

低轨卫星缺失时序数据的模式识别方法*

郭思尧¹, 颜 博², 翟 磊², 鲍军鹏¹, 宋 宴³, 张 超², 陈亚军²

(1 西安交通大学计算机科学与技术学院 西安 710049

2 中国电子科技集团公司第五十四研究所 石家庄 050081

3 陆军航空兵研究所 北京 101100)

摘要: 在航天应用中, 低轨卫星经常会由于原始数据缺失而影响卫星时序数据模式识别结果, 降低准确率。针对该问题提出了一种新型 MR-GRU 模型, 可有效处理缺失时序数据, 并获得较好的模式识别准确率。区别于传统模型的全缺失数据的方法, MR-GRU 模型直接在缺失时序数据上运用神经网络进行训练, 对传统门控循环单元结构进行了改进, 增加了两个新变量: 掩蔽项和衰减项。掩蔽项作用于输入, 衰减项作用于输入和隐层单元输出。MR-GRU 模型不仅能够保持时序数据固有的时间特性, 还能有效提高模式识别精度。在卫星时序数据上的模式识别试验表明, MR-GRU 模型准确率优于传统模型。

关键词: 数据缺失; 门控循环单元; 卫星时序数据; 模式识别

中图分类号: TP391.4 **文献标识码:** A **文章编号:** CN11-1780(2021)06-0038-09

DOI: 10.12347/j.ycyk.20201211001

引用格式: 郭思尧, 颜博, 翟磊, 等. 低轨卫星缺失时序数据的模式识别方法[J]. 遥测遥控, 2021, 42(6): 38-46.

A pattern recognition method for LEO satellite time series data with missing data

GUO Siyao¹, YAN Bo², ZHAI Lei², BAO Junpeng¹, SONG Yan³, ZHANG Chao², CHEN Yajun²

(1. School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China;

2. The 54th Institute of China Electronics Technology Group Corporation, Shijiazhuang 050081, China;

3. Army Aviation Research Institute, Beijing 101100, China)

Abstract: In aerospace applications, Low Earth Orbit (LEO) satellites often lose some parts of raw data, which will disturb pattern recognition on satellite time series data and decline accuracy. A novel MR-GRU model is proposed, which can achieve a high accuracy on satellite time series data while some data are randomly missing. The MR-GRU model directly trains a recurrent neural network on the incomplete time series data, instead of the traditional way that tries to complement the missing data. The common Gated Recurrent Unit (GRU) model is improved to MR-GRU model. Two terms are expanded, i.e. masking term and attenuation term. The masking term is applied to the input at each time, and the attenuation term is applied to the input and output of each hidden unit. Consequently, the inherent time characteristics of time series data are ensured by the MR-GRU model, while the accuracy of pattern recognition is increased. According to the experimental results on satellite time series data, it is shown that the MR-GRU model is superior to the traditional models.

Key words: Data missing; Gated recurrent unit; Satellite time series data; Pattern recognition

DOI: 10.12347/j.ycyk.20201211001

Citation: GUO Siyao, YAN Bo, ZHAI Lei, et al. A pattern recognition method for LEO satellite time series data with missing data[J]. Journal of Telemetry, Tracking and Command, 2021, 42(6): 38-46.

引 言

在航天领域以及医学、地球科学、天文学、生物学等诸多领域中, 时间序列数据无处不在。现实中

*基金项目: 国家自然科学基金项目“多应用负载性能干扰预测与隔离相关问题研究”(项目编号: 61672421)

收稿日期: 2020-12-11 收修改稿日期: 2021-06-18

由于处于非测控区、数据不便采集、意外事件、设备故障等原因,会造成部分观测数据丢失^[1]。解决时间序列数据缺失的一般思路是忽略缺失数据或者补全缺失值。最简单方案是直接忽略缺失数据,只对观察到的数据进行分析。但是当数据缺失率较高时,这种方法性能欠佳^[2]。另一种思路是用替换值来补全缺失数据,即数据填充。其中,光滑法^[2]、插值法^[3]、样条法^[4]等方法比较简单,在实践中很常用。然而,这些方法不能利用变量之间的相关性,也不能利用复杂模式进行预测评估。光谱分析^[5]、核方法^[6]、电磁算法^[7]、矩阵完备^[8]和矩阵分解^[9]等方法则用不同手段先预测评估缺失数据,然后再替代数据中的缺失值^[2]。但是计算评估方法有一定程度的缺陷。例如,难以有效处理随机丢失的数据,不能处理长度变化的时间序列数据,并且这些方法相对复杂度较高^[10]。

近几年,循环神经网络 RNN (Recurrent Neural Network),包括长短期记忆网络 LSTM (Long Short-Term Memory) 和门控循环单元 GRU (Gated Recurrent Unit) 等^[11],在序列数据相关应用领域中呈现出了较好效果,例如机器翻译^[12]和语音识别^[13]等。循环神经网络能够捕获长时间的依赖性以及观测可变长度数据,使其在解决时序数据问题时可发挥较好能力^[14]。

本文利用循环神经网络特性,以传统门控循环单元为基础,参照文献[2]中的门控循环单元衰减 GRUD (Gated Recurrent Unit Decay),提出了一种新模型——MR-GRU 模型,用以解决低轨卫星缺失时序数据的模式识别问题。MR-GRU 模型在文献[2]的衰变量基础上,对时序数据时间依赖特性进行深度分析,更新了衰变量表示形式,增加了数据时间间隔表示方法。因此,MR-GRU 模型包含两个新变量:掩蔽项 (M) 和衰减项 (R)。掩蔽项用来衡量数据缺失,作用于每个时刻的输入数据上。衰减项衡量输入时序数据的时间延迟,作用于各个时刻的输入和隐层单元输出部分。在低轨卫星时序数据上的实验表明,MR-GRU 模型优于现有模型,可有效解决低轨卫星时间序列数据有缺失情况下的模式识别问题。

1 现有处理缺失数据的方法

目前,针对有缺失时序数据,一般处理策略都是想办法补全缺失值。也就是用各种数据评估方法来处理时序数据中的缺失值。例如:

① Mean^[2]和 Forward^[2]是最基础的估算缺失值方法;

② KNN (K-Nearest Neighbors) ^[15]使用 K 邻近算法寻找相似样本数据并依据相似样本的平均权重去估算缺失值;

③ MICE (Multivariate Imputation Chained Equations) ^[16]是使用最为广泛的一种估算方法,使用链式方程为不同类型变量创建多重输入;

④ MF (Matrix Factorization) ^[9]将一个有缺失的矩阵分解为两个低维矩阵,并在其上使用矩阵因子分解方法来填充缺失值;

⑤ PCA (Principal Component Analysis) ^[17]使用主成分分析模型来估算缺失值。

MICE、MF 和 PCA 方法将每个时序数据都看作各自独立的一部分,分别对其进行评估,可用于有不同长度的时序数据。而 KNN 方法将每一个时间步看作一个输入样本,将每个时序数据看作是一个数据点,因此不适用于有不同长度的时序数据。

2 基于传统循环神经网络的方法

循环神经网络的固有特性使其在解决有关时间序列数据问题时有可靠的理论基础。一般情况下循环神经网络模型都适用于处理时间序列数据。LSTM 和 GRU 是两种典型的循环神经网络模型。

传统的 GRU 模型结构较为简单,包括重置门和更新门两部分。使用 GRU 模型实现有缺失时序数据模式识别任务时,每一个时间步都采用传统 GRU 模块,然后在最后一个时间步的隐层单元上用一个分类层完成模式分类^[18]。

LSTM 模型结构^[19]较 GRU 模型略微复杂,包括忘记门、更新门和输出门三个模块。LSTM 可以学习具有某种长期依赖关系的模式^[20]。在 LSTM 模型基础上加入注意力机制可以进一步提升性能^[21]。使

用基于注意力机制和 LSTM 的神经网络能够快速捕捉到全局联系, 不但能获取长期依赖关系, 还能关注元素间的联系^[22]。

然而, 基于注意力机制的 LSTM 模型在处理时序数据时也有不足之处, 由于注意力机制对时序数据进行处理时不考虑输入数据的元素顺序, 使其在解决相关任务时存在一定的局限性^[23]。因此, 本文提出了 MR-GRU 模型, 严格依据时序数据的时间依赖顺序, 在传统 GRU 模型上进行改进。相较于文献[2]中的 GRUD 方法, 本文的 MR-GRU 模型对衰减项概念进行了改进完善, 额外考虑了在时序数据输入时可能伴随的时间依赖特性。MR-GRU 模型克服了传统模型存在的一些问题, 提升了对缺失时序数据的处理能力。

3 MR-GRU 模型

传统 GRU 模型处理时序数据时, 没有考虑数据缺失情况。而本文提出的 MR-GRU 模型则引入了掩蔽项变量用以表示数据缺失信息。另外, 为了进一步改善对缺失时序数据的处理效果, 使其能够直接表示时序数据的时间依赖特性, MR-GRU 模型在传统门式循环单元基础上, 参照文献[2]中的 GRUD 方法, 引入了衰减项变量。

MR-GRU 模型与 GRUD 模型的区别在于重新定义了衰减项, 对输入和隐层单元输出分别进行不同方式衰减, 然后将其与掩蔽项一起应用于 GRU 模型中。MR-GRU 模型结构如图 1 所示, 图中 m_t 表示掩蔽项, r_{h_t} 和 r_{x_t} 表示衰减项。

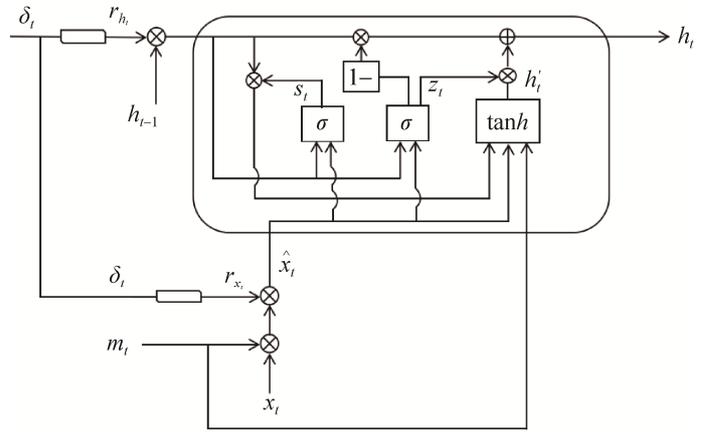


图 1 MR-GRU 模型结构图

Fig. 1 The structure of MR-GRU model

MR-GRU 模型中的衰减项包括 r_{h_t} 和 r_{x_t}

两部分。 r_{h_t} 表示 t 时刻隐层单元输出对应的衰减项。 r_{x_t} 表示 t 时刻输入单元对应的衰减项。 r_{h_t} 作用于隐层单元部分, 衡量相临两个数据之间的时间间隔。 r_{x_t} 则作用于输入单元部分, 衡量数据输入方向上的衰减。 r_{h_t} 计算公式如式 (1) 所示。

$$r_{h_t} = W_{hr} \times \exp\{-\max(0, W_r \delta_t + b_r)\} \quad (1)$$

其中, W_{hr} 表示待训练的隐层单元输出衰减项权值矩阵, W_r 表示待训练的隐层单元输出权值矩阵, b_r 表示待训练的隐层单元输出衰减项偏置, δ_t 表示相临两个隐层单元的时间间隔, 定义如下:

$$\delta_t = \begin{cases} 0, & t = 0 \\ s_t - s_{t-1}, & \text{other} \end{cases} \quad (2)$$

其中, s_t 表示 t 时刻的具体时间值, s_{t-1} 表示前一时刻的具体时间值。 r_{x_t} 计算公式如式 (3) 所示。

$$r_{x_t} = W_x r_{h_t} + b_x \quad (3)$$

其中, W_x 表示待训练的输入衰减项权值矩阵, b_x 表示待训练的输入衰减项偏置。由于考虑了数据输入方向上的衰减和数据缺失, 所以隐层单元每个时刻的输入由 x_t 变成了 \hat{x}_t , 如式 (4) 所示。

$$\hat{x}_t = r_{x_t} \otimes (m_t \otimes x_t) \quad (4)$$

其中, \otimes 表示向量对应元素相乘。 m_t 是掩蔽项, 表示时序数据中的数据缺失信息, 其定义为:

$$m_t = \begin{cases} 1, & \text{observed} \\ 0, & \text{other} \end{cases} \quad (5)$$

即, 如果 t 时刻没有数据缺失, 则 m_t 值为 1; 否则 m_t 值为 0, 表示该时刻缺失了数据。

由于 t 时刻的原始输入 x_t 在 sigmoid 激活前就经过 m_t 进行了掩蔽处理, 因此不同于 GRUD 模型,

MR-GRU 模型在 sigmoid 激活函数部分不再考虑 m_t 的作用, 而只在 tanh 激活函数计算输出时再考虑 m_t 的影响。

MR-GRU 模型中重置门函数 s_t 和更新门函数 z_t 分别如式 (6) 和式 (7) 所示:

$$s_t = \sigma(W_s \hat{x}_t + U_s \hat{h}_{t-1} + b_s) \quad (6)$$

$$z_t = \sigma(W_z \hat{x}_t + U_z \hat{h}_{t-1} + b_z) \quad (7)$$

其中, σ 表示 sigmoid 激活函数。 \hat{x}_t 表示 t 时刻原始输入 x_t 经过衰减项和掩蔽项处理过之后的输入, 如式 (4) 所示。 \hat{h}_{t-1} 表示隐层单元在上一时刻的输出 h_{t-1} 经过衰减项处理之后的值, 即:

$$\hat{h}_{t-1} = r_{h_t} \otimes h_{t-1} \quad (8)$$

W_s 、 U_s 、 b_s 分别表示重置门待训练的输入层权值矩阵、隐层单元权值矩阵和偏置。 W_z 、 U_z 、 b_z 则分别表示更新门待训练的输入层权值矩阵、隐层单元权值矩阵和偏置。

最终 t 时刻的隐层单元输出 h_t 由 t 时刻内部状态 h'_t 和上一时刻经过衰减项处理后的输出 \hat{h}_{t-1} 经过加权而成, 即:

$$h_t = (1 - z_t) \otimes \hat{h}_{t-1} + z_t \otimes h'_t \quad (9)$$

其中, z_t 是 t 时刻更新门输出。 t 时刻内部状态 h'_t 由式 (10) 计算:

$$h'_t = \tanh(W \hat{x}_t + U(s_t \otimes \hat{h}_{t-1}) + V m_t + b) \quad (10)$$

其中, W 、 U 、 V 分别表示待训练的权重矩阵, b 表示待训练的偏置。

与 GRUD 模型^[2]相比, MR-GRU 模型的计算过程有较大改变。在定义 δ_t 时, GRUD 模型不考虑缺失, 默认将缺失部分略过; 而 MR-GRU 模型把缺失数据用 0 值填充, 并将其视作正常数据。MR-GRU 模型根据输出衰减项 r_{h_t} 重新定义了输入衰减项 r_x , 然后由其更新 t 时刻的输入为 \hat{x}_t , 如式 (4) 所示。另外, MR-GRU 模型的重置门和更新门并不考虑掩蔽项 m_t , 只在计算 t 时刻内部状态时才综合考虑了所有影响因素, 包括 m_t 、 r_{x_t} 、 r_{h_t} 等。

在完整的模式识别过程中, MR-GRU 模型每个时刻 i 都输入 $\{x_i, m_i, \delta_i\}$ 3 个值和前一时刻的输出 h_{i-1} , 并产生 i 时刻的输出 h_i 。这 3 个值分别是 i 时刻的原始输入数据、 i 时刻原始输入数据的掩蔽项、 i 时刻与前一时刻间的时间间隔。然后 h_i 与下一时刻的

$\{x_{i+1}, m_{i+1}, \delta_{i+1}\}$ 再被输入到 MR-GRU 模型中产生下一时刻的输出 h_{i+1} 。如此重复, 直到观察窗口最后一个时刻数据经过 MR-GRU 产生输出之后送入一个全连接层, 最后利用 softmax 实现模式分类。整个识别过程如图 2 所示。

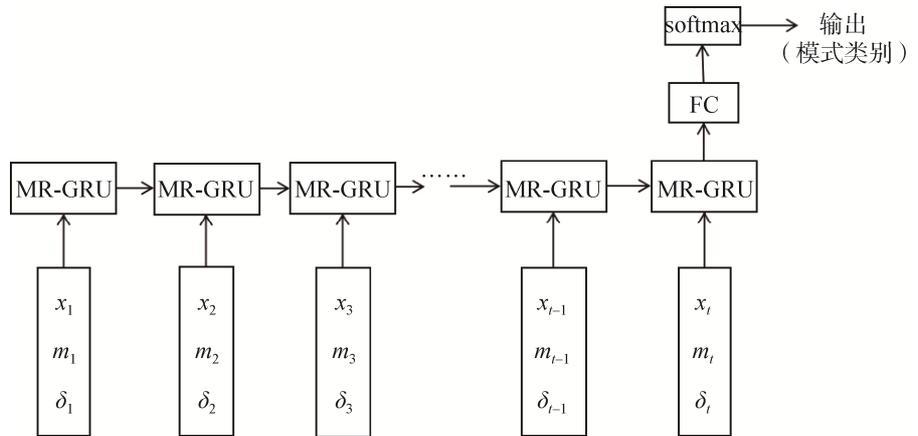


图 2 MR-GRU 模型的模式识别过程

Fig. 2 The pattern recognition process of MR-GRU model

4 数据处理及模型应用

4.1 数据预处理

在某卫星的五个遥测参数上对比测试本文所述模型, 包括: 卫星的负载功率、模拟式太阳敏感器 A、模拟式太阳敏感器 B、Y 分流调节器分流电流以及 Y 蓄电池温度, 以此来分别验证针对缺失卫星时序数据的模式识别问题。图 3 所示为 5 个参数在数据完整时的曲线图。表 1 列出了 5 组数据的预处理情况。

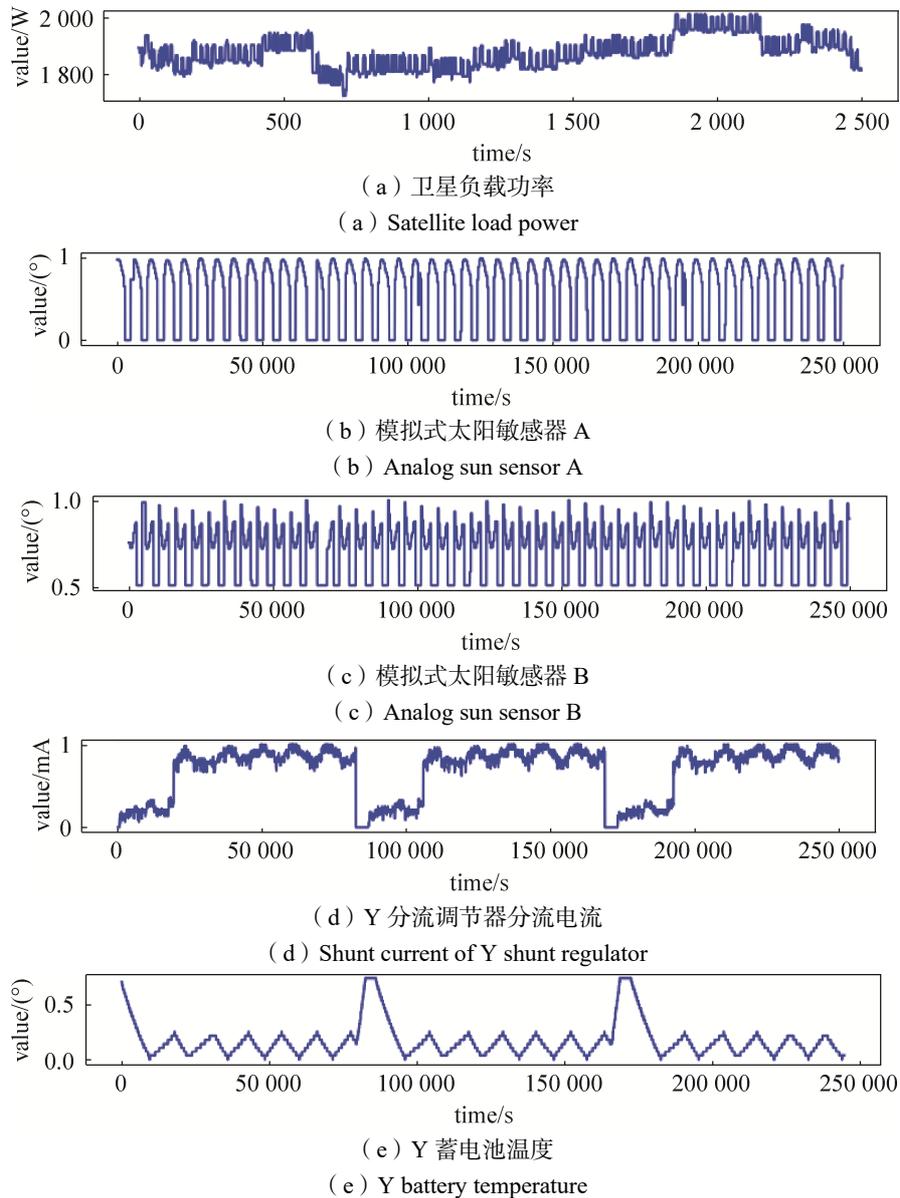


图 3 数据无缺失的 5 组遥测参数数据曲线图

Fig. 3 Five groups of telemetry parameter data curves without missing data

表 1 某卫星的五个遥测参数数据预处理情况

Table 1 Data preconditioning of five telemetry parameters of a satellite

数据名称	采样窗口大小 (s)	模式类别个数
负载功率	150	4
模拟式太阳传感器 A	5 000	3
模拟式太阳传感器 B	5 000	2
Y 分流调节器分流电流	1 000	2
Y 蓄电池温度	12 000	2

上述 5 组数据中, 卫星负载功率是描述该卫星运行期间功率的参数; 模拟式太阳传感器 A 及模拟式太阳传感器 B 是两个不同的传感器参数, 用于表示该卫星太阳传感器部件的工作状态; Y 分流调节器分流电流是描述该卫星分流调节器工作期间电流值的参数; Y 蓄电池温度是描述该卫星在运行期间蓄电池本身温度的参数。

根据表 1 中采样窗口大小, 分别将每组数据进行周期性划分并附上对应标签。以负载功率参数为例, 其他 4 组数据类似处理。首先, 按采样窗口 (150 s) 对初始数据 (图 3 (a)) 进行划分; 再对划分后每一段数据赋予相应模式类别标签值; 最后, 得到有 4 种不同模式类别的负载功率数据, 每个数据块长度为 150 s, 也就是说, 按采样窗口划分后的每一段数据就对应一个时序数据模式类别, 利用 MR-GRU 模型对该数据进行模式识别, 判定每段数据的模式类别。

4.2 缺失值处理策略

现有处理方法普遍都是使用一系列评估计算方法计算出一些值来代替缺失值, 这使得缺失值会有一定程度的不确定性, 并且很难保证评估值与实际值一样。为了避免这种问题, 本文不使用任意一种基线估算方法, 直接用 0 表示缺失值, 而正常观测到的值保持不变。这种方法实际上就是让正常数据和缺失数据融合在一起, 送入模型进行学习, 让模型自身去适应在数据有缺失情况下的分类学习, 这样才能最大限度地接近数据真实情况。

卫星的数据缺失现象通常是连续的, 即数据会表现出有一定随机长度的连续 0 值, 如图 4 (a) 所示。但是为了进一步讨论数据缺失影响, 本文还另外考虑了数据中有随机缺失的情况, 即数据中 0 值是随机出现, 并不是连续出现, 如图 4 (b) 所示。此外, 为与 0 值代替缺失值方法作比较, 本文还对比测试了以常见的 Mean 方法估算数据代替缺失值, 如图 4 (c) 和图 4 (d) 所示。

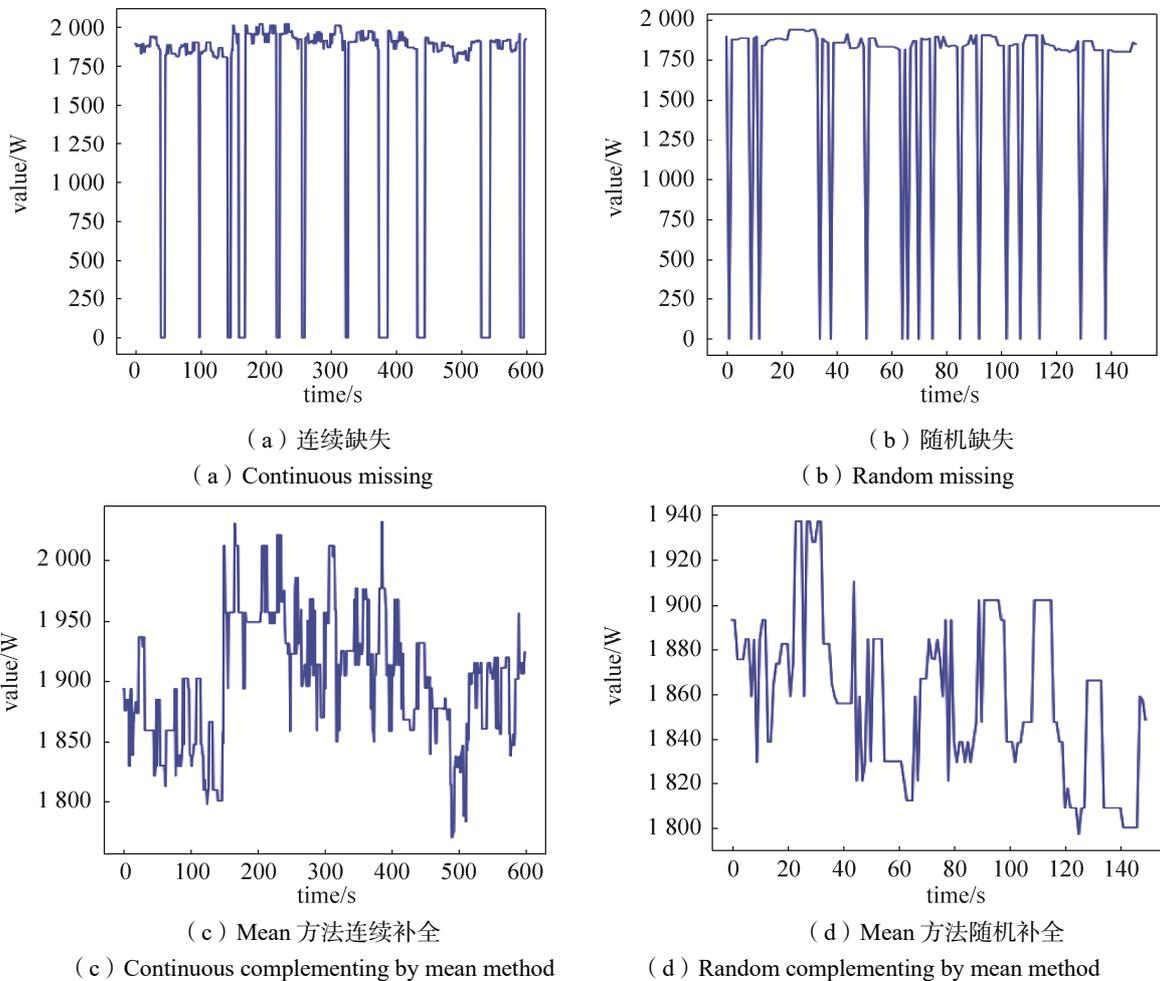


图 4 有缺失值的卫星负载功率数据曲线图
Fig. 4 Satellite load power data graph with missing values

4.3 结果对比

本文在上述有缺失值的卫星时序数据上对比了 MR-GRU 模型、传统 GRU 模型^[18]、基于注意力机制

的 LSTM 模型^[11]和 GRUD 模型^[2]。4 种模型在对比实验中设置了相同的训练轮数、学习率以及批量值 (batch_size), 都使用 Adam 优化器, 并用 MSE (Mean Square Error) 计算损失。

表 2 对比显示了 4 种不同模型在数据缺失率为 10%情况下的模式识别准确率结果。可以看到, 本文所提出的 MR-GRU 模型结果优于其他 3 种模型。图 5 则对比显示了在不同数据缺失概率情况下, 四种模型在测试数据集上模式识别结果的准确率曲线。从图中可以看出, MR-GRU 模型的识别准确率在不同测试数据上都能够一直保持最高, 并且十分稳定。而其他 3 种模型的识别率则有较大波动。MR-GRU 模型的准确率曲线明显优于 GRUD 模型。这说明本文针对衰减项所做的改进是成功的, 有效提升了模型对有缺失数据的识别准确率。

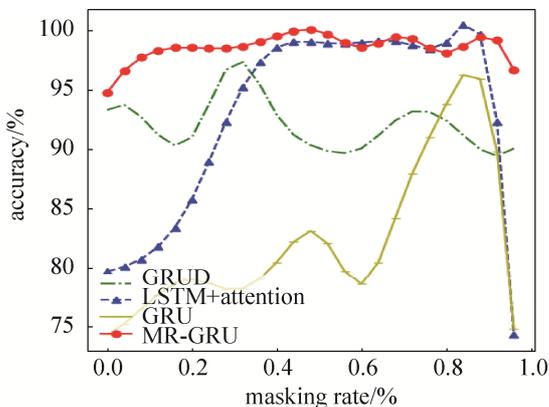
由实验结果可知, 相较于 GRU 模型、基于注意力机制的 LSTM 模型和 GRUD 模型, 无论是针对随机缺失还是连续缺失, MR-GRU 模型的准确率最高, 其次是 GRUD 模型, 再次是基于注意力机制的 LSTM 模型, 传统 GRU 模型效果最差。总体来看, 传统 GRU 模型与基于注意力机制的 LSTM 模型相差不多。使用这四种模型方法处理连续缺失的效果与处理随机缺失的效果差不多。此外, 使用 0 值表示缺失数据的效果较优于使用 Mean 估算值表示缺失数据。

表 2 4 种不同模型在数据缺失率为 10%时的模式识别准确率

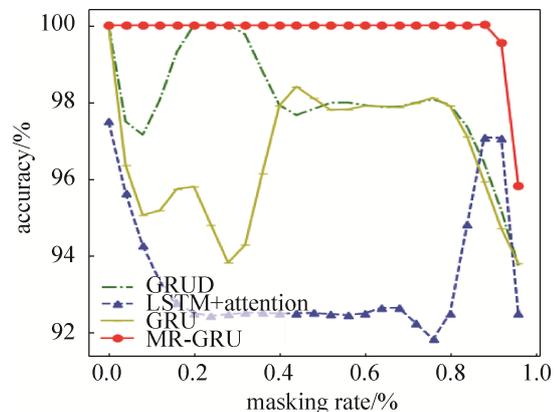
Table 2 The pattern recognition accuracy of four different models when the data missing rate is 10%

数据	缺失值 处理方式	GRU	LSTM + Attention	GRUD	MR-GRU	GRU	LSTM + Attention	GRUD	MR-GRU
		随机缺失				连续缺失			
负载功率	0 值	75.7%	81.22%	91.43%	98.11%	75.6%	80%	93.33%	95.28%
	Mean 估算	75.2%	76.11%	88.1%	95.75%	75.2%	76%	90%	94.81%
模拟式太阳 敏感器 A	0 值	95%	93.75%	97.5%	100%	95%	93.75%	97.5%	97.92%
	Mean 估算	95%	95.83%	95%	97.92%	92.5%	93.34%	92.5%	95.83%
模拟式太阳 敏感器 B	0 值	97.5%	97.92%	100%	100%	97.5%	97.92%	97.5%	97.92%
	Mean 估算	97.5%	97.92%	97.5%	97.92%	97.5%	96.67%	97.5%	97.92%
Y 分流调节器 分流电流	0 值	99.5%	100%	100%	100%	99.5%	100%	100%	100%
	Mean 估算	99.5%	100%	99.52%	100%	99.5%	100%	99.52%	100%
Y 蓄电池温度	0 值	94.4%	100%	100%	100%	94.4%	94.74%	100%	100%
	Mean 估算	94.4%	94.74%	100%	100%	94.4%	94.74%	100%	100%

在实际应用中卫星实时返回的各个参数数据中经常有缺失现象。这时若直接对这些参数数据进行模式识别, 利用现有方法不能达到满意效果。而本文提出的 MR-GRU 模型则可以对这种有缺失值的时间序列数据直接进行模式识别处理。由表 2 和图 5 结果可知, MR-GRU 模型比现有循环神经网络模型更稳定, 能更好地处理有缺失的时序数据, 并有效提升了模式识别效果, 具有更高可用性。



(a) 卫星负载功率数据
(a) Satellite load power data



(b) 模拟式太阳敏感器 A 数据
(b) Analog sun sensor A data

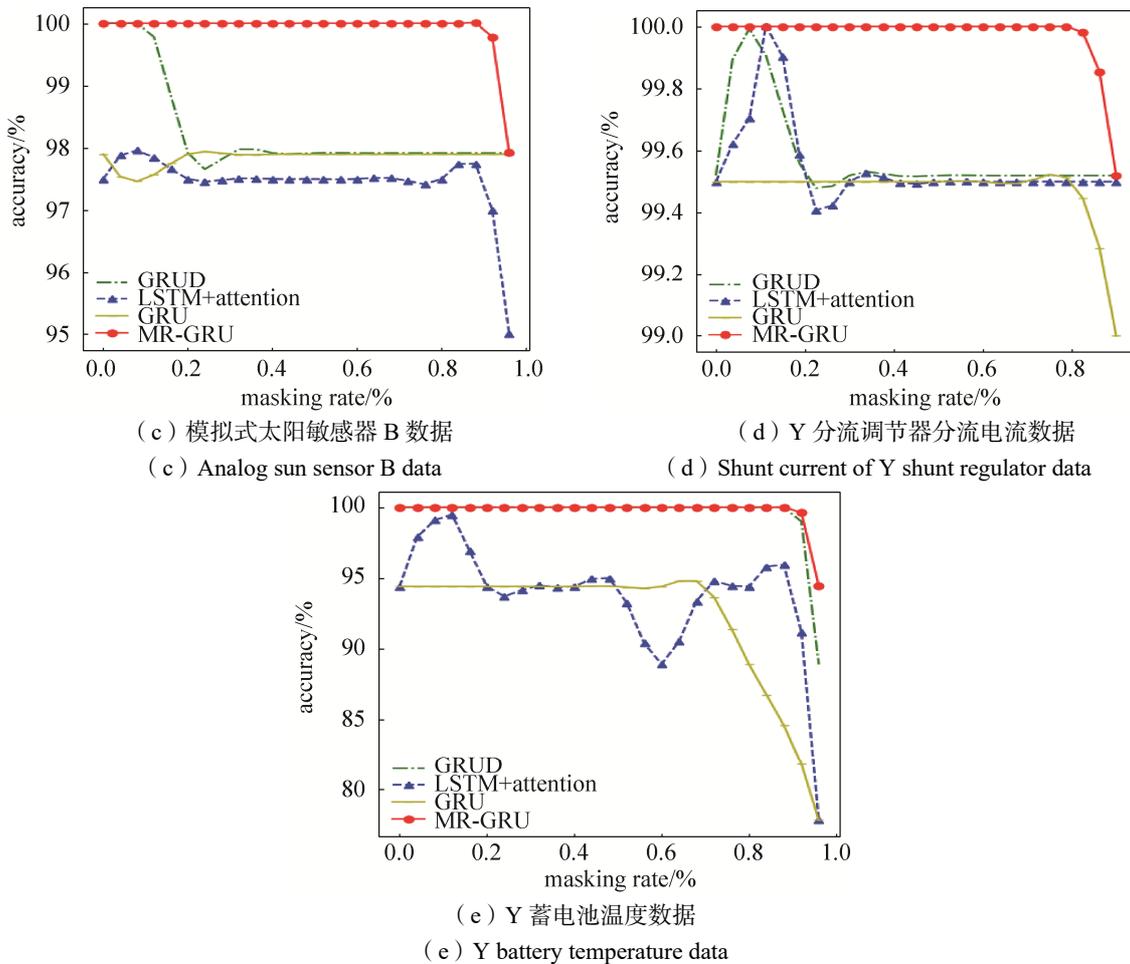


图 5 不同缺失概率情况下四种模型的准确率曲线
Fig. 5 Accuracy curves of four models on different missing probability

5 结束语

本文针对卫星时间序列数据有随机缺失值情况下的模式识别问题，提出了 MR-GRU 模型。该模型改进了传统门控循环单元 (GRU) 模型结构，增加了掩蔽项和衰减项。MR-GRU 模型不用补全缺失数据，而是直接在有缺失的时序数据上进行模式识别。本文在卫星时序数据上用不同的数据缺失概率对比了传统 GRU 模型、基于 LSTM 的注意力机制模型、GRUD 模型和 MR-GRU 模型的模式识别效果。结果表明相对于传统模型，MR-GRU 模型的识别准确率最高，泛化能力最好。

参考文献

- [1] SCHAFFER J L, GRAHAM J W. Missing data: our view of the state of the art[J]. Psychological Methods, 2007, 7(2): 147–177.
- [2] CHE Z, PURUSHOTHAM S, CHO K, et al. Recurrent neural networks for multivariate time series with missing values[J]. Scientific Reports, 2018, 8(1): 6085.
- [3] KREINDLER D M, LUMSDEN C J. The effects of the irregular sample and missing data in time series analysis[J]. Nonlinear Dynamics Psychology and Life Sciences, 2012, 10(2): 187–214.
- [4] 赵忠国, 张峰, 郑江华. 多元自适应回归样条法的滑坡敏感性评价[J]. 武汉大学学报(信息科学版), 2021, 46(3): 442–450. ZHAO Zhongguo, ZHANG Feng, ZHENG Jianghua. Landslide sensitivity evaluation by multiple adaptive regression spline method[J]. Journal of Wuhan University (Information Science Edition), 2021, 46(3): 442–450.
- [5] MONDAL D, PERCIVAL D B. Wavelet variance analysis for defective time series[D]. Washington: Annals of the Institute of Statistical Mathematics, 2010.
- [6] REHLFELD K, MARWAN N, HEITZIG J, et al. Comparison of correlation analysis techniques for irregularly sampled

- time series[J]. *Nonlinear Processes Geophysics*, 2011, 18(3): 389–404.
- [7] 杜永兴, 仝宗俊, 秦岭, 等. 基于改进 BP 算法的电磁涡旋成像方法[J]. *雷达科学与技术*, 2020, 18(5): 539–545.
DU Yongxing, TONG Zongjun, QIN Ling, et al. Electromagnetic vortex imaging method based on improved BP algorithm[J]. *Radar Science and Technology*, 2020, 18(5): 539–545.
- [8] MAZUMDER R, HASTIE T, TIBSHIRANI R. Spectral regularization algorithms for learning large incomplete matrices[J]. *The Journal of Machine Learning Research*, 2010, 11(11): 2287–2322.
- [9] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for references systems[J]. *Computer*, 2009, 42(8): 30–37.
- [10] WELLS B J, CHAGIN K M, NOWACKI A S, et al. Strategies for handling missing data in electronic health record derived data[J]. *EGEMS*, 2013, 1(3): 1035.
- [11] 吉长东, 朱锦帅, 黎虎, 等. BDS 卫星钟差短期预报的 LSTM 算法[J]. *导航定位学报*, 2021, 9(1): 44–52.
JI Changdong, ZHU Jinshuai, LI Hu, et al. A LSTM algorithm for BDS satellite clock difference short term prediction[J]. *Journal of Navigation and Positioning*, 2021, 9(1): 44–52.
- [12] ZHANG B, XIONG D Y, XIE J, et al. Neural machine translation with GRU-gated attention model[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(11): 4688–4698.
- [13] HINTON G, DENG Li, YU Dong, et al. Deep neural networks for acoustic modeling in speech recognition[J]. *IEEE Signal Processing Magazine*, 2012, 29(6): 82–97.
- [14] CHO K, MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//*The Conference Empirical Methods in Natural Language Processing*, Doha, Qatar, October 25, 2014.
- [15] 黄铭, 孙林夫, 任春华, 等. 改进 KNN 的时间序列分析方法[J]. *计算机科学*. 2021, 48(4): 22–34.
HUANG Ming, SUN Linfu, REN Chunhua, et al. Improved time series analysis method of KNN[J]. *Computer Science*. 2021, 48(4): 22–34.
- [16] AZUR M J, STUART E A, FRANGAKIS C, et al. Multiple imputation by chained equations: what is it and how does it work[J]. *International Journal of Methods in Psychiatric Research*, 2011, 20(1): 40–49.
- [17] 邓念东, 丁一, 邢聪聪, 等. 基于 PCA 和 BP 神经网络的风积砂充填体强度预测[J]. *矿业研究与开发*, 2021, 41(2): 109–113.
DENG Niandong, DING Yi, XING Congcong, et al. Strength prediction of aeolian sand filling body based on PCA and BP neural network[J]. *Mining Research and Development*, 2021, 41(2): 109–113.
- [18] MA Yan, LI Hongguang. A GRU network-based approach for steam drum water level predictions[J]. *Journal of Chemical Engineering of JAPAN*, 2020, 53(5): 11.
- [19] GAO Shuai, HUANG Yuefei, ZHANG Shuo, et al. Short-term run off prediction with GRU and LSTM networks without requiring time step optimization during sample generation[J]. *Journal of Hydrology*, 2020, 589: 125188.
- [20] YAN Haoran, QIN Yi, XIANG Sheng, et al. Long-term gear life prediction based on ordered neurons LSTM neural networks[J]. *Measurement*, 2020, 165: 108–205.
- [21] BAHDANAU D. Neural machine translation by jointly learning to align and translate[C]// *International Conference on Learning Representations*, San Diego, CA, USA, May 7-9, 2015.
- [22] CHOROWSKI J K, BAHDANAU D, SERDYUK D, et al. Attention-based models for speech recognition[C]//*In Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems*, Montreal, Quebec, Canada, December 7-12, 2015.
- [23] CHEN Long, ZHANG Hanwang, XIAO Jun, et al. Spatial and channel-wise attention in convolution NN for image captioning[J]. *Computer Vision and Pattern Recognition*, 2017, 6298–6306.

[作者简介]

- 郭思尧 1997 年生, 硕士研究生, 主要研究方向为时序数据在神经网络中的应用。
颜 博 1983 年生, 高级工程师, 主要研究方向为航天地面控制技术、智能优化技术。
翟 磊 1988 年生, 高级工程师, 主要研究方向为航天地面控制技术、卫星健康管理技术。
鲍军鹏 1974 年生, 博士, 博士生导师, 主要研究方向为机器学习、数据挖掘、人工智能。
宋 宴 1983 年生, 高级工程师, 主要研究方向为航天地面控制技术、智能优化技术。
张 超 1986 年生, 高级工程师, 主要研究方向为航天地面控制技术、智能优化技术。
陈亚军 1985 年生, 高级工程师, 主要研究方向为航天地面控制技术、智能优化技术。

(本文编辑: 傅 杰)