $A(\mathbf{Q})$; 而将其余 $m-u$ 个 A 用 \mathbf{V} 中对应的 $m-u$ 个元素的平均值替代, 从而将自注意力的时间复杂度降至 $O(dml \ln m)$ 。这2组 $A(\mathbf{Q})$ 和 A 的计算方法如下:

$\{\mathbf{K}_{t_i}^l\}_{i=-m+1}^0$ 中, 随机采样 u 个元素组成采样集 \mathbf{K}' , 计算 $\bar{\mathbf{M}}(\mathbf{Q}_{t_i}^l, \mathbf{K})$, 见下式:

$$\bar{\mathbf{M}}(\mathbf{Q}_{t_i}^l, \mathbf{K}) = \max_{\mathbf{K} \in \mathbf{K}'} \left\{ \frac{\mathbf{Q}_{t_i}^l \mathbf{K}^T}{\sqrt{d_K}} \right\} - \frac{1}{m} \sum_{\mathbf{K} \in \mathbf{K}'} \frac{\mathbf{Q}_{t_i}^l \mathbf{K}^T}{\sqrt{d_K}} \quad (16)$$

从序列 $\{\bar{\mathbf{M}}(\mathbf{Q}_{t_i}^l, \mathbf{K})\}_{i=-m+1}^0$ 中选取前 u 个最大 $\bar{\mathbf{M}}(\mathbf{Q}_{t_i}^l, \mathbf{K})$ 值对应的元素组成 \mathbf{Q}' ; 对每个 $\mathbf{Q} \in \mathbf{Q}'$ 利用式(13)分别计算 $A(\mathbf{Q})$ 。

第1组, 对每个 $l \in [1, n]$, 从长度为 m 的序列

第 2 组,对剩余的 $m - u$ 个 A ,利用下式计算:

$$A = \text{softmax} \left(\frac{\bar{V}_{t_i}}{\sqrt{d_{K_i}}} \right) \quad (17)$$

式中 $\bar{V}_{t_i} = \frac{\sum_{i=-m+1}^0 V_{t_i}^l}{m}$ 为 V^l 的均值。

利用式(14)计算第 l 头 ($l \in [1, n]$) 的权重 B^l 对应的行向量 β^l , 再把 B^l 按列依次拼接得到 $m \times d$ 维矩阵 B'' , 与网络参数 $W_0 \in \mathbb{R}^{d \times d}$ 相乘得到 $m \times d$ 维矩阵 B''' 。

3.4 解码器

将当前时刻 t_0 历史航迹标准化后,扩充标准化后的数据,添加长度为 n 的 0 向量,得到解码器的输入数据,如图 5 所示。

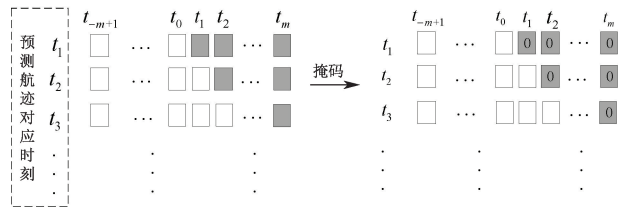


图 5 解码器中的数据掩码过程

Fig. 5 Masked data in decoder

掩码后的数据依次经过嵌入层、卷积块、掩码多头概率稀疏自注意力机制层得到矩阵 Q 。接着,将该矩阵与编码器传入的 K, V 组合,传入下一个多头概率稀疏自注意力层,继续计算注意力权重得到关系矩阵 B''' 。最后,通过前馈层进行训练得到最优模型参数,输入到线性层和激活函数,进而输出标准化的预测数据。

3.5 损失函数

在前馈层,将预测的航迹与真实值进行对比并计算损失函数,进行反向传播,调整优化训练集中各参数,得到最优模型。文中采用均方误差 (Mean Square Error, MSE) 作为损失函数,见下式:

$$\text{MSE} = \frac{1}{d} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^d (\bar{p}_{t_i,j} - \hat{p}_{t_i,j})^2 \quad (18)$$

式中: $\bar{p}_{t_i,j}$ 为航迹点 t_i 的特征 j 对应的归一化之前的预测值; $\hat{p}_{t_i,j}$ 为相应的反归一化之前的真实值。

4 预测模型参数与结果分析

4.1 数据集与预测模型参数

文中选取机型为空客 A321,时间段为 2023 年

11 月 11 日—2024 年 1 月 5 日,共 30 天的航班运行 ADS-B 数据作为航迹数据集,起飞机场和目的地机场分别为洛杉矶国际机场和肯尼迪国际机场。航迹数据包含航空器起飞爬升、巡航及进近各阶段共 135 458 个航迹点。

将这些航迹点分为训练集、验证集和测试集。输入数据包括纬度,经度,高度,速度,方向,时刻的六维特征。输出按时序预测的经度,纬度和高度四维特征值。编码器堆叠 2 层卷积块、多头概率稀疏自注意力机制与前馈层;解码器设置 1 层卷积块多头概率稀疏自注意力机与前馈层。其中, l, d, h 、训练周期、批量大小、学习率初值和激活函数等设置见表 2。

表 2 模型参数

Table 2 Model parameters

参数	设置
κ	3
l	2
d	512
多头注意力的头数	8
训练周期	10
批量大小	128
优化器	Adam
学习率初值	0.001
激活函数	GELU

预测模型运行的软硬件环境见表 3。

表 3 运行环境

Table 3 Operating environment

硬件环境	主频/GHz	2.10
	内存/G	120
	显存/GB	24
软件环境	编程语言	Python 3.8
	基础框架	PyTorch 2.0.0

4.2 基准模型

选择航迹预测常用的 Transformer 模型、Informer 模型、LSTM+卷积网络 (Convolution Neural Network, CNN) 模型以及 LSTM+CNN+ Attention 等 4 种神经网络模型作对比 F_1, F_2, F_3, F_4 。各个模型中,网络训练周期、批量大小、学习率初值和激活函数等参数设定同表 2 一致,其余参数设定如下:

1) 模型 F_1 : 编码器堆叠 2 层多头自注意力机与前馈层,解码器设置 1 层多头自注意力机与前馈层。

2) 模型 F_2 : 编码器堆叠 2 层多头概率稀疏自注意力机与前馈层,解码器设置 1 层多头概率稀疏

自注意力机与前馈层。

3) 模型 F_3 : 基于 LSTM 搭建的框架, 设定 32 个隐藏单元、2 层 LSTM。

4) 模型 F_4 : 在 LSTM 的框架中加入 CNN 和 Attention, 设定 256 个隐藏单元、2 层 LSTM 网络。

4.3 评价指标

采用均方根误差 (Root Mean Square Error, RMSE) 和平均绝对误差 (Mean Absolute Error, MAE) 作为误差评价指标。其中, RMSE 先对误差进行平方的累加后再开方, 放大了较大误差之间的差距, MAE 反映的是真实误差, 给予误差平等权重, 不会受到极端偏差的影响, 指标的计算见下式:

$$RMSE = \sqrt{\frac{1}{d} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^d (\bar{p}_{t_i,j} - \hat{p}_{t_i,j})^2} \quad (19)$$

$$MAE = \frac{1}{d} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^d |\bar{p}_{t_i,j} - \hat{p}_{t_i,j}| \quad (20)$$

式中: $\bar{p}_{t_i,j}$ 为航迹点 t_i , 特征 j 对应的归一化之前的预测值; $\hat{p}_{t_i,j}$ 为相应的归一化之前的真实值。

4.4 结果分析

1) 运行效率。首先, 对比模型 F_1 与改进 Transformer 模型 (简记为模型 F) 的运行效率差异。由于模型运行的速度随着数据集的增加而增大, 为统计不同规模数据集对应的运行时长, 将历史航迹分 2 组, 并设置 $m = 10$ 。第 1 组选取 15 天的历史航迹数据集共 64 391 个航迹点做测试, 其中, 前 13 天

数据为训练集, 剩余的第 14 天和第 15 天分别为验证集和测试集; 第 2 组选取 30 天的历史航迹数据集共 135 458 个航迹点做测试, 其中, 前 28 天数据为训练集, 剩余的第 29 和第 30 天分别为验证集和测试集。

经过多轮测试, 各个模型的运行时间统计见表 4。第 1 组中, 预测所用的整体时间不超过 160s, 每 epoch 平均时间不超过 16 s。其中模型 F_1 平均运行时间为 15.28 s, 模型 F 平均运行时间为 13.08 s。第 2 组中, 预测所用的整体时间不超过 360 s, 每 epoch 平均时间不超过 36 s, 其中, 模型 F_1 平均运行时间为 35.31 s, 模型 F 运行时间为 13.08 s。相比于模型 F_1 , 运行效率提升 14.4%。

表 4 运行速度对比

Table 4 Comparison of running speed

组别	模型	每 epoch (平均) 时间/s	整体时间/s
第 1 组	F_1	15.28	152.814
	F	13.08	130.886
第 2 组	F_1	35.73	357.36
	F	30.16	301.68

2) 预测偏差分析。选取 30 天的历史航迹作为数据集, 设置 n 为 1, 对比 5 个模型预测的纬度、经度和高度等 3 个特征对应的评价指标 MAE 和 RMSE, 见表 5。此外, “平均” 一列给出了各个模型对纬度、经度和高度等 3 个特征对应的评价指标 MAE 和 RMSE 偏差的平均值。

表 5 基准模型预测数据对比

Table 5 Prediction data comparison with benchmark models

模型	MAE				RMSE			
	纬度	经度	高度	平均	纬度	经度	高度	平均
F_1	0.044	0.092	0.116	0.084	0.05	0.113	0.136	0.099
F_2	0.028	0.032	0.047	0.035	0.035	0.041	0.055	0.043 3
F_3	0.034	0.058	0.062	0.047	0.042	0.075	0.077	0.0656
F_4	0.026	0.018	0.026	0.023	0.034 5	0.032	0.028	0.032
F	0.018	0.027	0.015	0.021	0.029	0.030	0.020	0.027

模型 F 预测的纬度和高度对应的 MAE 值分别为 0.018 和 0.015, 均小于其他基准模型的对应值。模型 F_4 预测的经度对应的 MAE 值 (为 0.018) 最小, 模型 F 的对应 MAE 值为 0.027, 与之相比高 0.011; 且预测的 MAE 的平均偏差为 0.021, 相比其他基准模型均减少, 与模型 F_1 的预测平均偏差 0.084 相比, 减小 0.063。

模型 F 与其他基准模型相比, 预测的经度、纬度和高度对应的 RMSE 值均减少; 且预测的 RMSE

的平均偏差为 0.027, 与模型 F_1 预测平均偏差 0.099 相比, 减小 0.072。

图 6 为航迹的三维可视化结果。其中, 图 6a 为历史航迹的三维可视化, 图 6b 为预测航迹高度和位置 (经度和纬度) 的三维可视化。模型 F_1 、模型 F_2 的预测曲线过于平滑, 位置偏差也较大; 而模型 F_3 的预测航迹与真实航迹位置偏差较大; 模型 F_4 的预测曲线在进近下降阶段位置偏差较大。模型 F 对拐点、曲率以及位置偏差方面均表现出更好的预测

性能和拟合效果。

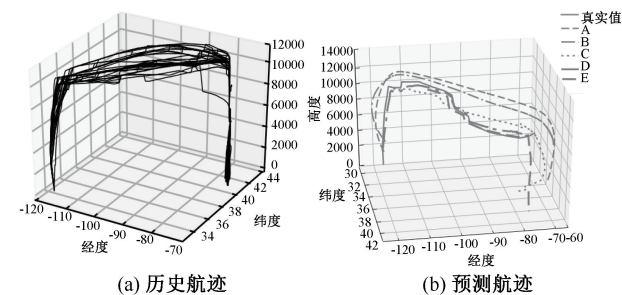


图 6 三维可视化

Fig. 6 3D-visualized latitude and longitude

图 7 为预测航迹高度的可视化结果。根据真实航迹中飞行器的高度变化,将预测航迹分为 3 个阶段:起始爬升阶段(预测航迹点 1 至预测航迹点 270)、巡航阶段(预测航迹点 271 至预测航迹点 3 450)及进近下降阶段(预测航迹点 3 450 至预测航迹点 3 750)。从图 6b 可以看出,起始爬升阶段,模型 F 与模型 F₂ 的预测航迹与真实航迹高度偏差较小,模型 F₁、模型 F₂ 和模型 F₃ 的预测航迹与真实航迹偏差较大。巡航阶段,模型 F 与模型 F₄ 灵敏地

捕捉到了飞行器平飞过程中高度的调整,模型 F₁、模型 F₂ 和模型 F₃ 的预测高度变化的灵敏度较为迟缓,且与真实航迹对应的高度偏差较大。进近下降阶段,模型 F 与模型 F₁、模型 F₃、模型 F₄ 的预测航迹与真实航迹高度偏差较小,模型 F₂ 的预测航迹与真实航迹偏差较大。对应于各个飞行阶段,预测模型的 MAE 和 RMSE 值参见表 6。

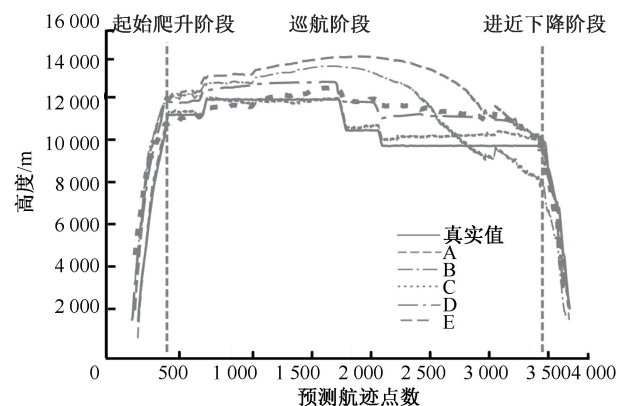


图 7 各飞行个阶段预测航迹可视化对比

Fig. 7 Visualization of predicted trajectories in three stages

表 6 不同阶段基准模型预测数据对比(预测序列长度为 1)

Table 6 Prediction data comparison with benchmark models in three stages (length of prediction series is 1)

阶段	起始爬升阶段				巡航阶段				进近下降阶段				
	模型	纬度	经度	高度	平均	纬度	经度	高度	平均	纬度	经度	高度	平均
F ₁	RMSE	0.132	0.015	0.091	0.079	0.051	0.108	0.145	0.101	0.027	0.120	0.030	0.059
	MAE	0.125	0.013	0.086	0.074	0.034	0.125	0.050	0.069	0.034	0.125	0.149	0.102
F ₂	RMSE	0.146	0.031	0.119	0.098	0.078	0.155	0.097	0.33	0.015	0.116	0.11	0.08
	MAE	0.128	0.026	0.114	0.089	0.155	0.096	0.038	0.096	0.013	0.113	0.108	0.078
F ₃	RMSE	0.014	0.045	0.118	0.059	0.066	0.072	0.052	0.063	0.040	0.048	0.045	0.044
	MAE	0.012	0.030	0.100	0.047	0.038	0.065	0.064	0.055	0.038	0.038	0.037	0.038
F ₄	RMSE	0.029	0.005	0.031	0.022	0.029	0.024	0.030	0.028	0.072	0.06	0.027	0.053
	MAE	0.027	0.010	0.029	0.022	0.022	0.015	0.030	0.027	0.071	0.059	0.026	0.052
F	RMSE	0.071	0.014	0.024	0.036	0.020	0.038	0.021	0.026	0.016	0.060	0.019	0.031
	MAE	0.055	0.01	0.019	0.028	0.015	0.029	0.018	0.020	0.018	0.048	0.010	0.025

模型 F₁ 预测航迹的特征偏差较大,后续主要考虑模型 F₂、模型 F₃ 和模型 F₄。

3) 预测序列长度的敏感性分析。为对比不同 n 对模型的影响,分别设置 $n=1, 16, 32, 48, 64, 80, 96$ 步,对应的时长 5 s、1 min 20 s、2 min 40 s、3 min 40 s、4 min、6 min 40 s、8 min。采用欧氏距离计算航迹特征的总体偏差 D 见下式:

$$D =$$

$$\sqrt{\frac{\sum_{i=1}^{|TE|} (\bar{p}_{i,1} - \hat{p}_{i,1})^2 + (\bar{p}_{i,2} - \hat{p}_{i,2})^2 + (\bar{p}_{i,3} - \hat{p}_{i,3})^2}{|TE|}}$$

(21)

式中: $\hat{p}_{i,1}, \hat{p}_{i,2}, \hat{p}_{i,3}$ 为归一化后第 t_i 个航迹点的纬度、经度和高度对应的真实值; $\bar{p}_{i,1}, \bar{p}_{i,2}, \bar{p}_{i,3}$ 为归一化后第 t_i 个航迹点的纬度、经度和高度的预测值; $|TE|$ 为预测集的元素个数。

图 8 为航迹特征的总体偏差随预测序列长度的变化趋势。从图 8 可以看出,模型 F₄ 和模型 F 的预测航迹特征的总体偏差随预测序列长度的增加而增大;模型 F₂ 和模型 F 的预测航迹特征的总体偏差随预测序列长度的增加基本保持平稳。当预测序列长度为 1 时,模型 F₄ 预测航迹特征的总体偏差是基准模型中最小的,偏差值为 0.049 6。模型 F 对应的偏差值为 0.041 3,与之相比,减小 0.008 3。预测序列

长度为其他设定步长时,改进 Transformer 模型相比于其他模型的变化参见表 7。

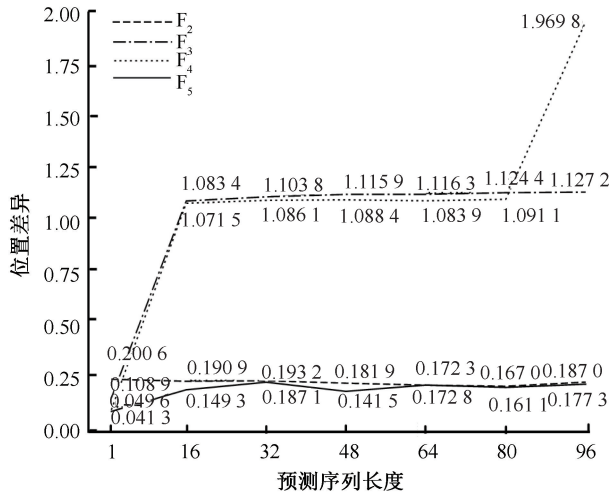


图 8 不同步长对比

Fig. 8 Comparison diagram of various steps

表 7 不同步长预测航迹特征的总体偏差变化(“-”为增大)

Table 7 Deviation difference with benchmark models under various steps(“-” indicates increase)

模型	1	16	32	48	64	80	96
F ₂	0.1593	0.0416	0.0061	0.0404	-0.005	0.0059	0.01
F ₃	0.0676	0.9341	0.9167	0.9744	0.9435	0.9633	0.9499
F ₄	0.0083	0.9222	0.8990	0.9469	0.9111	0.93	1.7925

5 结论

1) 相比于模型 F₁、模型 F₂、模型 F₃ 和模型 F₄ 等 4 种神经网络模型,模型 F 时间性能提高 14.4%;此外,采用 RMSE 和 MAE 作为评价指标,模型 F 对经度、纬度和高度等航迹特征预测的偏差更小,预测的稳定性更高。

2) 模型 F 预测表现优于传统的深度学习模型,包括模型 F₁、模型 F₂、模型 F₃,其平均 RMSE 和 MAE 分别为 0.027 和 0.021,比模型 F₂ 的平均偏差分别减小 0.009 和 0.012;比模型 F₃ 的平均偏差分别减小 0.031 和 0.024,比模型 F₄ 的平均偏差分别减小 0.001 和 0.002。此外,对于不同的预测序列长度的预测任务,改进 Transformer 模型的欧氏距离特征偏差相比于平均值,总偏差减少 0.37。

3) 在起飞爬升阶段,模型 F 与模型 F₄ 的预测航迹相比,经度、纬度和高度特征的偏差稍大,可在后续研究中通过进一步调参优化。

参考文献

- [1] International Civil Aviation Organization. Global air traffic management operational concept[Z]. 2005-05-02.
- [2] 张洪海, 汤一文, 许炎. TBO 模式下终端区进场交通流优化模型与仿真分析[J]. 航空学报, 2020, 41(7): 325-338.
ZHANG Honghai, TANG Yiwen, XU Yan. Optimizing arrival traffic flow in airport terminal airspace under trajectory based operation[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(7): 325-338.
- [3] PREVOST C G, DESBIENS A, HAGNON E. Extended Kalman filter for state estimation and trajectory prediction of a moving object detected by an unmanned aerial vehicle[C]. American Control Conference, 2007: 1 805-1 810.
- [4] 吕波, 王超. 改进的扩展卡尔曼滤波在航空器 4D 航迹预测算法中的应用[J]. 计算机应用, 2021, 41(增 1): 277-282.
LYU Bo, WANG Chao. Application of improved extended Kalman filtering in aircraft 4D trajectory prediction algorithm[J]. Journal of Computer Applications, 2021, 41(S1): 277-282.
- [5] ZHANG Junfeng, WU Xiaoguang, WANG Fei. Aircraft trajectory prediction based on modified interacting multiple model algorithm[J]. Journal of Donghua University: English Edition, 2015, 32(2): 180-184.
- [6] FAIRLEY G, MCGOVERN S. A kinematic/kinetic hybrid airplane simulator model: draft[C]. Proceedings of the ASME International Mechanical Engineering Congress & Exposition, 2008: 11-20.
- [7] FUKUDA Y, SHIRAKAWA M, SENOGUCHI A. Development and evaluation of trajectory prediction model[C]. Proceedings of the 27th International Congress of the Aeronautical Sciences, 2010: 19-24.
- [8] SCHUSTER W. Trajectory prediction for future air traffic management-complex manoeuvres and taxiing[J]. The

Aeronautical Journal, 2015, 119: 121–143.

- [9] 张军峰, 葛腾腾, 陈强, 等. 离场航空器四维航迹预测及不确定性分析[J]. 西南交通大学学报, 2016, 29(4): 800–806.
ZHANG Junfeng, GE Tengting, CHEN Qiang, et al. 4D trajectory prediction and uncertainty analysis for departure aircraft[J]. Journal of Southwest Jiaotong University, 2016, 29(4): 800–806.
- [10] 王兴隆, 许晏丰. 基于 AM-LSTM 的飞行区航空器滑行轨迹预测与冲突识别[J]. 中国安全科学学报, 2024, 34(1): 116–124.
WANG Xinglong, XU Yanfeng. Aircraft taxiing trajectory prediction and conflict risk identification in airfield area based on AM-LSTM[J]. China Safety Science Journal, 2024, 34(1): 116–124.
- [11] ZENG Weili, QUAN Zhibin, ZHAO Ziyu, et al. A deep learning approach for aircraft trajectory prediction in terminal airspace[J]. IEEE Access, 2020, 8: 151 250–151 266.
- [12] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]. In Advances in Neural Information Processing Systems, 2017: 6 000–6 010.
- [13] GUO Dongyue, ZHANG Zheng, ZHANG Jianwei, et al. FlightBERT++: a non-autoregressive multi-horizon flight trajectory prediction framework[C]. The Thirty-Eighth AAAI Conference on Artificial Intelligence, 2023: 127–134.
- [14] 冯霞, 孙琦琦, 左海超. 基于 Informer 的客机长时 4D 航迹预测方法[J]. 交通信息与安全, 2023, 41(4): 111–121.
FENG Xia, SUN Qiqi, ZUO Haichao. A method for predicting long-term 4D trajectory of airplanes based on Informer[J]. Journal of Transport Information and Safety, 2023, 41(4): 111–121.
- [15] 陈万通, 吴多, 刘庆, 等. 基于改进 PF 算法的 ADS-B 风场反演[J]. 传感技术学报, 2023, 36(7): 1 086–1 091.
CHEN Wantong, WU Duo, LIU Qing, et al. ADS-B wind field inversion based on improved PF algorithm[J]. Chinese Journal of Sensors and Actuators, 2023, 36(7): 1 086–1 091.
- [16] 张小江, 高秀华. 三次样条插值在机器人轨迹规划应用中的改进研究[J]. 机械设计与制造, 2008(9): 170–171.
ZHANG Xiaojiang, GAO Xiuhua. The research on the cubic splines in robot's trajectory planning[J]. Machinery Design and Manufacture, 2008(9): 170–171.

作者简介: 刘宏 (1985—), 男, 山西平遥人, 博士, 副教授, 主要从事空中交通管理、智能交通优化、无人机空管安全运行等方面的研究。E-mail: mliuhong@126.com。

