

面向智能运维的轨道交通转辙机模拟数据 生成器设计与验证^{*}

邹劲柏¹ 魏诗燕¹ 刘江² 沙泉¹ 吴杰³ 季国一¹

(1. 上海应用技术大学轨道交通学院, 201400, 上海; 2. 北京交通大学电子信息工程学院, 100044, 北京;
3. 上海地铁维护保障有限公司通号分公司, 200235, 上海)

摘要 [目的] 由于轨道交通各类设备的故障数据难以轻易获取, 导致在开展故障诊断与预测等机器智能算法研究时缺乏充足的数据支持。为了满足轨道交通智能运维对大量训练数据的迫切需求, 有必要设计轨道交通转辙机模拟数据生成器并对其进行验证。[方法] 对 S700K 型转辙机正常动作与缓变性故障的功率曲线特征进行了分析, 并探讨了故障发生原因。通过对比两种模拟数据生成方法, 基于 Borderline-SMOTE 算法设计出转辙机模拟数据生成器, 搭建转辙机模拟数据生成器平台, 利用 LSTM(长短期记忆) 预测模型学习功率数据的时间序列特征, 对生成的缓变性故障功率数据的峰值因子、标准差和方差等 3 个特征进行试验。[结果及结论] 通过转辙机模拟数据生成器生成的功率数据训练出的 LSTM 预测模型, 可以预测出 S700K 型转辙机的功率变化趋势。通过对比 LSTM 预测模型与周期性复制法计算得到的峰值因子、标准差、方差的均方根误差分别为 0.335 5、0.023 9 和 0.024 1, 误差较小, 证明转辙机模拟数据生成器的真实性及可行性。

关键词 轨道交通; 转辙机; 模拟数据生成器; 智能运维
中图分类号 U213.6⁺1

DOI:10.16037/j.1007-869x.2025.01.034

Design and Verification of Simulation Data Generator for Rail Transit Switch Machine O- riented to Intelligent Operation and Maintenance

ZOU Jinbai¹, WEI Shiyan¹, LIU Jiang², SHA Quan¹, WU Jie³, JI Guoyi¹

(1. School of Railway Transportation, Shanghai Institute of Technology, 201400, Shanghai, China; 2. School of Electronic and Information Engineering, Beijing Jiaotong University, 100044, Beijing, China; 3. Telecom & Signaling Branch, Shanghai Metro Maintenance Support Co., Ltd., 200235, Shanghai, China)

Abstract [Objective] Difficulty in easily obtaining fault data of various rail transit equipment leads to insufficient data to support the research on machine intelligence algorithms such as fault diagnosis and prediction. In order to meet the urgent need of intelligent rail transit operation and maintenance for a large amount of training data, it is necessary to design and verify the simulation data generator (hereinafter abbreviated as SD generator) of rail transit switch machine. [Method] The characteristics of S700K type switch machine power curves under normal operation and gradual fault conditions are analyzed, and causes of the faults are discussed. By comparing two simulation data generation methods, a switch machine SD generator is designed based on the Borderline-SMOTE algorithm. Through building a platform for SD generator, and using the time series features of learning the power data by LSTM (long short-term memory) prediction model, three characteristics of the generated gradual fault power data such as the crest factor, standard deviation, and variance are tested. [Result & Conclusion] The LSTM prediction model trained on the power data generated by SD generator can predict the power change trend of the S700K type switch machine. The root mean square errors of crest factor, standard deviation, and variance calculated by the LSTM prediction model and the periodic replication method are 0.335 5, 0.023 9, and 0.024 1 respectively. The relatively small errors prove the authenticity and feasibility of SD generator.

Key words rail transit; switch machine; simulation data generator; intelligent operation and maintenance

目前转辙机故障诊断和预测的方法很多^[1], 但都未能从数据层面解决故障数据缺乏的问题。由于转辙机发生故障的频率较低^[2], 历史故障数据稀缺, 在轨道交通智能运维中, 仅利用某一转辙机在

^{*} 上海市科技创新行动计划项目(21210750300); 上海市地方能力建设项目(20090503100); 北京市自然科学基金项目(L191014); 科技创新行动计划“一带一路”中老铁路工程国际联合实验室项目(21210750300); 地方能力建设“轨道交通智能运维关键技术研究”项目(20090503100)

某种条件下的故障数据很难进行故障诊断与预测的机器智能研究。解决故障数据稀缺难题的手段之一是进行模拟数据生成器的开发。模拟数据生成器能够模拟转辙机多种故障数据的变化情况,实现现场故障数据的仿真回放,从而为各种故障诊断或预测技术提供海量故障数据^[3]。本文研究的主要对象为转辙机,分析了转辙机功率曲线特征,介绍了模拟数据生成器的数据生成方法,构建针对转辙机的模拟仿真平台,通过对所生成的故障数据进行试验,验证了模拟数据生成器的真实性和可行性。

1 S700K 型转辙机动作功率曲线分析

S700K 型转辙机因其具有良好的工作性能在轨道交通中得到广泛运用。本文从 S700K 型转辙机功率-时间数据出发,对其动作功率曲线特征进行了分析。

1.1 S700K 型转辙机工作的主要参数

S700K 型转辙机在三相四线制的交流电下工作,单独分析转辙机的工作电流和电压,只能片面了解该转辙机的运行状态。

$$P = UI \cos \theta \quad (1)$$

式中:

P ——功率;

U ——电压;

I ——电流;

θ ——电压与电流之间的夹角。

$$P = \pi n v F / (30 \Omega \eta_t) \quad (2)$$

式中:

n ——电机转动圈数,单位圈;

v ——平移速度,单位 m/s;

Ω ——角速度,单位 rad/s;

η_t ——电动机的传动效率,单位 %;

t ——时间,单位 s;

F ——转换力,单位 N。

对电动转辙机微机监测测量的主要参数包括:转辙机在动作过程中的输出力,电机电流及功率随时间的变化情况。由式(1)和式(2)可知:转辙机的电压、电流及转换力的变化均可在功率数据上体现^[4],故本文对转辙机的工作功率数据进行研究。

1.2 转辙机正常动作与缓变性故障功率曲线分析

图 1 为 S700K 型转辙机正常工作时的功率-时间曲线。由图 1 可见:前 0.1 s 为转辙机需要克服阻力的启动阶段,表现为骤然上升且很快达到峰

值。随后,转辙机内部完成解锁进入转换阶段,功率急速降至 0.5 kW 左右并趋于平稳。此后进入锁闭阶段,由于三相交流电还存在两相小电流,因此在 5 s 左右出现一个小台阶,直至表示电路被切断,功率降至 0。

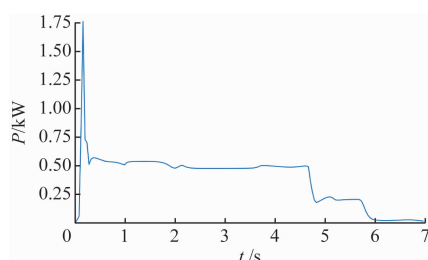


图 1 转辙机正常工作功率曲线

Fig. 1 Power curve of the switch machine under normal operation

图 2 为转辙机发生某种缓变性故障时的功率曲线。由图 2 可见:在转换阶段转辙机功率大幅波动,且整个转辙机的工作时间延长至 14 s。该类故障发生的原因如下:滑板床撒沙,滑板床缺油,滑板床固定部件松动。这三个原因都会增加滑板床和轨道之间的摩擦阻力,日积月累,最终会影响道岔正常转换。

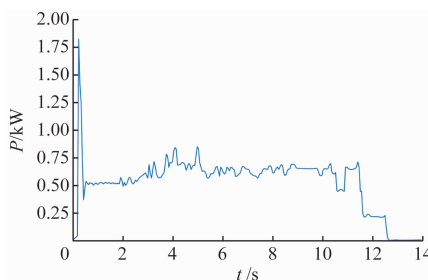


图 2 转辙机发生缓变性故障时的功率曲线

Fig. 2 Power curve of the switch machine with gradual fault

2 模拟数据生成方法

为了得到大量上述的故障数据,本文采用 SMOTE(合成少数类过采样技术)的改进算法 Borderline-Smote。Borderline-Smote 算法解决了 SMOTE 对所有少数样本一视同仁的问题^[5],该算法仅对靠近多数样本的少数样本进行数据增强。其主要改进方法为:计算出每个少数样本 x_i 的近邻样本 x_j 属于多数样本的数量 N ,对于符合 $j/2 < N < j$ 要求的少数样本 x_i 进行合成。

将正常数据样本作为边界样本,其中掺杂着少

数样本,应用 Borderline-Smote 算法,人为生成新的少数样本^[6-7]。图 3 为使用 Borderline-Smote 算法生成数据前后对比图。由图 3 可见:Borderline-Smote 算法生成的新数据包含了原本少数样本数据和多数样本数据的特征。

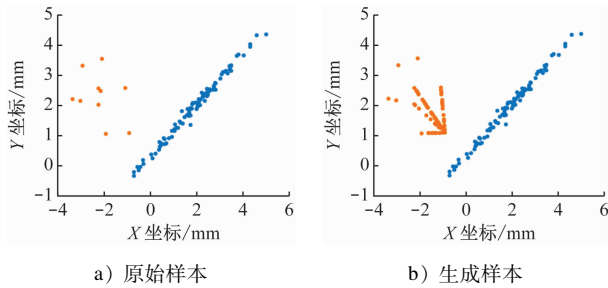


图 3 使用 Borderline-Smote 算法生成数据前后对比图

Fig. 3 Comparison charts before and after data generation using the Borderline-Smote algorithm

3 模拟数据生成器平台搭建及验证

3.1 基于 LSTM 的验证试验设计

LSTM(长短期记忆)模型是卷积神经网络模型的一种改进形式,可以学习长期的依赖信息,而转辙机功率时间数据具备时间序列特性,为本试验提供了理论支持^[8]。

为验证模拟数据生成器的真实性及可行性,对转辙机多特征缓变性故障进行了验证试验设计。试验设计框图见图 4。由图 4 可见:本试验设计包括 S700K 转辙机模拟数据生成器平台与 LSTM 预测模型训练两部分。

LSTM 预测模型训练的具体实施流程如下:将转辙机模拟数据生成器平台得到的 3 种特征值作为 1 个单元,分别输入到 LSTM 预测模型中训练,利用当前输入的转辙机特征值 g_t 和上一时刻门控状态值 h_{t-1} 一起训练得到三个状态,即图 4 中的 Z_f 、 Z_i 和 Z_o 。这三种门控状态均由 h_{t-1} 、 g_t 拼接成的向量与权重矩阵相乘之后,通过 sigmoid 激活函数,得到最终的输出值。

$$c_t = Z_f \otimes c_{t-1} + Z_i \quad (3)$$

$$h_t = Z_o \otimes \tanh c_t \quad (4)$$

$$y_t = \sigma(W h_t) \quad (5)$$

式中:

c_t ——当前时刻 t 的信息状态值;

σ ——sigmoid 函数,取值为 0~1;

W ——权重矩阵;

y_t ——本试验指定最终的输出。

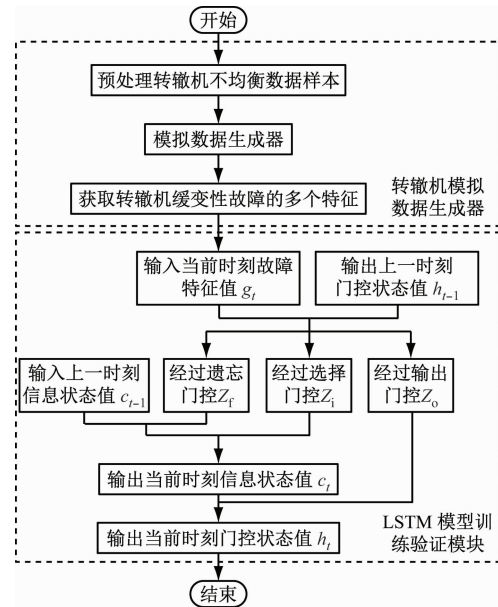


图 4 试验设计框图

Fig. 4 Block diagram of experimental design

训练好 3 个状态后,将上一时刻门控状态值 h_{t-1} 、当前时刻输入的转辙机特征值 g_t 、上一时刻信息状态值 c_{t-1} 作为当前 LSTM 预测模型的输入。从式(3)、式(4)和式(5)可见:通过 LSTM 预测模型中的参数可以计算得到该模型的输出。

3.2 试验结果与分析

本文试验的主要实现步骤是:先训练训练集,利用规定步长预测后一个数值,再对测试集进行测试。本试验选取由转辙机模拟数据生成器平台生成的 200 组故障数据作为基本数据,其中 160 组故障数据为训练集,40 组故障数据为测试集。本试验的输入为功率数据、峰值因子、标准差及方差,输出为相应的数值。转辙机微机监测系统的采样频率为 25 Hz,即每隔 40 ms 采集 1 个功率数据,因此针对 1 组转换时间为 14 s 的缓变性故障,可采集到 350 个功率数据。设置本试验的步长为 25,即每 25 个数据作为 1 个时刻的输入,第 26 个数据作为该时刻的标签值,训练 500 轮,试验的输出即每个时刻的标签值。

图 5 为转辙机缓变性故障功率训练结果。由图 5 可见:虚线、实线分别代表 LSTM 预测模型的输入和输出数据,虚线和实线大致贴合,说明该模型的训练效果贴合原始数据的大致变化趋势。图 6 为转辙机缓变性故障功率测试结果。由图 5 与图 6 可

见:由转辙机模拟数据生成器生成的功率数据可以很好地训练 LSTM 预测模型,并且能够预测出该功率的大致变化趋势。

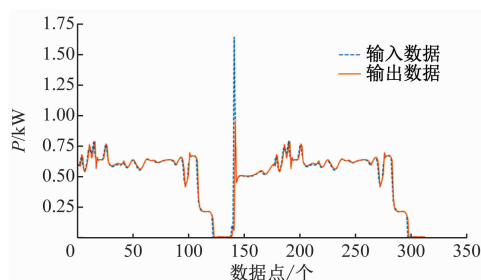
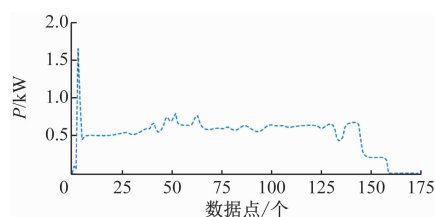
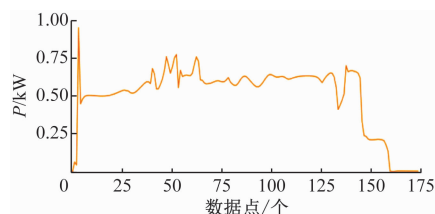


图 5 转辙机缓变性故障功率训练结果

Fig. 5 Training results of the switch machine power with gradual faults



a) 原始输入的功率数据



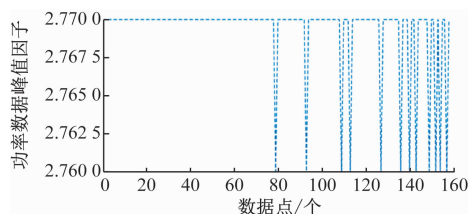
b) 预测输出的功率数据

图 6 转辙机缓变性故障功率测试结果

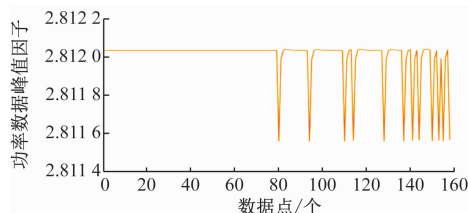
Fig. 6 Test results of the switch machine power with gradual faults

同样以上述试验步骤对转辙机功率数据的峰值因子进行训练。图 7 为转辙机功率数据的峰值因子特征测试结果。由图 7 可见:试验预测出了原始功率数据峰值因子特征的大体趋势,但存在着一定的误差。本试验主要是为了证明转辙机模拟数据生成器平台生成的数据可训练 LSTM 预测模型,因此该误差在允许范围内。图 8 为转辙机功率数据的标准差测试结果。图 9 为转辙机功率数据的方差测试结果。

从上述试验分析可知,由转辙机模拟数据生成器平台的生成数据训练出的 LSTM 预测模型,较适用于转辙机功率数据及其特征的故障时间序列预测,且能够预测出特征数据的大致波动趋势,证明



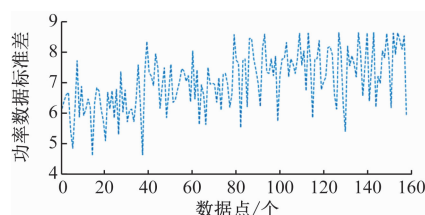
a) 原始输入的功率数据峰值因子



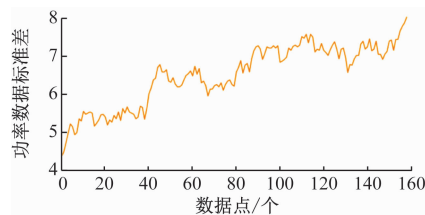
b) 预测输出的功率数据峰值因子

图 7 转辙机功率数据的峰值因子测试结果

Fig. 7 Test results of the crest factor of the switch machine power data



a) 原始输入的功率数据标准差



b) 预测输出的功率数据标准差

图 8 转辙机功率数据的标准差测试结果

Fig. 8 Test results of the standard deviation of the switch machine power data

了转辙机模拟数据生成器平台的生成数据能够训练出较完美的 LSTM 预测模型。

使用周期性复制法对 6 组转辙机缓变性故障功率数据样本扩增至 200 组,并进行上述试验。选择 RMSE(均方根误差)作为度量这两组试验测试结果的标准^[9]。在 LSTM 预测模型的测试中,应用周期性复制法和 Borderline-Smote 算法生成数据的特征值,计算得到峰值因子的 RMSE 为 0.335 5,标准差的 RMSE 为 0.023 9,方差的 RMSE 为 0.024 1,误差较小,证明了转辙机模拟数据生成器的真实性及可行性。

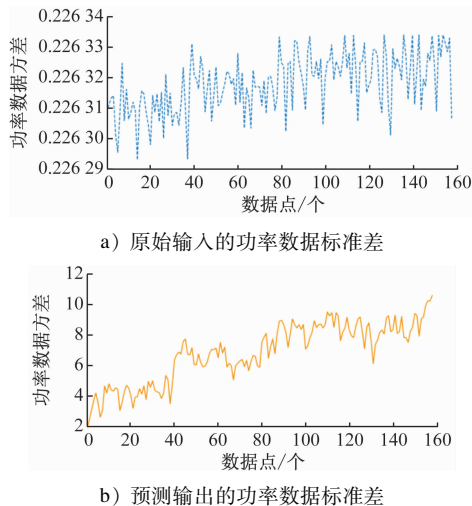


图9 转辙机功率数据的方差测试结果

Fig. 9 Test results of the variance of the switch machine power data

4 结语

本文设计了一种利用 Borderline-Smote 算法生成海量数据的模拟数据生成器,构建了转辙机模拟数据生成器平台,供地铁转辙机模拟相关数据,从而达到故障诊断与预测的目的。本文分析了 S700K 转辙机正常工作与发生缓变性故障功率数据曲线特征,并对模拟数据生成器中的相关方法进行了介绍,构建了转辙机模拟数据生成器平台,利用该平台生成的多特征缓变性故障数据训练 LSTM 预测模型,验证模拟数据生成器生成数据的真实性和可行性。结果表明,模拟数据生成器生成的数据可以很好地训练 LSTM 预测模型,但不能利用该模型对转辙机故障进行预测。后续将尽可能生成转辙机性能退化数据,建立退化性能指标,对转辙机退化性能进行预测。

参考文献

- [1] 高利民,许庆阳,李锋,等. 基于 SOM-BP 混合神经网络的道岔设备退化状态研究[J]. 中国铁道科学, 2020, 41(3): 50.
GAO Limin, XU Qingyang, LI Feng, et al. Research on degradation state of turnout equipment based on SOM-BP hybrid neural

- network[J]. China Railway Science, 2020, 41(3): 50.
- [2] LU K X, XIAO J. Analysis and treatment of circuit faults of S700K electric switch machine [C] // Proceedings of SPIE: Machine Intelligence and Pattern Recognition in Remote Sensing 2019. Wuhan: SPIE-The International Society for Optical Engineering, 2020.
- [3] LI Z, YIN Z, TANG T, et al. Fault diagnosis of railway point machines using the locally connected autoencoder[J]. Applied Sciences, 2019, 9(23): 5139.
- [4] 侯大山. 基于神经网络的道岔故障预测和诊断[D]. 北京: 北京交通大学, 2019.
HOU Dashan. Points fault prognosis and diagnosis based on neural network[D]. Beijing: Beijing Jiaotong University, 2019.
- [5] 刘新发,魏文军. 基于模糊聚类方法的 S700K 转辙机故障诊断[J]. 中南大学学报(自然科学版), 2019, 50(9): 2148.
LIU Xinfu, WEI Wenjun. Fault diagnosis of S700K switch machine based on fuzzy cluster method[J]. Journal of Central South University (Science and Technology), 2019, 50(9): 2148.
- [6] 赵楠,张小芳,张利军. 不平衡数据分类研究综述[J]. 计算机科学, 2018, 45(增刊1): 22.
ZHAO Nan, ZHANG Xiaofang, ZHANG Lijun. Overview of imbalanced data classification [J]. Computer Science, 2018, 45 (S1): 22.
- [7] 邓梦菲. S700K 转辙机故障诊断与预测[D]. 北京: 北京交通大学, 2021.
DENG Mengfei. Fault diagnosis and prediction of S700K switch machine[D]. Beijing: Beijing Jiaotong University, 2021.
- [8] 石洪波,陈雨文,陈鑫. SMOTE 过采样及其改进算法研究综述[J]. 智能系统学报, 2019, 14(6): 1073.
SHI Hongbo, CHEN Yuwen, CHEN Xin. Summary of research on SMOTE oversampling and its improved algorithms [J]. CAAI Transactions on Intelligent Systems, 2019, 14(6): 1073.
- [9] 李婉婉. 基于 LSTM 的提速道岔故障预测研究[D]. 兰州: 兰州交通大学, 2020.
LI Wanwan. Study on fault prediction of speed-up switch based on LSTM[D]. Lanzhou: Lanzhou Jiaotong University, 2020.

· 收稿日期:2023-09-11 修回日期:2023-11-02 出版日期:2025-01-10
Received:2023-09-11 Revised:2023-11-02 Published:2025-01-10
· 通信作者:邹劲柏,教授,Zoujb@sit.edu.cn
· ©《城市轨道交通研究》杂志社,开放获取 CC BY-NC-ND 协议
© Urban Mass Transit Magazine Press. This is an open access article under the CC BY-NC-ND license

欢迎投稿《城市轨道交通研究》

投稿网址:tougao. umt1998. com