

文章编号：1005-8451（2017）07-0066-04

CRH380B（L）型动车组中央控制单元数据的分析与挖掘

谢名源¹, 周斌¹, 许帅帅²

(1. 上海铁路局 高速铁路运维技术中心, 上海 201812;

2. 上海铁路局 上海动车段, 上海 201812)

摘要：对CRH380B（L）型动车组中央控制单元（CCU）采集的数据进行梳理和分析，并从便于用户使用和分析的角度对数据进行清洗和关联，同时基于日常运用经验和需求对故障数据进行分析和挖掘，查找潜在的故障关联关系，探索动车组故障的预警预测。

关键词：动车组；中央控制单元；数据挖掘

中图分类号：U266.2 : TP39 **文献标识码：**A

Data analysis and mining of central control unit of CRH380B (L) EMU

XIE Mingyuan¹, ZHOU Bin¹, XU Shuaishuai²

(1. Operation and Maintenance Technical Center of High-speed Railway, Shanghai Railway Administration, Shanghai 201812, China;

2. Shanghai EMU Depot, Shanghai Railway Administration, Shanghai 201812, China)

Abstract: This article combed and analyzed the data collected by the central control unit (CCU) of the CRH380B (L) EMU, cleaned and related the data from the angle of user's ease use and analysis, analyzed and mined failure data based on daily usage experience and requirements, sought the potential fault association, explored the early warning and prediction of EMU faults.

Keywords: EMU; central control unit (CCU); data mining

目前，我国广泛投入运用的CRH380B（L）型动车组中央控制单元（CCU）采用SIBAS32系统，主要安装在动车组两端的司机室，每个动力单元配置一套，采用主、从双机热备份方式进行部署确保系统能够有充足的冗余性。

CCU通过车辆总线（MVB）连接本单元的牵引控制单元（TCU）、制动控制单元（BCU）、安全回路控制系统（SL）、车门控制系统、旅客信息系统（PIS）、火灾报警系统（FAS）等影响列车运行安全和旅客乘坐舒适性的诸多控制模块，并通过列车总线（WTB）与其他牵引单元进行连接和通信，实现列车的联挂和分解功能。

CCU可以实现列车的运行控制、故障诊断、系统维护和调试功能，在动车组的日常维护过程中，通常会使用SIBAS32 MONITOR软件查看和分析动

车组中央控制单元数据，并进行数据的下载和分析，以排查报警或故障发生的原因，消除故障隐患。

1 中央控制单元数据分析

目前，从CRH380B（L）型动车组CCU下载的车载信息主要分为4种：故障环境数据、故障历史数据、故障统计数据、协议数据，包含了动车组运行过程中产生的故障记录、发生时间、故障时车组状态等信息。

1.1 故障环境数据

该数据包含车组号、车厢号、故障代码、故障发生时间、故障恢复时间、列车状态等信息，部分故障发生时也会同时记录制动信息、列车通信设备状态信息及牵引变流器部分状态信息，对于不同的故障代码记录的环境数据会有较大的差异。

1.2 故障历史数据

该数据包含车组号、车厢号、故障代码、故障

收稿日期：2017-03-10

基金项目：中国铁路总公司科技研究开发计划课题（2016J007-B）。

作者简介：谢名源，工程师；周斌，工程师。

描述、故障发生时间、故障恢复时间，是动车组运用过程中故障信息的简要记录。

1.3 故障统计数据

该数据包含车组号、车厢号、故障代码、次数统计、故障描述、维护模式。

1.4 协议数据

该数据包含车组号、车厢号、协议代码、协议描述、发生时间、恢复时间，通过协议代码的形式记录了动车组控制指令及部分主要部件的状态变化情况。

车载信息按照不同的使用需求将动车组的故障信息或状态信息记录成4种不同形式，但在日常的动车组故障排查和分析过程中，往往需要对这4种信息都进行对比和印证分析才能做出更准确的判断。但由于SIBAS32 MONITOR软件在进行车载数据下载时，将这4种信息分别保存为一个文本文档格式的文件，所以在查找和对比故障信息时存在很大的不便，因此，如何对这4种信息进行解析、清洗和关联显得尤为重要。

2 中央控制单元数据的清洗和关联

2.1 数据的清洗

由于目前下载的4种车载信息采用文本文档的格式进行存储，所以要进行清洗就必须对文件进行批量解析，通过分析4种信息包含的字段发现故障历史数据、协议数据、故障统计数据具有固定的字段信息，对于每一条故障/协议描述都包含了相同的字段，而故障环境数据因故障代码的不同会包含不同的字段，且包含的字段数量也各不相同。因此，选择故障历史数据为主要解析对象，将车组的故障历史数据逐条解析成为包含10个字段的故障记录，如图1所示。将协议数据逐条解析成包含9个字段的协议信息，故障环境数据按照车组号、车厢号、故障代码、发生/恢复时间的不同解析成不同的信息块，而故障统计数据的信息可通过故障历史数据进行体现，所以暂不解析。

根据以上要求对上海动车段2013年~2015年CRH380B(L)型动车组下载的中央控制单元数据进行解析后，共解析出故障记录13 751 311条，

car_group_number	car_unit_number	id	fail_code	fail_description	set_date	set_time	reset_date	reset_time	n_mode
3571	01	1	66DC	Contactor urgency	2014-12-10	00:41:17	(null)	(null)	0
3571	01	2	610A	UB request by EB	(null)	(null)	2014-12-10	00:36:27	0
3571	01	3	66D5	Urgency brake req	(null)	(null)	2014-12-10	00:36:27	0
3571	01	4	610B	Contactors EB req	(null)	(null)	2014-12-10	00:36:26	0
3571	01	5	610B	Contactors EB requ	2014-12-10	00:35:45	(null)	(null)	0
3571	01	6	610A	UB request by EB	2014-12-10	00:35:35	(null)	(null)	0
3571	01	7	66D5	Urgency brake req	2014-12-10	00:35:35	(null)	(null)	0
3571	01	8	66DC	Contactor urgency	(null)	(null)	2014-12-10	00:32:55	0
3571	01	9	610A	UB request by EB	(null)	(null)	2014-12-10	00:30:57	0
3571	02	10	6305	Vacuum circuit bree	(null)	(null)	2014-12-10	00:20:44	0
3571	02	11	6305	Vacuum circuit bree	2014-12-10	00:30:39	(null)	(null)	0
3571	01	12	684F	No automatic train	(null)	(null)	2014-12-10	00:30:01	0
3571	01	13	6836	Fault switch AT&Fa	(null)	(null)	2014-12-10	00:29:58	0
3571	01	14	6075	CCUmester/slave c	(null)	(null)	2014-12-10	00:29:55	0
3571	01	15	610A	UB request by EB	2014-12-10	00:29:43	(null)	(null)	0
3571	01	16	6858	Storage status whe	(null)	(null)	2014-12-10	00:29:41	0
3571	01	17	6858	Storage status whe	2014-12-10	00:29:49	(null)	(null)	0
3571	01	18	684F	No automatic train	2014-12-10	00:29:35	(null)	(null)	0
3571	01	19	6412	Urgency brake val	2014-12-10	00:29:24	(null)	(null)	0
3571	01	20	6836	Fault switch AT&Fa	2014-12-10	00:29:19	(null)	(null)	0
3571	01	21	6316	Shutdown profiled	(null)	(null)	2014-12-10	00:29:19	0
3571	01	22	66A8	ASD not tested	(null)	(null)	2014-12-10	00:29:19	0
3571	01	23	6075	CCUmester/slave c	2014-12-10	00:28:34	(null)	(null)	0
3571	01	24	6412	Urgency brake val	(null)	(null)	2014-12-10	00:28:34	0
3571	01	25	66DC	Contactor urgency	2014-12-09	23:17:35	(null)	(null)	0

图1 故障记录包含的字段

通过对比解析的故障数据与动车组真实故障的记录，可以看出中央控制单元数据记录的故障数量，远远超过动车组实际发生的故障数量，通过分析总结出以下几条主要原因：

(1) 维护模式下进行测试的数据也被记录成故障信息；

(2) 每条故障信息发生和恢复时都被独立记录成一条故障信息；

(3) 动车组在正常运用、检修过程中，由于主控的投入/退出、升降操作等常伴随有动车组自检过程中产生的故障信息，随着自检的完成故障信息会自动消除。

基于以上的分析，有针对性地制定了动车组故障记录清洗规则，对解析的故障记录进行逐步清洗，消除无效的故障记录，为后续的故障关联和数据挖掘做好预处理工作。制定的清洗规则如下：

(1) 清除故障信息不完整的数据；

(2) 清除维护模式下的故障信息；

(3) 清除故障恢复时保存的故障信息；

(4) 清除重复下载产生的重复数据；

(5) 清除不包含在TD故障代码表内的故障信息；

(6) 清除同一个月内同一车组同一故障代码发生次数>30，且发生天数>25的故障。

通过清洗，动车组故障记录剩余208 334条，下降为原始故障数据的1.52%，故障记录数量有了极大的下降，数据的清洗过程如图2所示。

2.2 数据的关联

为了对故障信息进行分析排查，需要将故障发生的车组号、车厢号、发生时间进行精确定位，同



图2 故障数据的清洗过程

时将故障发生前相关的协议数据进行分析，将这3种信息有效地关联在一起，将极大地提高故障分析和排查的效率。

同时通过对故障环境数据、故障历史数据、协议数据包含的内容进行对比分析可以看出，这3种信息均包含了车组号、车厢号、（故障/协议代码）发生时间，因此可将故障环境数据、故障历史数据、协议数据的信息进行关联。

(1) 以车组号、车厢号、发生日期、发生时间为关联条件，将故障历史数据与协议数据进行关联；

(2) 以车组号、车厢号、发生日期、发生时间、故障代码为关联条件，将故障历史数据与环境数据进行关联。

通过以上条件，将动车组车载故障信息按照车组号、车厢号、故障代码、发生日期、发生时间逐条进行关联，同时可自由设定车组号、故障代码、发生时间段等条件对故障信息进行查询，从而便于技术人员对故障的排查和分析，关联完成的故障信息形式如图 3 所示。

图3 动车组故障信息

其中，协议数据默认以故障代码发生时间为起点，向前倒查 15 min 内相关的协议代码，并可通过

手动设置倒查时间对更早期的协议代码进行查询，关联的协议数据如图 4 所示。

协议数据追踪							
序号	车组号	车厢号	协议代码	协议描述	日期	时间	
T05	3572	01	903S	WC3 OFF by ITW [C01]	2014-12-22	01:11:41	
T12	3572	01	914B	WC3 lock by TCU [D1 s]	2014-12-22	01:11:41	
T13	3572	01	914A	No WC3 release traction (TCW1 or TCW2)	2014-12-22	01:11:41	
T14	3572	01	90CD	WC3 OFF by traction [TCW]	2014-12-22	01:11:41	
T15	3572	02	0039	WC3 arm TU is ON	2014-12-22	01:11:40	
T16	3572	01	903R	C005: WC3 release	2014-12-22	01:11:40	
T17	3572	01	903S	C004: WC3 release	2014-12-22	01:11:39	
T19	3572	02	909Z	WC3 arm TU is ready for closing	2014-12-22	01:11:29	
4028	3572	01	9071	WC3 OFF by C0W [C005]	2014-12-22	01:11:41	
4031	3572	01	914B	WC3 lock by TCU [D1 s]	2014-12-22	01:11:40	
4032	3572	01	914A	No WC3 release traction (TCW1 or TCW2)	2014-12-22	01:11:40	
4035	3572	02	9070	WC3 arm TU is ON	2014-12-22	01:11:40	
4036	3572	01	9080	C005: WC3 release	2014-12-22	01:11:40	
4036	3572	02	909Z	WC3 arm TU is ready for closing	2014-12-22	01:11:29	
7889	3572	08	9005	WC3 OFF by ITW [C01]	2014-12-22	01:11:41	
7892	3572	08	914B	WC3 lock by TCU [D1 s]	2014-12-22	01:11:41	
7893	3572	08	914A	No WC3 release traction (TCW1 or TCW2)	2014-12-22	01:11:41	
7894	3572	08	90CD	WC3 OFF by traction [TCW]	2014-12-22	01:11:41	
7895	3572	08	903Y	C005: WC3 release	2014-12-22	01:11:40	
7897	3572	08	903E	C004: WC3 release	2014-12-22	01:11:39	
7898	3572	08	0001	Command WC3 by train driver	2014-12-22	01:11:39	
7905	3572	08	9011	Drivers cab occupied	2014-12-22	01:11:28	
7905	3572	08	902B	Indicator light doors displays door(s) open	2014-12-22	01:11:28	
7907	3572	08	912E	E8 request by E8-K05	2014-12-22	01:11:28	
7908	3572	08	9005	Direction switch forward	2014-12-22	01:11:28	
7909	3572	08	9093	Key switch on	2014-12-22	01:11:28	
1191	3572	08	914B	At least one door is open in car 1	2014-12-22	01:11:01	
1192	3572	08	914D	All doors closed	2014-12-22	01:06:54	

图4 协议数据

环境数据按照车组号、车厢号、故障代码、发生时间为标识，将环境数据解析为每个故障代码对应一个环境数据信息块，关联的环境数据如图 5 所示。

环境数据追踪	
HC Car number	= 13303572
Driver identifier	= 0
Sequence position number	= 3.17
Diagnostic code	= 6M48
Frequency	= 2
in maintenance mode	= 0
Feedback tube heater implausible	
Serial number of occurrence	= 1712
Train status	
CPU is master.....	= 0
Position 5 MYB-wide	= 0
Position 8 MYB-wide	= 1
Drivers cab occupied, DS in %...	= 0
Own drivers cab occupied.....	= 0
Own TU leading.....	= 0
Own TU guided.....	= 1
Tractive effort controller 0 pos...	= 1
Own TU PANTO up.....	= 1
PDU PANTO up.....	= 0
VCU 1 on.....	= 0
TCU 1 or TCU 2 have start-up command	= 1
Main key switch in position on.....	= 0
External power supply active.....	= 0
Maintenance mode in activated.....	= 0

图5 环境数据

通过以上的方式对 CCU 数据进行清洗和关联后，车载故障记录已形成格式化的规范数据，同时建立的相互关联关系使得故障记录更加清晰和有序，提升了故障原因的分析排查效率。

3 故障数据的分析和挖掘

通过对数据的清洗和关联，可以为技术人员排查故障提供高效的工具，但这样的方式仅适合于故障发生后对故障原因的排查，为了挖掘故障数据潜在的关联关系，对清洗过后的故障信息采用数据挖掘的方法进行关联性分析，使得结果更加合理和明

表1 故障关联分析结果

序号	关联故障代码数量	故障代码关联式	支持度	置信度 (%)	符合规则次数	左端发生次数	右侧代码是否为TD	故障代码描述
1	4	66D6 6836 683D →65B0	0.33	98.89	89	90	TRUE	(通过制动控制器请求紧急制动)(TCR 故障开关44-S27 置在关闭位置)(GFX-3A 开关44-S13 置在位置开) → (MCB 44-F11 ETCS BT 断开)
2	3	6700 675B →66EB	0.37	93.88	92	98	TRUE	(紧急制动环路状态控制系统和端车 BCU 不一致(司机))(由于故障, 紧急制动阀未启动(司机)) → (请求紧急制动, 接触器43-K24 未闭合(司机))
3	4	6042 6052 62FC → 68C4	0.26	93.42	71	76	TRUE	(CCU1 失效)(CCU2 失效)(由于故障旁路 TCU1, VCB 释放闭锁) → (至少一个对轮轴承测量温度故障)
4	4	63D6 64A7 651D →64A9	0.26	90.48	57	63	TRUE	(21-S01: 紧急关闭按钮启动)(KS 外部供电 31-S01 置在位置'外部')(电池充电器无法启动加载操作) → (运行外部供电)
5	3	6384 6740 →63D6	0.32	87.78	79	90	TRUE	(启动紧急模式)(转向架监控环路故障开关 43-S25 置在'断开'位置) → (21-S01: 紧急关闭按钮启动)

确, 对分析的结果预先进行限定, 要求如下:

- (1) 关联故障代码必须为同一车组发生的故障;
- (2) 关联故障代码必须遵循同样的时间顺序, 即故障是相继发生且顺序一样, 其中时间间隔可调整;
- (3) 关联故障代码最后一个代码必须为 TD 故障代码表内的故障信息;

根据以上要求, 对经过清洗后的 208 334 条故障记录进行关联性分析, 通过使用成熟的商用软件进行建模和计算后, 部分计算结果如表 1 所示。

以上计算结果是基于对故障代码发生的时间间隔设定为 24 h 以内得出, 可以看出, 在对车载故障代码进行分析和挖掘时, 生成的发生次数较多、置信度较高的关联故障代码往往可能是列车运用、检修过程中人为操作造成, 如计算结果中的第 4 条故障代码关联式反映的是动车组入库进行 I2 修采用外接电源的方式进行作业时, 依次操作产生的故障代码。

4 结束语

随着 CRH380B(L) 型动车组不断投入实际运用, 对动车组 CCU 记录的故障数据规律的分析和研究也将不断加深, 通过采用计算机软件进行建模分析, 结合专业技术人员的实际运用经验, 能够探索和建立起动车组故障预警和预测的有效手段, 为动

车组的安全运行保驾护航。

参考文献:

- [1] 黄学文, 刘春明, 冯璨, 等. CRH3 高速动车组故障诊断系统 [J]. 计算机集成制造系统, 2010, 16 (10): 2311-2318.
- [2] 汲磊举. 大数据环境下动车组故障关联关系分析关键技术研究与实现 [D]. 北京: 北京交通大学, 2016.
- [3] 马海漫. 动车组状态修中故障知识获取模型的研究 [D]. 北京: 北京交通大学, 2013.
- [4] 魏海洋, 周纪超, 穆俊斌, 等. CRH3 型动车组故障诊断代码编码规则研究 [J]. 大连交通大学学报, 2014, 35 (4): 102-105.
- [5] 沙森, 王伟, 徐建波. CRH5 型和 CRH3 型动车组列车网络控制系统的比较 [J]. 铁道车辆, 2012, 50 (1): 25-27.
- [6] 张密哲, 李蕊, 王振宏. 浅谈 CRH3 型高速动车组远程数据传输系统 [J]. 现代制造技术与装备, 2013 (1): 57-58.
- [7] 田静宜, 杨业, 杨雪峰, 等. 高速列车智能化故障诊断方法 [J]. 化工自动化及仪表, 2013 (4): 531-533.
- [8] 曾忠禄. 大数据分析: 方向、方法与工具 [J]. 情报理论与实践, 2017, 40 (1): 1-5.

责任编辑 陈蓉