Data specifications to final stage

1. Introduction

To promote the development of PHM in the rail transit industry and attract more scholars, engineers, and students to join studies on related techniques, the Committee of PHM-Beijing 2024 organized the PHM Data Challenge with the background of fault diagnosis for subway train transmission systems.

2. Dataset overview

Datasets used at the final stage are provided by the State Key Laboratory of Advanced Rail Autonomous Operation of Beijing Jiaotong University. **Based on the normal and 16 single typical fault classes in the preliminary training datasets, 2 component-level compound-fault classes and 3 system-level compound-fault classes are further given at the final stage.** The datasets are available for everyone. In addition to the competition, participants are encouraged to use the datasets to verify original fault diagnosis algorithms and submit papers to the proceedings of the PHM-Beijing2024. The following article can be referenced when you use the datasets:

- *A. Ding, Y. Qin, B. Wang, L. Guo, L. Jia, and X. Cheng, Evolvable graph neural network for system-level incremental fault diagnosis of train transmission systems, Mechanical Systems and Signal Processing, 210 (2024) 111175.*

3. Experimental platform

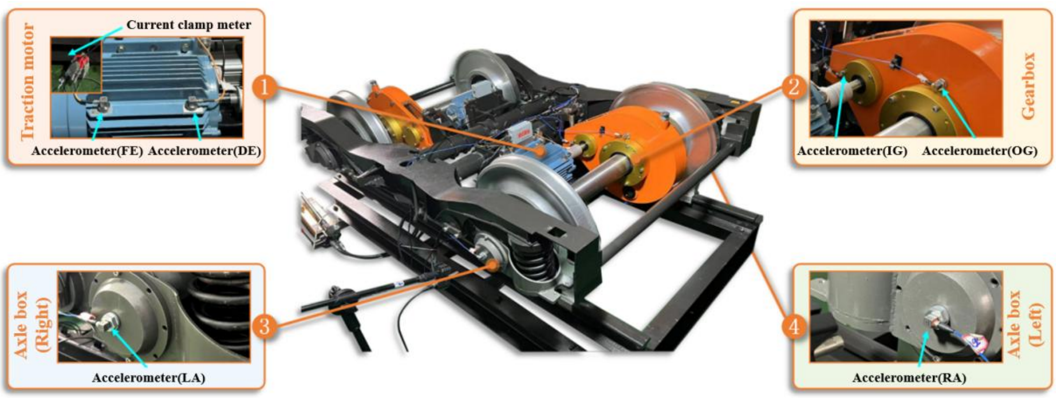

Fig.1 Experimental platform for fault simulation of subway train transmission systems.

The data is collected from the experimental platform for fault simulation of subway train bogies, as shown in Fig. 1. This platform is designed and adjusted according to the real bogies of subway trains, and the scale of the platform to the real bogie is 1:2. The single transmission chain includes a motor, a gearbox, and two axle boxes. The transmission chain is driven by a three-phase asynchronous AC motor. The load is applied via electro-hydraulic load equipment. The motor bearing type is SKF 6205-2RSH. The gears of the reduction gearbox are helical gears. The number of the driving gear teeth is 16, while the number of the driven gear teeth is 107. The axial bearing type of driving gear is HRB 32305. The axle box bearing type is HRB 352213.

4. Working conditions

The datasets consider **9 working conditions,** as listed in Table 1. Different motor speed is to simulate different train speed, while different lateral load is to simulate trains going straight or round corners. The positive value of load denotes that the loading direction points to the motor side. Otherwise, it denotes that the loading direction points to the gearbox side.

Table1 Working condition list

Working condition	Motor speed/lateral load	Working condition	Motor speed/lateral load
WC1	20Hz/0kN	WC6	40Hz/-10kN
WC2	20Hz/10kN	WC7	60Hz/0kN
WC3	20Hz/-10kN	WC8	60Hz/10kN
WC4	40Hz/0kN	WC9	60Hz/-10kN
WC5	40Hz/10kN		

5. Sampling setting

A total of 21 channel signals including vibration and current are collected with a sampling frequency of 64kHz in the experiments. The deployment of sensors is shown in Fig. 1. The information on sensor channels is given in Table 2.

Table 2 Summary of sensor channel information

Channel	Components	Deployment location	Signal type
CH1	Motor	Motor (drive end)	Tri-axial acceleration
CH2			
CH3			
CH4		Motor (fan end)	Tri-axial acceleration
CH5			
CH6			
CH7		Motor (cable)	Three-phase current
CH8			
CH9			
CH10	Gearbox	Gearbox (input axle)	Tri-axial acceleration
CH11			
CH12			
CH13		Gearbox (output axle)	Tri-axial acceleration
CH14			
CH15			
CH16	Axle box (left)	Axle box (end cover)	Tri-axial acceleration
CH17			

CH18			
CH19 CH20 CH21	Axle box (right)	Axle box (end cover)	Tri-axial acceleration

6. Fault types and Sample labels

Note that the labeling way at the final stage is different from that at the preliminary stage. Please read this paragraph carefully. The basic health states of the components **(motors, gearboxes, axle boxes)** in datasets are given in Table 3. Fig. 4 shows the photographs of fault simulations in the experiments. Examples are listed in Table 4, which presents the labels of classical single-fault classes, component-level compound-fault classes, and system-level compound-fault classes based on basic health states. **The labeling at the final stage should be subject to this way.**

Table3 Health state in the datasets

Component	Basic health state	Description
Motor	M0	Normal condition
	M1	Short circuit
	M2	Broken rotor bar
	M3	Bearing fault
	M4	Bowed axis
Gearbox	G0	Normal condition
	G1	Gear cracked tooth
	G2	Gear worn tooth
	G3	Gear missing tooth
	G4	Gear chipped tooth
	G5	Bearing inner race fault
	G6	Bearing outer race fault
	G7	Bearing rolling element fault
	G8	Bearing cage fault
Axle box (left)	LA0	Normal condition
	LA1	Bearing inner race fault
	LA2	Bearing outer race fault
	LA3	Bearing rolling element fault
	LA4	Bearing cage fault
Axle box (right)	RA0	Normal condition

Table4 Label labeling examples (Important)

Example		Description of the fault	Label
Typical single-fault classes	Example 1	Short circuit (Motor)	M1_G0_LA0_RA0
	Example 2	Gear cracked tooth (Gearbox)	M0_G1_LA0_RA0
	Example 3	Bearing inner race fault (Left axle box)	M0_G0_LA1_RA0
Component-level compound-fault classes	Example 1	Gear chipped tooth (Gearbox) + Bearing inner race fault (Gearbox)	M0_G4+G5_LA0_RA0
	Example 2	Bearing inner race fault (Left axle box) + Bearing outer race fault (Left axle box) + Bearing rolling element fault (Left axle box) + Bearing cage fault (Left axle box)	M0_G0_LA1+LA2+LA3+LA4_RA0
System-level compound-fault classes	Example 1	Short circuit (Motor) + Bearing inner race fault (Left axle box)	M1_G0_LA1_RA0
	Example 2	Gear missing tooth (Gearbox) + Bearing inner race fault (Left axle box)	M0_G3_LA1_RA0
	Example 3	Short circuit (Motor) + Bearing inner race fault (Motor) + Bearing inner race fault (Right axle box)	M1_G0_LA1_RA1

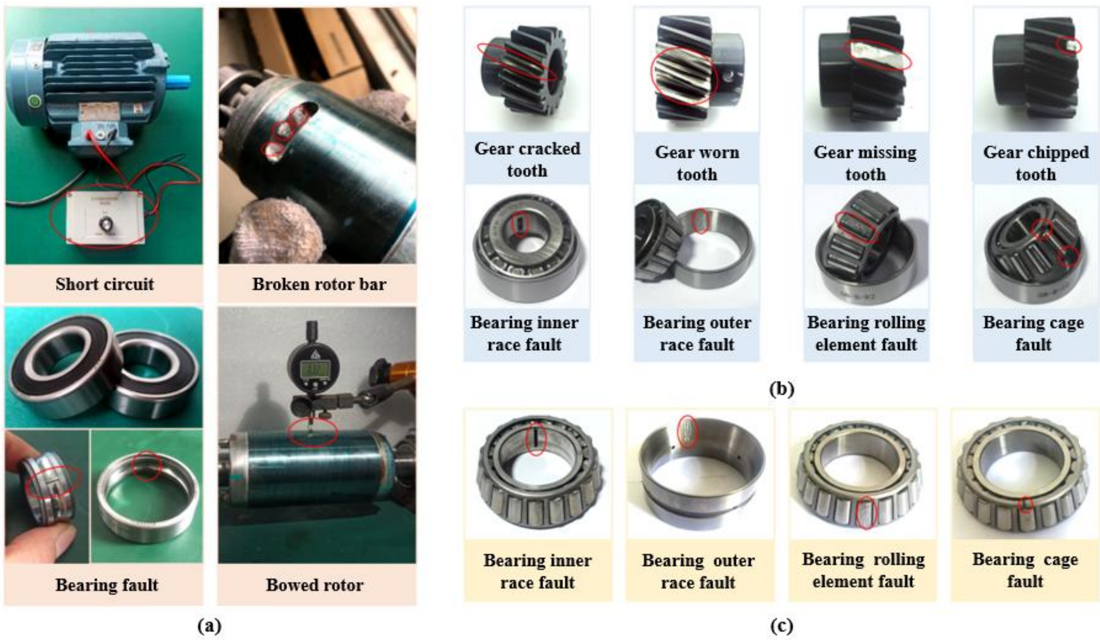

Fig. 2. Photographs of fault simulations. (a) Motor. (b) Gearbox. (c) Axle box

7. Folder structures

The structure of the dataset folder provided at the final stage is shown in Fig. 3 . Training datasets are placed in the ‘Training’ folder. Its subfolders are named based on the sample labels. For component-level compound-fault and system-level compound-fault classes, only representative class examples are provided. Several samples are provided for each health state. Each sample contains 4 CSV files, which are sensor data corresponding to the signal channels of the traction motor, reduction gearbox, and two axle boxes. Test datasets are placed in the ‘Test’ folder. The test samples do not provide labels. After entering the ‘Test’ folder, you will see several subfolders named based on the sample serial number. The internal structure of each sample is the same as that in the training datasets.

It should be noted that the working conditions of the training samples and test samples are not provided, and the working conditions of the test samples are different from those of the training samples, thereby increasing the challenge of competition. Importantly, the test datasets are updated every two weeks, and the classes contained in the test datasets may be changed. Also, the test datasets may contain samples belonging to unseen classes in the training datasets.

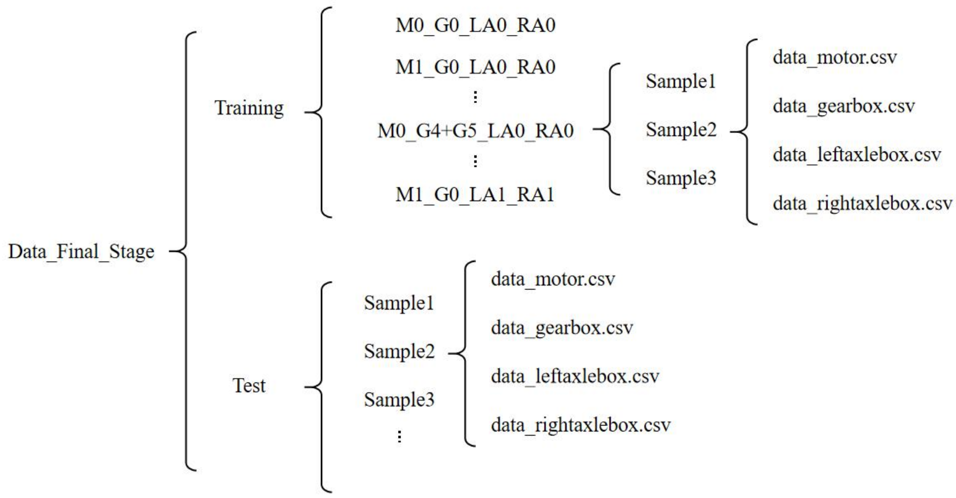

Fig.3 Folder structures of provided datasets

8. CSV files

For the sensor data in the CSV file provided at the final stage, different columns represent different sensor channels, and different rows denote different sampling points. The meaning of the channel has been given in Table 2. Fig. 4 demonstrates an example of motor sensor data, and other CSV files are similar.

9. Contact information

If you have any questions about the datasets, please contact us at any time.
Email: data_challenge2024@outlook.com
Website: https://www.icphm.org

	A	B	C	D	E	F	G	H	I
1	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8	CH9
2	0.290704	-0.01399	0.1209424	-0.262116	-0.979872	0.6350178	25.614489	-0.089915	26.447301
3	-0.288692	-0.121477	-0.452686	-0.88972	-0.289381	0.008954	25.971591	-0.162642	26.522159
4	-0.693015	0.7695266	1.2887994	-0.49862	0.0735342	0.8283818	24.969318	-0.892472	26.349006
5	0.3054926	0.4993003	-0.184208	0.6932759	0.129233	0.568729	25.677983	-0.90142	26.324432
6	-0.47647	1.0982002	1.109972	0.9206978	0.4957514	-0.439305	27.056676	-0.404688	27.773722
7	-0.413583	0.3290961	0.4540013	1.1456427	0.1092991	0.0890916	26.690767	0.0015625	25.653693
8	-0.875968	0.8681059	0.2040235	-0.657975	-0.008316	-0.004381	25.210227	0.3943182	25.71108
9	-1.196723	0.7857259	0.0347528	0.7256518	-0.775403	0.0697116	24.855682	0.8088068	25.726563
10	-2.157148	1.4343143	-0.667469	0.2742875	-0.235384	0.0247417	26.300284	1.2693182	25.682528
11	-1.801821	0.3446203	0.3786482	0.6591436	0.1320772	0.1294022	26.749858	1.0911932	26.909233
12	-1.787363	1.2567622	-0.211183	0.9324216	0.3448718	0.4215266	25.835795	0.4288352	26.536364
13	-1.333511	0.558513	-0.706121	0.999743	0.6303599	0.308486	26.555966	0.3034091	26.447017
14	-1.308101	-0.259854	0.4493213	0.8663701	1.3765526	-0.058269	27.91108	0.8163352	27.063068
15	-0.286586	0.8207243	-1.133605	0.4020575	0.7002985	-0.426428	27.483523	1.3140625	25.745881
16	-0.870924	-0.510621	0.4572946	0.3172099	0.3249419	1.2335956	25.966051	1.5835227	26.264205
17	-0.764129	0.4269143	-0.321666	-0.153714	-0.037818	-0.174752	25.560227	2.0830966	24.739489
18	-0.785936	0.0997994	-0.10845	-0.436696	0.4959278	0.192652	27.01733	2.5377841	26.078977
19	-0.298872	0.1066639	0.1651683	-0.439674	0.4788073	0.1675765	27.451136	2.2840909	26.189063
20	0.3122097	0.3908983	0.2707092	-0.779603	0.210169	0.999214	26.529119	1.6025568	25.721449
21	0.0542548	-1.234367	0.0296785	-1.027044	0.1734357	0.9000405	27.31179	1.3772727	26.249148
22	0.6257749	-0.861269	0.6432357	-0.757699	0.0454295	-0.129791	28.602841	1.9169034	26.600142
23	0.9889263	0.7985231	-0.585875	-0.484391	-0.689225	0.0392949	28.070597	2.4553977	26.08267
24	0.5982351	-0.002936	-0.474942	-1.119085	-0.671464	-0.542561	26.596165	2.718608	25.072727
25	1.1829453	-0.831289	-0.459802	0.2545029	-0.384526	0.4562585	26.180114	3.1798295	25.43608
26	1.2597271	-0.429642	-0.770439	0.3745548	0.2502223	0.3489181	27.505824	3.6413352	25.003551
27	1.5157374	-0.960453	0.1299426	-0.321933	0.4692506	0.1069843	28.008097	3.4090909	25.709517
28	1.3316819	-0.414386	-0.651648	-0.708737	0.5688505	0.0870105	27.165909	2.63125	24.343466
29	0.7579705	-0.485933	-0.555449	0.8835711	0.9414307	-0.165118	27.93608	2.3475852	26.833097
30	0.5318468	-1.197357	-1.320359	1.5213926	1.4738857	-0.405559	29.124858	2.9208807	25.902557
31	0.8421325	0.0182833	-2.389396	0.6497789	0.2884002	-0.798002	28.577699	3.4910511	25.78267
32	0.929744	-0.19066	-1.056518	1.103982	-0.718237	-0.573018	27.132386	3.6421875	25.25767
33	0.1109329	-0.688907	-0.965326	0.8679409	-0.666545	-0.473019	26.490483	4.2015625	24.427131
34	0.4629109	-1.122622	-1.470887	1.1804182	-0.437555	0.2224015	27.82983	4.678267	22.695455
35	0.5809612	-1.035559	-1.464272	-0.08302	-1.358412	-0.489361	28.390341	4.3098011	24.28267
36	-0.301513	0.5829836	-1.229372	-0.677369	-0.984215	-0.606642	27.588494	3.6015625	26.223438
37	-0.61714	-0.770887	-0.32566	-0.743467	-0.81964	-0.724072	28.458239	3.2298295	23.604403
38	-0.117106	-0.617897	0.0360718	0.0351036	0.4339003	-0.79159	29.520313	3.8931818	26.720455

Fig.4 An example of a CSV file containing sensor data of the motor