文章编号:1671-4598(2026)01-0051-08 DOI:10.16526/j.cnki.11-4762/tp.2026.01.007 中图分类号:TP391.4 文献标识码:A

基于时空门控 VAE 的 ADS-B 数据异常检测方法

蒋东旭，刘 蕾
（中国电子科技集团公司 第十五研究所，北京 100083）

摘要：广播式自动相关监视是新一代空中管理系统重要的组成部分，但由于 ADS-B 报文以明文形式广播且缺乏数据加密和认证，导致其极易受到欺骗干扰；针对以上问题，提出一种基于时空门控变分自编码器的 ADS-B 数据异常检测算法；该算法编码器通过采用双向 LSTM 建模局部时序特征，结合 3 层 8 头 Transformer 提取全局时空特征，并利用门控网络动态融合时空特征；引入变分推理生成潜在空间分布，约束模型对正常飞行模式的概率建模；解码器采用单层 LSTM 与 2 层 Transformer 的级联结构通过全连接层同步重建多维飞行参数；经实验测试，在不同攻击场景下，该模型可有效检测出 ADS-B 数据的各类异常，性能优于相关基线算法，为提升空中管理系统安全性提供了可行性方案。

关键词：广播式自动相关监视（ADS-B）；异常检测；长短期记忆神经网络（LSTM）；Transformer；变分自编码器（VAE）

Anomaly Detection Method for ADS-B Data Based on Spatio-Temporal Gated VAE

JIANG Dongxu，LIU Lei
（The 15th Research Institute of China Electronics Technology Group Corporation，Beijing 100083，China）

Abstract：Automatic dependent surveillance-broadcast (ADS-B) is a critical component of the new generation air traffic management system. However，its vulnerability to spoofing interference arises from the plaintext broadcasting of ADS-B messages without data encryption and authentication. To address this issue，a spatio-temporal gated variational autoencoder-based anomaly detection algorithm for ADS-B data is proposed. The encoder of this algorithm employs bidirectional long short-term memory (LSTM) modelling to capture its local temporal features，to extract global spatio-temporal features by combining a Transformer with three layers and eight heads，and to dynamically fuse these features by using a gating network. The variational inference is introduced to generate the distribution of latent space，which constrains the model to model the probability of normal flight patterns. The decoder of this algorithm incorporates a cascaded structure with single-layer LSTM and two-layer Transformer to simultaneously reconstruct multi-dimensional flight parameters through a fully connected layer. Experimental results demonstrate that the proposed model effectively detects various anomalies in ADS-B data under different attack scenarios，with a superior performance over relevant baseline algorithms，which provides a feasible solution for enhancing the security of air traffic management systems.

Keywords：ADS-B；anomaly detection；LSTM neural network；Transformer；variational autoencoder (VAE)

0 引言

广播式自动相关监视技术（ADS-B，automatic dependent surveillance broadcast）通过集成机载传感器与全球导航卫星系统获取参数，并向地面接收站以及临近空域的航空器广播飞机的高度、速度、经纬度等信息的监视技术[1]。该技术因具备定位精度高、覆盖范围广以及部署成本低等优势，被列为新一代空管系统的核心方案。然而，其开放的数据广播机制和公开的报文编码标准，导致极易受到恶意干扰和欺骗攻击[2]。攻击者通过模拟或篡改 ADS-B 数据以误导其他飞行器或地面控制系统甚至导致严重的空中事故[3]，因此针对 ADS-B 欺骗干扰的异常检测成为了 ADS-B 研究领域中的重要问题之一。

收稿日期:2025-06-24； **修回日期**:2025-09-29。
作者简介:蒋东旭(2000-),男,硕士研究生。
引用格式:蒋东旭,刘 蕾. 基于时空门控 VAE 的 ADS-B 数据异常检测方法[J]. 计算机测量与控制,2026,34(1):51-58.

针对 ADS-B 安全问题，国内外学者开展了很多研究，涌现出了很多方法。主要可分为两大类：广播认证和位置认证[4]。广播认证是在 ADS-B 通信数据中增加身份认证机制以防止非法攻击。可分为加密方案和非加密方案。加密的广播认证方案是消息发送前对 ADS-B 报文加密处理，收到消息后用存储的密钥进行数据还原，能够很好保证数据可靠传输，但是对内存和处理能力要求高，加密和解密过程会带来延迟[5]。非加密的广播认证方案包括基于硬件的指纹识别方案[6]，其存在适用范围窄仅对少数航班有效的问题。位置认证能通过对飞行器的飞行数据进行建模或统计分析，从而验证接收到的飞行器位置信息的准确性[7]。但是大规模部署难度较大，实用性不高，且容易受到多径效应的影响[8]。

随着机器学习技术的深入发展，机器学习算法不断应用在 ADS-B 的异常检测中。传统算法有：一类支持向量机方法（OC-SVM）[9]：将正常数映射到高维并构造最优超平面将偏离点判定异常，孤立森林算法（IF）[10]：通过快速隔离异常点实现异常高效实时筛查，局部异常因子（LOF）[11]：通过计算局部密度偏离度有效检测轨迹空间的离群点。除此之外，文献［12］将 LSTM 网络应用于 ADS-B 异常检测，各项指标优于传统方法。文献［13］在 LSTM 的基础上引入了差分处理思想，其模型能够学习到航迹数据的变化规律。文献［14］改进了传统 Transformer 算法设计出 Anomaly Transformer 算法，通过剪枝注意力头数极大提升速度和效率。

本文在前人研究的基础上提出了一种基于时空门控变分自编码器（ST-Gated VAE）的 ADS-B 数据异常检测算法。本文主要创新点如下：

1）动态时空融合机制。设置门控加权单元，通过可学习权重系数动态调节 LSTM 时序特征和 Transformer 空间特征的贡献度，有效捕捉了 ADS-B 轨迹的时空依赖性。

2）鲁棒概率空间建模。在潜在空间中引入变分贝叶斯约束，通过 KL 散度正则化迫使潜在编码服从高斯分布，有效提升模型对噪声干扰的鲁棒性。

3）层次化解码结构。采用 LSTM-Transformer 混合解码器，在重建阶段分别捕获局部时序连贯性和全局空间合理性。

实验结果表明，该算法对四类数据异常有较高的精确率、准确率和召回率，具有高效和广泛的实用性。

1 基础模型

1.1 LSTM 模型

LSTM[15]是一种特殊的循环神经网络，解决了传统 RNN 网络存在的长程依赖问题。LSTM 基本单元包括：细胞状态、输入门、遗忘门以及输出门。其结构如图 1 所示。其中 $\boldsymbol{h}_{t-1}$ 表示 $t-1$ 时刻 LSTM 的输出，$\boldsymbol{x}_t$ 表示 t 时刻输入，$\boldsymbol{C}_{t-1}$ 表示 $t-1$ 时刻细胞状态。

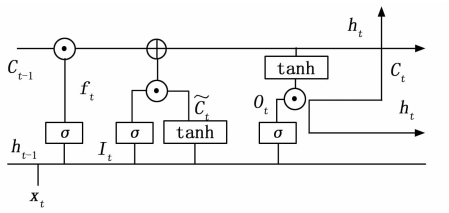

图 1 LSTM 结构图

LSTM 单元执行步骤的第一步，遗忘门 $\boldsymbol{f}_t$ 决定的是决定细胞状态中哪些信息需要被遗忘。它通过对 $\boldsymbol{h}_{t-1}$ 和 $\boldsymbol{x}_t$ 实施加权和偏移，再通过 sigmoid 函数得到 $\boldsymbol{f}_t$。可表示为：

$$\boldsymbol{f}_t = \delta(\boldsymbol{W}_f(\boldsymbol{x}_t, \boldsymbol{h}_{t-1}) + \boldsymbol{b}_f) \tag{1}$$

第二步，输入门 $\boldsymbol{I}_t$ 决定的是当前输入的新信息中哪些部分需要更新到细胞状态。首先通过对 $\boldsymbol{h}_{t-1}$ 和 $\boldsymbol{x}_t$ 加权和偏移并通过 sigmoid 得到 $\boldsymbol{I}_t$。其次通过对 $\boldsymbol{h}_{t-1}$ 和 $\boldsymbol{x}_t$ 加权和偏移并通过 tanh 函数得到候选细胞状态 $\widetilde{\boldsymbol{C}}_t$。可表示为：

$$\begin{cases} \boldsymbol{I}_t = \delta(\boldsymbol{W}_i(\boldsymbol{x}_t, \boldsymbol{h}_{t-1}) + \boldsymbol{b}_i) \\ \widetilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{W}_c(\boldsymbol{x}_t, \boldsymbol{h}_{t-1}) + \boldsymbol{b}_c) \end{cases} \tag{2}$$

第三步，实现当前细胞状态的更新。通过遗忘门 f_t 和 C_{t-1} 相乘确定遗忘与保留。通过输入门 I_t 与候选细胞状态 $\widetilde{C}_t$ 相乘确定新增与更新。再将二者相加便得到 t 时刻细胞状态的更新值 C_t。可表示为：

$$\boldsymbol{C}_t = \boldsymbol{f}_t * \boldsymbol{C}_{t-1} + \boldsymbol{I}_t * \widetilde{\boldsymbol{C}}_t \tag{3}$$

第四步，输出门 $\boldsymbol{O}_t$ 决定的是细胞状态中的哪些信息需要输出到下一个隐藏状态。通过对 $\boldsymbol{h}_{t-1}$ 和 $\boldsymbol{x}_t$ 加权和偏移并通过 sigmoid 得到 $\boldsymbol{O}_t$。最终的输出 $\boldsymbol{h}_t$ 由 $\boldsymbol{O}_t$ 和 $\boldsymbol{C}_t$ 共同决定。可表示为：

$$\begin{cases} \boldsymbol{O}_t = \delta(\boldsymbol{W}_o(\boldsymbol{x}_t, \boldsymbol{h}_{t-1}) + \boldsymbol{b}_o) \\ \boldsymbol{h}_t = \boldsymbol{O}_t * \tanh(\boldsymbol{C}_t) \end{cases} \tag{4}$$

其中：$\boldsymbol{W}$ 和 $\boldsymbol{b}$ 分别为 LSTM 的权重和偏移项，下标 f、i、o 分别表示遗忘门、输入门和输出门。

1.2 Transformer 模型

Transformer[16]是首个完全基于自注意力机制的序列建模框架，它摒弃了传统循环神经网络和卷积神经网络的串行计算结构，而通过并行化的注意力机制处理输入输出[17]。其核心采用编码器－解码器架构，编码器将输入序列转化为高维隐藏表示，解码器则基于该表示生成目标序列。通过自注意力层，Transformer 模型显著提升了长距离依赖的捕捉能力[8]。

Transformer 的自注意力机制包括：描述我们正在查找什么的查询矩阵 $\boldsymbol{Q}$，错略描述元素提供了什么的键

值 $\boldsymbol{K}$ 以及值向量 $\boldsymbol{V}$，计算过程为：

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^T}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{5}$$

其中：d_k 为键向量的维度，$1/\sqrt{d_k}$ 是缩放因子以避免向量积结果过大，softmax 为归一化指数函数。

多头自注意力机制通过并行计算多个自注意力机制，并将结果拼接在一起，从而捕捉序列中不同子空间的依赖关系[19]。多头自注意力机制的计算公式为：

$$\begin{cases}\text{Multihead}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \text{Concat}(head_1,\cdots,head_n)\boldsymbol{W}^O \\ head_i = \text{Attention}(\boldsymbol{Q}\boldsymbol{W}_i^Q,\boldsymbol{K}\boldsymbol{W}_i^K,\boldsymbol{V}\boldsymbol{W}_i^V)\end{cases} \tag{6}$$

其中：$\boldsymbol{W}^O$、$\boldsymbol{W}_i^Q$、$\boldsymbol{W}_i^K$、$\boldsymbol{W}_i^V$ 分别为查询、键、值和输出的权重矩阵，Concat 为拼接函数，$head_i$ 为第 i 个注意力头的输出结果。

前馈神经网络由两个全连接层和一个 RELU 激活函数构成。运算过程为：

$$\text{FFN}(\boldsymbol{x}) = \boldsymbol{W}_2(0,\boldsymbol{W}_1\boldsymbol{x}+\boldsymbol{b}_1)+\boldsymbol{b}_2 \tag{7}$$

其中：$\boldsymbol{W}_1$、$\boldsymbol{W}_2$、$\boldsymbol{b}_1$、$\boldsymbol{b}_2$ 为前馈神经网络的权重和偏置项。

1.3　VAE 模型

VAE[20] 是自编码器与变分推断相结合的生成模型。概率模型图如图 2 所示，它分为编码器和解码器。编码器将原始数据 $\boldsymbol{X}=\{x_1, x_2, \cdots, x_n\}$ 从高维特征分布映射到一个连续的隐变量空间（均值为 μ，方差为 σ^2）并输出均值和方差。然后从隐变量空间采样，得到 $\boldsymbol{Z}=\{z_1, z_2, \cdots, z_n\}$，解码器从潜在变量 Z 采样，将隐变量映射回原始空间，重构数据 X'。编码器和解码器可以采用不同的神经网络结构，从而使其适用于更多的任务和应用领域[21]。

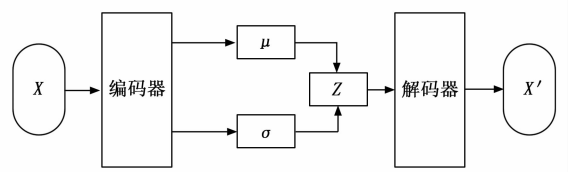

图 2　VAE 结构图

为了保证采样数据 Z 的概率分布符合，假设 Z 关于 X 的后验分布为 $P_\theta(Z \mid X)$，此处后验分布应尽可能近似于正态分布，从而达到先验分布 $P_\theta(Z)$ 近似正态分布假设。

$$P_\theta(Z) = \sum\nolimits_x P_\theta(Z \mid X)P_\theta(X) = \sum\nolimits_x N(0,1)P_\theta(X) = N(0,1)\sum\nolimits_x P_\theta(X) = N(0,1) \tag{8}$$

此时，先验分布与后验分布就都符合标准正态分布了。

2　ADS-B 异常检测算法设计

2.1　模型设计

本文提出一种基于 ST-Gated VAE 的 ADS-B 异常检测模型，该模型是一个无监督深度学习模型，其结构如图 3 所示。

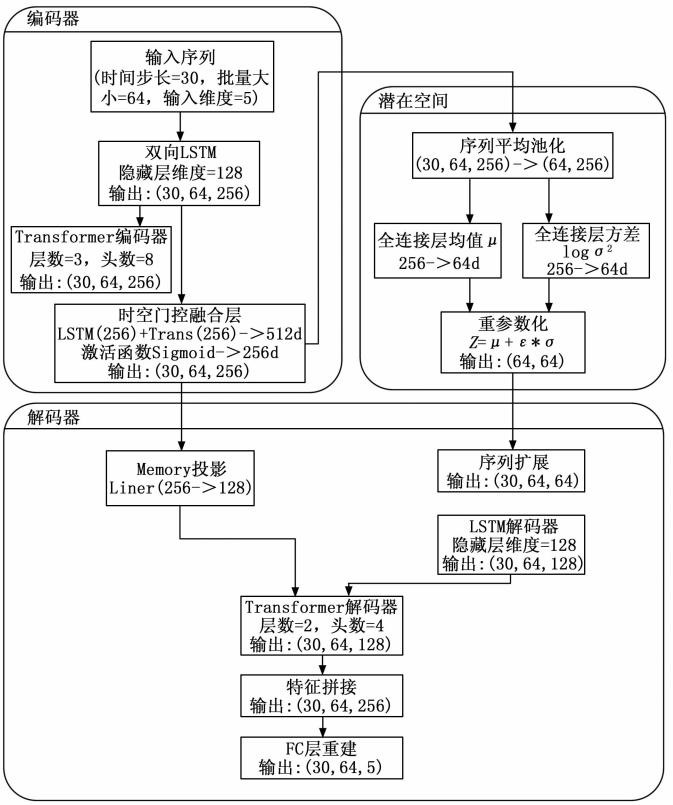

图 3　ST-Gated VAE 模型图

该架构采用编码器－潜在空间－解码器的三级设计，编码器通过双向 LSTM 和 Transformer 双分支提取时空特征，经门控单元动态融合；潜在空间采用 VAE 结构进行概率空间建模；解码器用 LSTM 和 Transformer 重构轨迹数据。

2.2　双向时空编码器

2.2.1　双向 LSTM 局部时序建模

LSTM 是一种特殊的循环神经网络，能有效捕捉时间序列数据中长期依赖关系。在本文 ST-Gated VAE 算法中，采用双向 LSTM 进行局部时序建模，分别从正向和反向两个方向处理输入序列，以同时捕捉过去和未来的上下文信息。

设输入序列 $\boldsymbol{X}=\{x_1, x_2, \cdots, x_T\}$，其中 $\boldsymbol{x}_t$ 是一个 5 维向量（高度，速度，航向角，经度，纬度）。前向 LSTM 由 $\boldsymbol{x}_1$ 到 $\boldsymbol{x}_T$ 处理输入序列：$\overrightarrow{\boldsymbol{h}_t}=LSTM(x_t, \overrightarrow{h_{t-1}})$，反向 LSTM 由 $\boldsymbol{x}_T$ 到 $\boldsymbol{x}_1$ 处理输入序列：$\overleftarrow{\boldsymbol{h}_t}=LSTM(x_t, \overleftarrow{h_{t-1}})$。每个 LSTM 单元包含 128 个隐藏单元，通过双向结构：$\boldsymbol{h}_t^{\text{LSTM}}=[\overrightarrow{h}_t; \overleftarrow{h}_t]$，每个时间步的

隐藏状态维度为 256，最终得到时序特征 $\boldsymbol{H}^{\mathrm{LSTM}}=\{h_1, h_2, \cdots, h_T\}$，其中 $\boldsymbol{h}_t$ 是一个 256 维向量，编码了该位置的时序信息。

2.2.2 Transformer 全局时空建模

虽然 LSTM 能够捕捉时序依赖关系，但难以建模序列中的全局依赖关系。因此，在模型中引入 Transformer 编码器进行全局时空建模。Transformer 编码器由多个自注意力层堆叠而成，每个自注意力层包含多头注意力机制和前馈网络。

在 ST-Gated VAE 算法中，Transformer 编码器包含 3 个编码层。每个层有 8 个注意力头，模型维度为 256，这种设计使模型能处理任意位置之间的依赖关系，从而捕捉全局模式。其中 Transformer 编码器的多头注意力机制可以表示为：

$$\mathrm{MultiHead}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{Concat}(head_0,\cdots,head_h)*\boldsymbol{W}^O \tag{9}$$

其中：$\boldsymbol{Q}$ 为查询，$\boldsymbol{K}$ 为键，$\boldsymbol{V}$ 为值矩阵，h 表示注意力头的数量，$\boldsymbol{W}^O$ 是用于线性变换的可学习参数矩阵。

2.2.3 自适应时空门控

为了发挥 LSTM 和 Transformer 各自优势，在算法中引入自适应时空门控机制。通过动态的调整 LSTM 和 Transformer 的输出权重，实现自适应融合。

分别将 LSTM 和 Transformer 的输出拼接，得到一个 512 维的向量。之后通过一个线性曾将其映射到 1 个单元，并使用激活函数 sigmoid 得到门控值 g_t。门控值 g_t 用于加权融合 LSTM 和 Transformer 的输出，从而实现自适应时空门控。门控机制的数学表达式可以表示为：

$$g_t=\sigma(W_g[H^{\mathrm{LSTM}};H^{\mathrm{Trans}}]+b_g \tag{10}$$

其中：σ 表示 sigmoid 函数，W_g 和 b_g 是可学习参数。W_g 和 b_g 作为模型整体的一部分，在训练过程中通过端到端的反向传播进行优化，采用 Adam 优化器更新，并应用梯度裁剪（阈值为 1.0）确保稳定性。最后融合特征表示为：

$$H^{\mathrm{Fused}}=g_t\odot H^{\mathrm{LSTM}}+(1-g_t)H^{\mathrm{Trans}} \tag{11}$$

2.3 概率空间建模

在编码器提取特征后，使用变分自编码器（VAE）建模潜在空间。VAE 通过重新参数化技巧，将潜在变量从输入数据中解耦，从而实现数据的生成和异常检测。具体实现为：ST-Gated VAE 算法将融合后的特征张量沿时间维度取平均，得到一个 256 维的全局特征向量，然后通过两个独立的线性层分别计算潜在空间的均值 μ 和对数方差 $\log\sigma^2$ 每个层的输出纬度为 64，经实验验证，纬度为 64 在重构精度与正则化强度间取得最佳平衡。最后通过重新参数化技巧，我们从这些参数中采样潜在变量 z。z 的计算可表示为：

$$\begin{cases}\mu=\boldsymbol{W}_\mu*\mathrm{Avg}(H^{\mathrm{Fused}})+\boldsymbol{b}_\mu(\boldsymbol{W}_\mu\in R^{256\times 64})\\ \log\sigma^2=\boldsymbol{W}_{\log\sigma^2}*\mathrm{Avg}(H^{\mathrm{Fused}})+\boldsymbol{b}_{\log\sigma^2}\\ z=\mu+\in\odot\exp(0.5\log\sigma^2)(\in\sim\boldsymbol{N}(0,1))\end{cases} \tag{12}$$

其中：$\boldsymbol{W}_\mu$、$\boldsymbol{b}_\mu$、$\boldsymbol{W}_{\log\sigma^2}$ 和 $\boldsymbol{b}\log\sigma^2$ 是可学习参数。

2.4 多机制重构解码器

解码器包括 LSTM 和 Transformer 两个部分。LSTM 解码器用于重构时间序列，而 Transformer 解码器通过交叉注意力机制利用解码器的全局特征重构长程依赖。通过这种有机的结合，能更全面的重构输入数据和评估数据的异常程度。

具体地，潜在变量 z 被扩展为与输入序列相同长度的序列，然后通过 LSTM 解码器和 Transformer 解码器进行处理。LSTM 解码器由一个 LSTM 组成，输入维度为 64，隐藏层大小为 128，Transformer 解码器由 2 个 Transformer 解码器层组成。每层有 4 个注意力头，维度为 128，最后将 LSTM 和 Transformer 的输出拼接，通过线性曾映射带输入维度 5，重构过程的数学表达式为：

$$\begin{cases}z_{\mathrm{expanded}}=z\otimes 1_{L\times 1\times d_z}\\ \mathrm{LSTM\ decoded}=\mathrm{LSTM}_{\mathrm{dec}}(z_{\mathrm{expanded}})\\ M=W_mH^{\mathrm{enc}}+b_m\\ \mathrm{Transformer\ decoded}=\\ \quad\mathrm{Transformer}_{\mathrm{dec}}(\mathrm{LSTM}_{\mathrm{dec}},K=V=M)\\ \mathrm{Reconstruction}=\\ \quad W_{\mathrm{out}}[LSTM\ \mathrm{decoded};\mathrm{Transformerdecoded}]+b_{\mathrm{out}}\end{cases} \tag{13}$$

其中：L 是序列长度，d_z 是潜在变量的维度，$1_{L\times 1\times d_z}$ 是一个 $L\times 1\times d_z$ 的全 1 矩阵，用于将潜在变量扩展为与输入序列相同长度的序列，M 是编码器输出的特征投影，H^{enc} 是编码器 transformer 的输出，$\boldsymbol{W}_m$、$\boldsymbol{b}_m$、$\boldsymbol{W}_{\mathrm{out}}$ 和 $\boldsymbol{b}_{\mathrm{out}}$ 是可学习参数。

2.5 多目标损失函数

$$\mathrm{Reconloss}=\frac{1}{N}\sum_{i=1}^{N}(x_i-Recon_i)^2 \tag{14}$$

其中：N 是数据点的数量，x_i 表示输入数据的第 i 个数据点，$Recon_i$ 表示重构数据的第 i 个数据点。

KL 散度损失是衡量编码器输出的潜在变量分布和先验分布的差异：

$$KLloss=-\frac{1}{2}\sum_{i=1}^{N}1+\log\sigma_i^2-\mu_i^2-\sigma_i^2) \tag{15}$$

其中：μ_i 和 δ_i^2 分别是第 i 个数据点的潜在变量的均值和方差。

本模型的损失函数采用的是重构损失和 KL 散度损失的加权和：

$$Loss=Reconloss+\lambda*KLloss \tag{16}$$

其中：λ 是一个超参数，控制 KL 散度损失的权重。本实验 λ 取 0.001。

3　实验设计与结果分析

3.1　数据集的构建

3.1.1　数据获取与划分

本实验数据选用的是公共服务平台飞常准 ADS-B 中的数据，飞常准的 ADS-B 系统是基于广播式自动相关监视技术构建的航班追踪平台，目前航班数据实现中国国内全覆盖，且数据更新频率可达每秒 1 次，远超 Flightradar24 等平台，适合作为实验数据。

为增强算法的健壮性，实现各种距离、地形、不同时段航班和不同航空公司的飞行数据异常检测，数据采集如表 1 所示，本文选取以北京为起点，分别在北京的东向、南向和西向选择各方向的短途、中途、长途共 9 个目的地 D1～D9，且每个目的地选取不同时段与不同航空公司的近 7 天的航迹，每条航迹是一个 json 格式数据，其中包含完整的从起到落的 ADS-B 数据，总共约 30 万个 ADS-B 数据点。分别在 9 个目的地的航线中，以 7∶1.5∶1.5 的比例划分为训练集、验证集和测试集。

表 1　数据采集

方向	距离/km		
	短途(<1 000)	中途(1 000～2 000)	长途(>2 000)
东向	D1:北京－>大连:80 条	D2:北京－>上海:60 条	D3:北京－>福州:60 条
南向	D4:北京－>郑州:80 条	D5:北京－>武汉:60 条	D6:北京－>广州:60 条
西向	D7:北京－>西安:80 条	D8:北京－>兰州:60 条	D9:北京－>成都:60 条

3.1.2　数据预处理

对获取的数据集进行预处理，处理的流程如图 4 所示。其步骤为（1）数据加载与处理。进行读取数据，将数据转换为 Pandas DataFrame 格式，并按照时间戳 updatetime 字段对每条行航迹数据进行时间排序，后续处理都是将原始数据硬编码提取为五维 ADS-B 数据（高度，速度，航向角，经度，纬度）用于后续实验。（2）缺失值处理，对中间数据采用三次样条插值法，通过拟合多项式曲线估计缺失值并进行填补，此方法优于均值/中位数填补，适用于时序连续性数据。对首尾缺失采用线性插值方法，防止样条过冲。（3）滑动平均去噪。使用大小为 5 的窗口，对每个特征分别进行滑动平均以实现去噪并保留趋势特征。（4）数据归一化。采用 Z-score 标准化方法，通过对数据减去均值并除以标准差，将数据转化为均值为 0 标准差为 1 的分布，有助于模型的训练[22]。最后将处理后的数据转回 json 格式并保留浮点精度，结构化存储预处理结果。

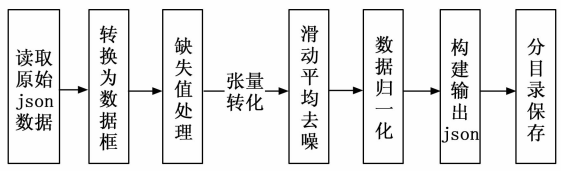

图 4　数据预处理流程

3.2　攻击构建

由于获取的 ADS-B 数据都是正常数据，为模拟真实飞行状况可能遇到的各种欺骗攻击，本实验在测试集数据中，针对部分数据进行攻击模拟。分别针对高度、速度、航向角以及经纬度 5 个特征设计异常数据，分别为 4 大类：GPS 缓变欺骗、航向缓变、分阶段爬升以及速度突增。具体攻击类型和攻击如表 2 所示。

3.3　评价指标

本文异常检测属于二分类问题，预测与真实类别的关系分为四种情况，如表 3 所示。TP 正常样本被正确预测为正常样本的数量，FN 指正常样本被错误预测为异常样本的数量，FP 指异常样本被错误预测为正常样本的数量，TN 指异常样本被正确预测为异常样本的数量。[23]

表 2　攻击操作与方法

攻击类型	数据操作	具体攻击方法
GPS 缓变欺骗	①逐步偏移经纬度	D1:经度总偏移 6°纬度总偏移 2.5°,每步变化 0.035°,SNR 基准 38.5 dB。
	②航向角渐变	D4:经度总偏移 7.2°,纬度总偏移 3°,每步变化 0.04°,SNR 基准 36 dB。
	③卫星信噪比抛物线衰减	D7:经度总偏移 5°,纬度总偏移 2°,每步变化 0.03°,SNR 基准 40 dB。
航向缓变	①分阶段航向角变化	D2:15°分 3 阶段完成,叠加 2°正弦干扰,横滚率限制 3°/s。
	②叠加正弦震荡	
	③横滚率动态计算	D8:12°分 4 阶段完成,叠加 1.8°正弦干扰,横滚率限制 2.5°/s。
分阶段爬升	①多阶段高度爬升	D3:两阶段爬升 1 200－>800 ft/min,垂直加速度 0.2 m/s^2。
	②垂直速度计算	
	③垂直加速度约束	D6:两阶段爬升 1 800－>600 ft/min,垂直加速度 0.3 m/s^2。
速度突增	①加速曲线生成速度	D5:150 节两次加速,滚动速度变化均值 18.7 节/5 点。
	②标记速度脉冲	
	③滑动窗口速度变化率计算	D9:180 节线性增长,滚动速度变化峰值 22.3 节/5 点。

表 3 二分类情况

真实	预测	
	预测为正常样本	预测为异常样本
正常样本	TP	FN
异常样本	FP	TN

本实验通过计算精确率 P（Precision）、召回率 R（Recall）、F_1 分数来衡量模型能力。其公式为：

$$\begin{cases} P = \dfrac{TP}{TP + FP} \\ R = \dfrac{TP}{TP + FN} \\ F_1 = \dfrac{P * R}{P + R} \end{cases} \tag{17}$$

3.4 实验设定

本实验的实验参数设定如下：采用滑动窗口法分割时序数据，窗口长度为 30、步长为 15，以确保序列重叠覆盖；训练时批量大小为 64，学习率为 0.000 1，共训练 300 轮；使用 Adam 优化器并施加梯度裁剪（阈值 1.0）防止梯度爆炸。

实验在支持 CUDA 的 GPU 环境下运行，输入数据经预处理后验证为（30，5）维度，训练过程实时监控训练集和验证集损失，仅保留验证损失最低的模型。模型轻量化设计保证实时性。异常检测通过计算重构误差动态设定阈值。高精度且低延迟的识别异常 ADS-B 数据。

3.5 异常检测及类型识别流程

异常检测及类型识别总体流程如图 5 所示。

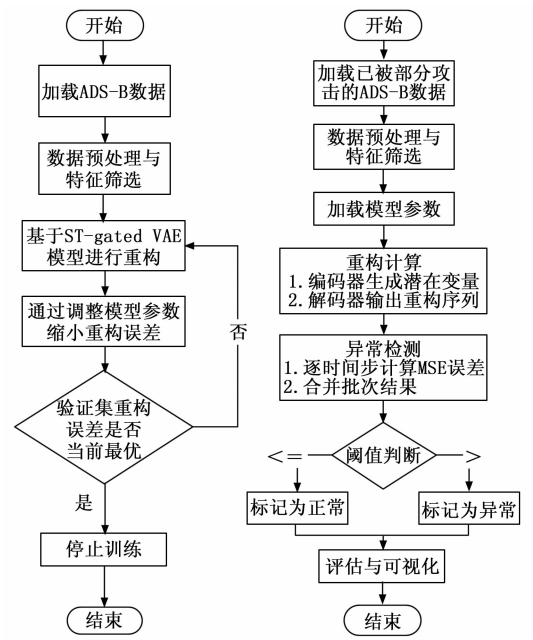

图 5 a 训练阶段（左）b 测试阶段（右）

图 5 a 是训练阶段的整体流程，首先加载 ADS-B 数据并进行数据预处理与筛选，将预处理后的数据输入到本实验模型 ST-Gated VAE 中，通过其编码器－解码器的结构实现数据的重构，通过调整模型 ST-Gated VAE 参数以缩小误差，并判断是否为当前最优，是则保存模型至停止训练，否则继续训练。

图 5 b 是测试阶段的整体流程，首先加载已经被部分攻击和打了标签的测试集数据，并对数据进行预处理与筛选，之后加载已经训练好的模型参数。编码器生成潜在变量，解码器输出重构序列，并逐时间步计算 MSE 误差，之后合并批次结果。若误差小于等于阈值则标记为正常序列，反之标记为异常数据，最后将预测结果与标签对照计算精确率、召回率和 F_1 分数等评价指标。

3.6 性能分析

3.6.1 实验结果

训练过程的训练损失和验证损失如图 6 所示。通过综合模型收敛效果、测试精度以及训练时间成本，本实验训练 300 epoch 时模型收敛且精度高、训练成本低。

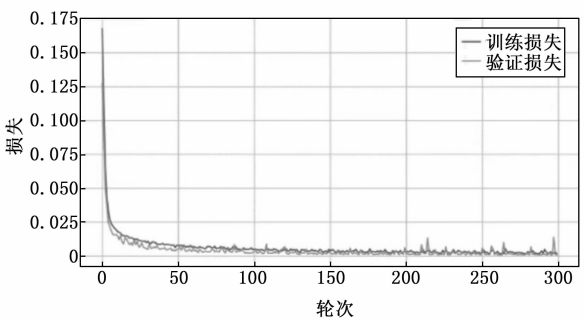

图 6 训练损失图

各类攻击的检测准确率如表 4 所示。

表 4 实验结果

评价指标	GPS 缓变欺骗	航向缓变	分阶段爬升	速度突增	全局平均
精确率/%	95.7	95.7	94.4	90.8	98.6
召回率/%	97.2	99.9	99.8	99.4	98.8
F_1 分数/%	96.5	97.7	97.0	94.9	98.7

1）GPS 缓变欺骗检测：以 96.5% 的 F_1 分数实现微小位置偏移的精准识别，97.2% 召回率验证模型对渐变特征的敏感捕获能力。

2）航向缓变攻击检测：性能最优，99.9% 召回率证实模型有效解耦分阶段航向变化与正弦干扰的复合特征。

3）分阶段爬升检测：99.8% 高召回率证明模型对多阶段爬升模式及垂直加速度约束的鲁棒识别能力。

4）速度突增检测：99.4% 召回率突显模型对脉冲式速度异常的强敏感性。

综上，各类别的精确率、召回率和 F_1 分数都达到

了 90%以上，且全局预测（正常样本与异常样本总体预测）的精确率、召回率和 F_1 分数都在 98%以上，通过结果验证了本实验模型的低成本与高精度的特点，验证了算法的有效性。

3.6.2　计算效率

算法的计算效率充分考虑对 ADS-B 数据实时检测需求：

1）时间复杂度恒定：固定序列长度为 30，避免了长序列填充计算开销，使计算复杂度与航班数量无关。

2）硬件加速优化：自动选择最佳计算设备，batch_size 设置为 64，充分利用 GPU 并行效率。

3）实时性保障：解码器仅依赖固定长度隐变量，无递归依赖，且无动态调整开销。

算法在实际部署中可满足：推理延迟实测＜20 ms，吞吐量≥3 200 样本/秒。可支持 2 000＋架次/日的机场实时处理需求。

3.6.3　对比实验

为了客观验证模型的性能，选取目前具有代表性 4 个算法作为对比实验，以验证模型的性能。分别是：一类支持向量机方法（OC-SVM）、孤立森林（IF）、局部异常因子（LOF）以及 Anomaly Transformer。实验结果如表 5 所示，本文模型 ST-Gated VAE 模型的各项指标均优于其余对比实验模型，验证了本文模型的性能与应用价值。

表 5　对比实验结果

消融实验模型	评价指标	GPS 缓变欺骗	航向缓变	分阶段爬升	速度突增
OC-SVM	精确率/%	69.8	75.3	79.8	72.1
	召回率/%	66.6	71.8	77.0	68.5
	F_1 分数/%	68.2	73.5	78.4	70.3
IF	精确率/%	74.2	72.5	67.3	90.2
	召回率/%	70.5	67.8	63.6	88.0
	F_1 分数/%	72.3	70.1	65.4	89.1
LOF	精确率/%	83.1	81.2	74.0	76.8
	召回率/%	80.0	79.2	69.7	73.6
	F_1 分数/%	81.5	80.2	71.8	75.2
Anomaly Transformer	精确率/%	92.6	92.4	89.7	85.3
	召回率/%	86.1	87.3	83.1	93.8
	F_1 分数/%	89.2	89.8	86.3	89.3
ST-Gated VAE	精确率/%	95.7	95.7	94.4	90.8
	召回率/%	97.2	99.9	99.8	99.4
	F_1 分数/%	96.5	97.7%	97.0	94.9

深入分析表 5 对比实验结果，不同基线模型的性能局限反映了其对 ADS-B 数据特性适配性的差异：OC-SVM 受限于核方法在高维时空序列上的表征能力；IF 对点异常（如速度突增）敏感但对需捕捉长期演化的缓变异常效果骤降；LOF 依赖空间密度，在非空间主导异常上表现不佳；即使是先进的 Anomaly Transformer，其关联差异机制对序列整体偏离的敏感性仍显不足。ST-Gated VAE 的核心优势在于其门控融合机制（LSTM 局部动态＋Transformer 全局依赖）与概率生成框架（VAE）的结合：前者精确建模时空特征演化，后者通过重构误差直接量化序列整体异常程度，使其对隐蔽性极强的缓变异常具备更强的检测能力。

3.6.4 消融实验

为验证各组件和本模型创新点的有效性，本组实验设置消融实验单一改变每个组件与元模型对比，验证各组件有效性。[24] 本组实验共设计五组消融实验，分别是：

1）移除门控机制。将原模型中编码器中的门控机制移除，改用线性投影。其余不变。

2）仅用 LSTM 编码器。移除 Transformer 编码器，将 LSTM 输出维度调整为 256，删除门控机制，直接使用 LSTM 输出，修改 VAE 输入维度为 256，保留原解码器。

3）仅用 Transformer 编码器。移除 LSTM 编码器，增强 Transformer 编码器为 6 层，8 头注意力。其余所有组件与流程与原模型一致。

4）移除 VAE 重参数化。将变分自编码器改为普通自编码器（AE），以检验 VAE 核心组件概率编码和潜在空间随机性的实际贡献。

实验结果如表 6 所示，ST-Gated VAE 模型的各项指标均优于其余四组消融实验，验证了动态门控融合机制、概率空间建模和多机制编码与解码结构对本模型的积极作用，验证了各组件和创新点的有效性。

表 6　消融实验结果

消融实验模型	评价指标	GPS 缓变欺骗	航向缓变	分阶段爬升	速度突增
(1)Remove_gata	精确率/%	83.6	82.9	77.0	65.8
	召回率/%	87.9	87.0	79.4	77.6
	F_1 分数/%	85.6	84.9	78.2	71.2
(2)OnlyLstm_VAE	精确率/%	88.7	88.5	85.2	77.5
	召回率/%	81.7	83.2	81.9	83.2
	F_1 分数/%	85.0	85.8	83.5	80.2
(3)OnlyTrans_VAE	精确率/%	82.5	82.3	78.1	68.3
	召回率/%	81.9	83.8	84.9	86.8
	F_1 分数/%	82.2	83.1	81.4	76.4
(4)ST-Gated_AE	精确率/%	86.1	86.3	82.6	73.3
	召回率/%	80.1	85.2	84.7	83.3
	F_1 分数/%	82.9	85.8	83.7	77.9
ST-Gated VAE	精确率/%	95.7	95.7	94.4	90.8
	召回率/%	97.2	99.9	99.8	99.4
	F_1 分数/%	96.5	97.7	97.0	94.9

4 结束语

ST-Gated VAE 算法，通过结合时空门控机制与变分自编码器，有效的捕捉了 ADS-B 数据中的时序依赖和全局关系。该模型通过门控机制动态融合 LSTM 和 Transformer 的特征表示，同时利用变分自编码器实现数据的高效编码与重构。实验结果表明，与目前已经存在的传统算法相比，ST-Gated VAE 表现出更高的精确率和召回率，验证了 ST-Gated VAE 的高性能。通过消融实验验证了本文的创新点和各部分组件的有效性。因此，该方法兼具结构可解释性与工程实用性，可为 ADS-B 异常检测提供高可信度解决方案。

参考文献：

[1] 郑 超. 基于 ADS-B 的 DOA/GNSS 融合定位方法 [J]. 电子技术与软件工程，2021 (7)：82 - 83.

[2] 孙保明，陈 娟，谷志鸣. ADS-B 系统面临的风险及应对措施 [J]. 通信学报，2024，45 (s1)：114 - 118.

[3] COSTIN A，FRANCILLON A. Ghost in the air (traffic)：on insecurity of ADS-B protocol and practical attacks on ADS-B devices [J]. Computers & Security，2021，109：102450.

[4] STROHMEIER M，LENDERS V，MARTINOVIC I. On the security of the automatic dependent surveillance-broadcast protocol [J]. IEEE Communications Surveys & Tutorials，2015，17 (2)：1066 - 1087.

[5] 陈 敏，马志坤，吴仁彪. 一种适用性广的 ADS-B 异常数据检测方法 [J]. 信号处理，2023，39 (5)：875 - 885.

[6] DANEV B，ZANETTI D，CAPKUN S. On physical-layer identification of wirelessdevices [J]. ACM Computing Surveys，2012，45 (1)：1 - 29.

[7] 张晓磊. 面向 ADS-B 通信数据的多变量时间序列异常检测算法研究 [D]. 成都：电子科技大学，2024，18 - 20.

[8] DING J L，ZOU Y K，WANG J，et al. ADS-Banomaly data detection model based on deep learning [J]. Acta Aeronautica et Astronautica Sinica，2019，40 (12)：323220.

[9] LI K L，HUANG H K，TIAN S F，et al. Improving one-class SVM for anomaly detection [C] //Proceedings ofthe International Conference on Machine Learning and Cybernetics，2003：3077 - 3081.

[10] ZHANG K Z，KANG X D，LI S T. Isolation forest for anomalydetection in hyperspectralimages [C] //Proceedings of the IGARSSIEEE International Geoscience and Remote Sensing Symposium，2019：437 - 440.

[11] BREUNIG M M，KRIEGEL H P，NG R T，et al. LOF：identifying density-based local outliers [C] //Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. Dallas，TX，USA：ACM，2000：93 - 104.

[12] HABLER E，SHABTAI A. Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages [J]. Computers&Security，2018，78：155 - 173.

[13] FRIED A，LAST M. Facing airborne attacks on ADS-Bdata with autoencoders [J]. Computers&Security，2021，109：102390.

[14] XU J，WU H，WANG J，et al. Anomaly Transformer：Time Series Anomaly Detection with Association Discrepancy [C] //Proceedings ofthe International Conference on Learning Representations (ICLR). 2022：1 - 10.

[15] HOCHREITER S，SCHMIDHUBER J. Long short-term memory [J]. Neural Computation，1997，9 (8)：1735 - 1780.

[16] VASWANI A，SHAZEER N，PARMAR N，et al. Attention is All You Need [C] //Proceedings of the 31st International Conference on Neural Information Processing Systems，Red Hook，NY，USA：Curran Associates Inc.，2017：6000 - 6010.

[17] 李石峰，罗 晰，刘晓茹，等. 基于 Transformer 架构的端到端视频异常检测方法 [J]. 计算机技术与发展，2025，35 (6)：49 - 55.

[18] 杨 斌. 基于 Transformer 的预训练模型综述 [J]. 现代计算机，2024，30 (24)：16 - 22.

[19] 高建树，郝世宇，党一诺. 基于长短期记忆网络-Transformer 模型参数优化的锂离子电池剩余使用寿命预测[J/OL]. 汽车工程师，2025. https://kns.cnki.net/kcms2/article/abstract?v=7KCVMXbQLqKSiHB3Yi7sz-HqtM9_oIvB_DRDPN44JICRFK9AubH6QCGmvw1sYB_ZI5UNsWSAs12hBeC2VM1V4_CksXJ2gR8rjE6xWuyHx9A_Ls2ROKl6Y3SIA7xGnYQkrnd3inQxrGaCHhAl0-2uty0iyDUx_bb6EGuw1iAho2JsZEfy_sIY9w==&uniplatform=NZKPT&language=CHS

[20] KINGMA D P，WELLING M. Auto-encoding Variational Bayes [J]. ArXiv Preprint ArXiv：1312.6114，2013.

[21] 史绪钊. 变分自编码器 VAE 在推荐系统中的应用 [D]. 北京：北京邮电大学，2023.

[22] 俞文静，王代涛，黄舒怡，等. 基于 Swin-Transformer 的岩石自动分类识别 [J]. 现代计算机，2024，30 (13)：15 - 20.

[23] 罗 鹏，王布宏，李腾耀. 基于 BiGRU-SVDD 的 ADS-B 异常数据检测模型 [J]. 航空学报，2020，41 (10)：281 - 291.

[24] 丁建立，张琪琪，王 静，等. 基于 Transformer-VAE 的 ADS-B 异常检测方法 [J]. 系统工程与电子技术，2023，45 (11)：3680 - 3689.