



# 高速列车运行控制系统车载子系统故障预测方法研究

臧钰<sup>1</sup>, 蔡伯根<sup>1,2\*</sup>, 上官伟<sup>1</sup>, 王化深<sup>1</sup>

1. 北京交通大学电子信息工程学院, 北京 100044

2. 北京交通大学轨道交通控制与安全国家重点实验室, 北京 100044

\* 通信作者. E-mail: bgcai@bjtu.edu.cn

收稿日期: 2019–10–15; 修回日期: 2019–12–21; 接受日期: 2020–01–24; 网络出版日期: 2020–04–13

国家自然科学基金重大项目 (批准号: 61490705) 资助

**摘要** 列车运行控制系统是高速列车信息控制系统的“神经中枢”, 其结构复杂、组件繁多, 其中, 车载子系统是列车运行控制系统的核心组成部分, 是保证行车安全、提高运行效率的关键. 目前, 车载子系统的故障数据处理方式仍以人工处理实现故障定位为主, 尚未深入到系统故障机理层面, 无法实现有效的故障预测. 本文分析了车载子系统的结构及系统故障处置现状, 总结了现存的主要问题, 结合车载子系统结构及各模块性能参数, 提出了基于贝叶斯网络 (Bayesian network, BN) 的系统级故障预测模型构建方法. 利用实际系统现场运行数据, 采用基于贝叶斯网络的方法进行故障预测, 分别在 20, 200, 2000, 20000 组数据条件下实施了验证, 故障预测准确率分别为 5%, 27%, 92%, 96.3%, 在 2000 组数据条件下同时对隐马尔科夫模型 (hidden Markov model, HMM)、神经网络 (neural network, NN) 与本文所提出的方法进行了对比, 预测结果验证了贝叶斯网络在系统级故障预测方面的显著优势.

**关键词** 高速列车, 列车运行控制系统, 车载子系统, 故障预测, 贝叶斯网络

## 1 引言

近年来, 我国高速铁路一直处于快速发展时期, 截至 2019 年底, 我国全年新增高铁里程超过 4000 公里, 总里程突破 3.5 万公里, 稳居世界第一. 《中长期铁路规划网》(2016~2030) 中规划了“八纵八横”的高速铁路网, “交通强国”战略引领高速铁路发展进入新的阶段, 更高的效率、更优的系统能力对高速列车的运行安全提出了更加严苛的要求.

列车运行控制系统作为高速列车信息控制系统的关键组成部分, 是铁路运输的“神经中枢”, 是保障高速铁路行车安全、提高运输效率的核心<sup>[1]</sup>. 高速列车系统规模庞大、结构复杂、组件繁多, 本文

**引用格式:** 臧钰, 蔡伯根, 上官伟, 等. 高速列车运行控制系统车载子系统故障预测方法研究. 中国科学: 信息科学, 2020, 50: 511–526, doi: 10.1360/SSI-2019-0227  
Zang Y, Cai B G, Shanguan W, et al. Research on the fault prediction method of an on-board subsystem in high-speed train control systems (in Chinese). Sci Sin Inform, 2020, 50: 511–526, doi: 10.1360/SSI-2019-0227

以 CTCS-3 级 (China train control system level 3) 列车运行控制系统为研究对象, 该系统基于 GSM-R (global system for mobile-railway) 无线通信系统实现地面与列车控制信息双向实时传输<sup>[2]</sup>, 系统主要由 3 部分构成: 地面子系统、车载子系统以及车地通信网络. 其中, 车载子系统由车载安全计算机、轨道电路信息读取模块、应答器模块、无线传输模块、人机界面模块、列车接口模块、测速测距模块和司法记录模块等组成, 根据地面子系统提供的列车信息、线路参数、临时限速信息生成目标距离速度曲线, 监控列车安全运行. 车载子系统是列车运行控制系统的核心组成部分, 是保证行车安全、提高运行效率的关键<sup>[3]</sup>.

故障预测及健康管理是保证列车行车安全、提高运行效率的有效手段, 其中故障预测是实现健康管理的基础. 故障预测的定义是指根据系统现在或历史性能状态预测部件或者系统在未来何时发生故障及设备的剩余寿命<sup>[4]</sup>. 故障预测方法总体上可以分为 3 类: 基于模型、数据驱动、统计可靠性方法. 顾名思义, 基于模型的方法要求已知系统的精确数学模型. 文献 [5,6] 建立了轴承的机械模型, 通过对两个缺陷传播模型进行参数估计, 实现剩余寿命评估. 数据驱动方法利用系统在设计、运营、维护等各阶段的历史数据, 提取与故障相关的性能参数训练模型, 然后输入观测数据进行验证. 文献 [7] 应用神经网络 (neural network, NN) 方法训练历史数据建立轴承故障预测模型, 从而进行剩余使用寿命预测. 上述两类方法中, 基于模型的方法能够深入对象系统本质的性质 (失效机理) 进行实时故障预测, 但要求对象系统的数学模型已知. 数据驱动方法以采集的数据为基础, 通过各种数据分析处理方法挖掘其中的隐含信息进行预测操作. 但是数据一般都具备不确定性及不完整性. 统计可靠性方法中贝叶斯网络 (Bayesian network, BN) 为典型代表, 通过概率密度函数涵盖不同结点之间潜在的关联关系, 克服了其他两类方法的缺陷, 非常适合复杂系统. 另外, 工程设备中的设备性能退化过程都服从 Weibull 分布<sup>[8]</sup>, 因此利用 Weibull 分布表征设备故障分布, 从而实现寿命预测也是该类方法中的一个方案.

但是, 目前对车载子系统的故障研究主要集中在故障诊断方面, 针对故障预测方面的研究尚不多见, 北京交通大学梁潇等<sup>[9]</sup> 将贝叶斯方法应用到车载设备的故障诊断中, 从历史数据中获得故障征兆数据, 并利用粗糙集方法进行属性简约, 从而降低了模型训练的复杂度, 并利用专家知识改进了贝叶斯模型. 北京交通大学赵阳等<sup>[10]</sup> 同样也将贝叶斯网络应用到车载设备的故障诊断中, 并通过不同结构学习算法的比较获得最优的网络结构, 最终与神经网络方法比较证明了该方法的优越性. 文献 [11] 将文本挖掘方法应用到车载设备故障诊断中, 建立了适用于车载设备的贝叶斯结构学习算法, 并验证了该方法的有效性. 文献 [12] 将案例推理方法应用到车载设备的故障诊断中, 通过对典型故障进行案例表示, 对案例的组织、修正、学习及维护, 提高车载设备的维护维修效率. 文献 [13] 利用文本信息抽取的方法建立了车载设备的故障字典, 并建立了车载子系统故障检测的自动化处理平台. 文献 [14] 按照硬件软件对车载设备故障进行分类, 然后利用模糊贝叶斯网络实现了故障诊断. 可见, 上述研究的基本思想均针对文本文件, 实现故障定位, 提高现场数据分析工区的工作效率, 无法实现故障预测.

高速列车列控车载子系统内部组件繁多, 故障机理复杂, 无法建立精确的数学模型, 因此基于模型的方法不适用. 另外, 列控车载子系统设备均采用独立超限报警机制, 无法通过安置传感器获取大量的性能参数数据进行模型训练, 因此数据驱动的方法不适用, 如何有效利用现场实际数据实现车载子系统故障预测成为一大难点. 利用实际现场统计数据, 通过贝叶斯网络的结构学习和参数学习算法, 反演高速列车列控车载子系统状态演变, 并扩充样本数据实现系统级故障预测成为可行方案. 本文首先介绍了 CTCS-3 级车载子系统结构, 然后分析了高铁运营管理现场数据处理现状, 总结了现存的突出问题. 然后介绍了故障预测相关概念及算法, 基于车载子系统结构及各模块性能参数构建故障预测模型, 利用贝叶斯网络在不同数据量情况下进行故障预测, 并与隐马尔科夫模型 (hidden Markov model, HMM) 及神经网络两种算法进行对比验证.

## 2 问题描述

### 2.1 系统结构

目前在用的列控车载子系统型号有 CTCS-300T, CTCS-300S, CTCS-300H, CTCS-200H, CTCS-200C 5 种, 本文以 CRH3 (Chinese high-speed railway) 系列动车组 300T 型列控车载子系统为例进行研究, 300T 型列控车载子系统由列车自动防护控制单元 (automatic train protection control unit, ATPCU)、CTCS-2 控制单元 (CTCS2 control unit, C2CU)、应答器信息接收单元 (Balise transmission module, BTM) 及应答器天线 (compact antenna unit, CAU)、速度距离处理单元 (speed and distance processing unit, SDP)、轨道电路信息读取器 (track circuit reader, TCR) 及接收天线、测速测距单元 (speed and distance unit, SDU)、安全数字输入输出 (vital digital input/output unit, VDX) 单元、安全无线传输单元 (safe transmission unit-vehicle, STU-V)、无线通信单元 (remote terminal unit, RTU)、多普勒雷达 (radar)、速度传感器 (odometer, ODO)、人机界面 (driver-machine interface, DMI)、司法记录单元 (judicial recorder unit, JRU)、列车网关 (train signaling gateway, TSG) 等组成。

在 CTCS-3 运行模式下, ATPCU 负责监控静态速度、动态速度和目标距离, 监控 CTCS-2 (China train control system level 2) 的运行状态. C2CU 负责列车在 CTCS-2 运行过程中监控静态速度、动态速度和目标距离等. TSG 主要用于实现各模块之间的数据交换. SDP 软件负责速度和距离数据的处理, 以及车载子系统单元和设备通信的平台主控器处理. DMI 主要用于显示列车运行的速度、距离、制动状态等信息, 以及司机的输入和操作. SDU 软件负责处理从速度传感器和测速雷达接收到的脉冲数据. VDX 用于输出紧急制动并采集制动反馈. STU-V 主要负责对无线数据进行加密和安全传输. RTU 软件负责车载子系统无线通信功能, 处理无线数据的传输. TCR 的主要功能是对接收到的轨道电路编码的解码器. BTM 将处理后的有效信息通过多功能车辆总线 (multi-function vehicle bus, MVB) 发送给 ATPCU 和 C2CU. 雷达通过频率的改变数值, 计算出目标与雷达的相对速度. 速度传感器通过检测轮轴转速的方式, 得到当前列车运行速度. JRU 用于记录列车运行中, 车载设备采集的原始信息和车载设备输出的控制信息. 以上模块协同工作, 通过过程现场总线 (process field bus, PROFIBUS)、MVB 总线进行信息交互, 完成车载子系统功能, 各单元模块排列情况如图 1 所示, 该车载子系统结构图来自《CRH3, CRH380B, BL 型车 CTCS3-300T 列控车载设备维护说明书 V2.3.0》.

### 2.2 故障处置现状

以 300T 型车载子系统为例, 高速列车列控设备现场运营管理工作中, 目前的车载子系统故障数据主要为 AE (application event) Log 文件. 当发生重要事件时, 例如启动失败、系统异常时, 把执行成功的语句或者错误代码写入车载子系统内部模块的一块非易失存储区, 通过 AE 命令从终端导出, 即使系统断电仍可以保存并导出. 由于存储空间有限, AE Log 只能记录系统运行过程的重要进程, 如果超过限制, 旧的记录将会被新的记录覆盖.

目前, 现场运营过程对于此类文本文件主要采用人工阅读方式处理, 如果在阅读过程中发现问题, 则将问题概况告知负责人; 如果可以定位故障点, 则按照案例库中的处理意见进行处理; 如果仍然无法定位故障, 则需要联系厂家确定故障点及处理办法. 在整个流程中, 做好记录工作. 因此利用 AE Log 可以初步断定在何时发生了何种故障. 可见目前分析的数据为文本文件, 对车载子系统运行时的事件进行记录, 对设备底层信息描绘较少, 以文字描述为主, 由人工处理, 时间成本大, 虽然可直接读取, 但可用信息深度不够, 人工处理时间消耗较大, 此类数据只能实现故障检测、故障定位, 无法用于故障预测.

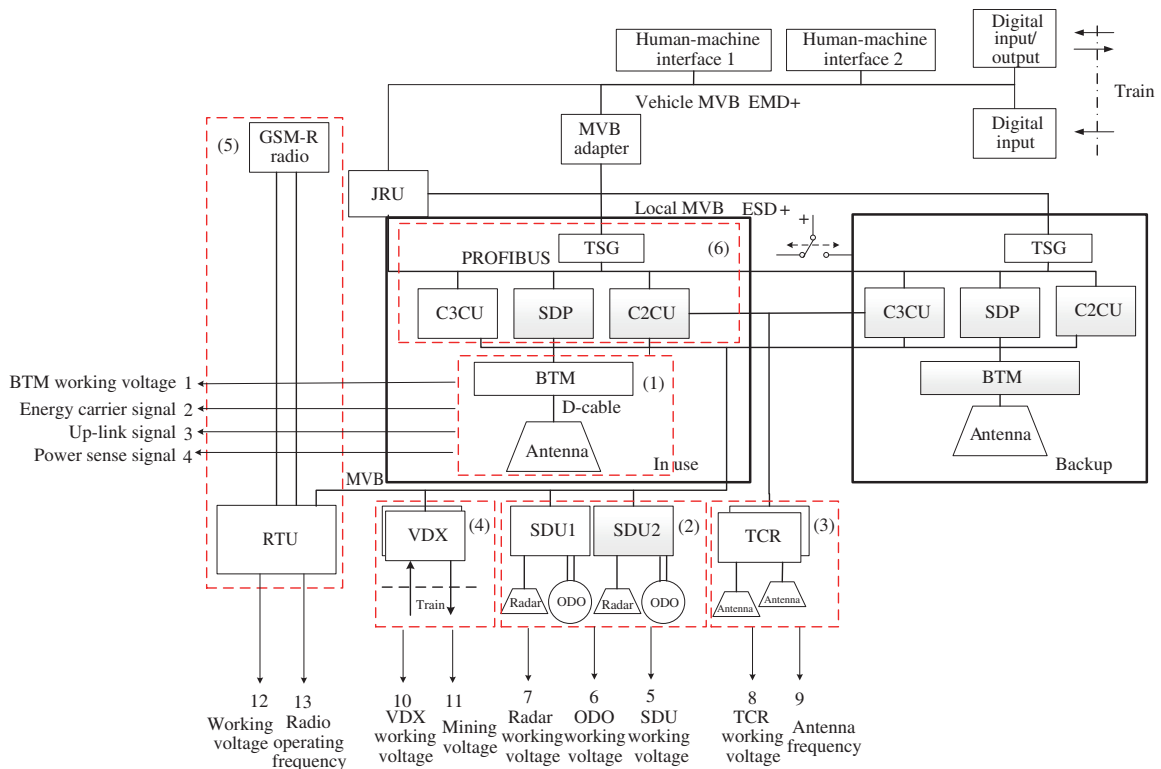


图 1 (网络版彩图) 300T 型列车车载子系统结构

Figure 1 The structure of 300T on-board subsystem in train control system

另一类文本文件针对设备底层电气特性, 采用专用测试平台, 由人工进行测试并记录, 对照元器件基本电气特性所发现的电路、芯片等底层故障现象. 动车段继电器在更换前均需在专用的测试平台进行测试, 测试温度、湿度、电压、电阻等相关电气特性参数, 对照设备的参考范围判断该继电器是否合格. 该类数据是针对继电器设备级的参数, 相对以上数据来说, 层次较深, 但目前动车段工区仅限于判断产品是否合格, 并未研究参数变化与继电器故障之间的关系.

结合现场设备资料及日志分析发现, 300T 车载子系统中 VDX 更换量较大. 受限于工厂 VDX 设备产能和测试台检测周期, 设备返修周期相对较长, 备品备件供应也比较紧张, 因此给现场的维护工作带来了较大的压力. 针对这一问题, 工作人员经统计分析, 发现 VDX 单元发生较为集中的为时序相关问题, 该问题一般都在启机制动测试过程中发生, 是造成现场 VDX 更换量较大的主要原因. 为缓解现场 VDX 单元的返修压力, 以下针对该类问题进行故障机理分析, 并提出短期内 VDX 单元的现场应急维护建议, 证明现场已经有针对基于故障机理分析的故障预测出现.

### 2.3 需求分析

如上述讨论, 现场数据存在信息深度不够、以文字表述为主、人工处理等特点, 导致时间人力成本消耗过大, 且仅能实现故障检测及定位. 目前已有高速铁路运营单位初步建立了列控车载子系统智能化检测平台, 该平台可以大大提高铁路局现场日志数据处理效率, 从而提高维修效率, 但是主要针对日志数据, 且主要目的是实现故障定位, 对于状态数据的分析较少, 很难深入到故障机理层面, 从而无法实现故障预测. 故障预测需要基于设备故障机理, 提取表征设备故障机理的性能参数, 目前该类数据

并未用于故障预测的研究,导致现场对于故障的处理长期处于被动,即故障后才能安排故障检测及维护等活动,对于列车的安全运行十分不利,因此,故障预测算法对于提高列车的安全运行,主动式维修方式的改变至关重要,本文基于高铁列控车载子系统运营维护现状及现场实施故障预测的需求,在故障预测算法设计、车载子系统状态预测及对比验证等方面开展具体工作.

### 3 算法设计

故障诊断是一个执果索因的过程,即在系统发生故障时,利用模块之间的因果关系以及专家经验定位系统具体的故障点.故障预测是一个执因求果的过程,在系统正常运行的情况下,通过与故障相关关键信号测试点的变化趋势建立起来的预测模型,利用先进的预测算法推理系统在当前运行状况下发生故障的概率.贝叶斯网络算法是一个典型的推理算法,相较于神经网络算法,其优势在于其推理能力,且对先验概率的要求较低,对数据量要求较小<sup>[15]</sup>;相较于隐马尔科夫方法,其优势在于贝叶斯网络可以动态调整网络结构与参数,预测能力更高<sup>[16]</sup>,因此贝叶斯网络被越来越广泛地应用<sup>[17,18]</sup>.2.2小节介绍了目前实际现场的故障数据类型,为文本类文件,多描述功能性故障,没有针对系统电气特性参数的记录与分析,因此很难提供大量的数据支撑神经网络算法.另外,专家及各类标准等先验知识信息对于高铁列控车载子系统的故障预测至关重要,因此贝叶斯网络在车载子系统的故障预测方面优于隐马尔科夫模型.本文将设计基于贝叶斯网络的列控车载子系统故障预测算法实现车载子系统状态的预测.

#### 3.1 贝叶斯网络算法

贝叶斯网络是一个有向无环图,结点是离散变量,代表所应用模型中的变量,结点之间有向连接弧表示结点之间的因果关系,在父节点给定的条件下,结点的概率值由结点参数值表示<sup>[19]</sup>.

数学模型上,用  $D$  表示有向无环图,存在一个离散变量集合  $F = F_1, F_2, \dots, F_n$ , 且  $F$  中的元素与  $D$  中的结点一一对应,并且满足

$$P(F_1, F_2, \dots, F_n) = \prod_u p(F_u | \text{pa}_u),$$

其中  $P$  为联合概率分布,  $\text{pa}_u$  是  $D$  中  $F_u$  的第一个父结点.

贝叶斯网络的基本算法包括结构学习算法和参数学习算法,不同的学习算法下,又分为数据完整和数据不完整两种情况.

##### 3.1.1 结构学习算法

结构学习算法分为数据完整和数据不完整两种情况,数据完整情况下的典型算法为 K2 算法<sup>[20]</sup>,K2 算法的基本思想为首先给定结点顺序,将此作为先验信息,利用贝叶斯网络打分方法评价网络模型与数据的贴合程度,然后不断向网络中添加边,寻找出最佳的结构.

贝叶斯打分方法原理如下,设模型结构为随机变量,用  $B_s$  表示,给定随机变量参数  $\theta_{B_s}$ ,结构网络的先验概率为  $P(B_s)$ ,参数  $\theta_{B_s}$  的先验概率为  $P(\theta_{B_s} | B_s)$ ,则可以得到

$$P(B_s, \theta_{B_s}) = P(B_s)P(\theta_{B_s} | B_s).$$

假设完备的数据集为  $V$ ,计算后验概率分布的公式为

$$P(B_s, \theta_{B_s} | V) = P(V | B_s, \theta_{B_s})P(B_s, \theta_{B_s}).$$

$P(B_s, \theta_{B_s} | V)$  即为贝叶斯打分模型, K2 算法分为以下两步:

步骤 1. 寻找结构  $B_s$ , 且确认其后验概率最大.

步骤 2. 对于步骤 1 中找到的  $B_s$ , 进行贝叶斯参数打分, 选取分数最高的网络结构:

$$B_s^* = \arg \max_{B_s} P(B_s | V).$$

寻求后验概率最大的网络结构的过程实质上是最优化的过程, 目标函数为

$$\log P(B_s, \theta_{B_s}) = \log P(V | B_s) + \log P(B_s).$$

数据不完备情况下的结构学习比较复杂, 目前研究相对较少, 故本文采用 K2 算法进行结构学习.

### 3.1.2 参数学习算法

参数学习也分为数据完整和数据不完整两种情况, 数据完整情况下的经典参数学习方法为最大似然估计 (maximum likelihood estimation, MLE) 方法<sup>[21]</sup>. 该方法的原理为, 令  $\alpha$  为一个未知参数,  $W$  为训练样本集,  $D$  为网络结构, 在给定  $D$  和  $W$  时, 具有最大后验概率时参数  $\alpha$  的取值即为所求参数值, 其计算公式为

$$P(\alpha | W, D) = \frac{p(W | \alpha, D) P(\alpha | D)}{p(W | D)}.$$

$P(\alpha | D)$  为给定网络结构  $D$  下参数  $\alpha$  的先验概率数据. 先验概率  $P(\alpha | D)$  服从 Dirichlet 分布, 参数  $\alpha$  的后验概率为

$$P(\alpha | W, D) = \text{Dir}(\alpha | \theta_1, \dots, \theta_z) = \frac{T(\theta)}{\prod_z T(\theta_z)} \prod_z \alpha^{\theta_z - 1},$$

其中  $\theta_1, \dots, \theta_z$  为一组超参数,  $\sum_z \alpha_z = 1$  且  $0 \leq \alpha_z \leq 1$ .

不完备情况下的参数学习算法为期望最大化算法 (expectation maximization, EM), 本文先研究数据完整情况下的预测, 后续会进一步研究数据不完整情况.

## 3.2 车载子系统贝叶斯网络模型构建

车载子系统部件众多, 每个部件的信号监测点繁多, 很难通过数据统计方法进行预测, 而且车载子系统是一个典型的安全苛求系统, 无法安装传感器获取所需信号的信息, 各种监测信号之间的相互作用机理不明确, 各功能单元模块通过电缆总线传输信息, 故障并不独立, 每个模块的故障机理也无法全部明确, 因此提取相关监测信号, 经过贝叶斯算法的推理, 对车载子系统的故障进行预测是一个合理可行的方法.

基于贝叶斯网络的车载子系统故障预测步骤如图 2 所示.

首先, 结合系统结构功能分析建立车载子系统的贝叶斯网络结构, 基于该结构, 利用现场统计数据对贝叶斯网络结点参数学习, 最终建立车载子系统贝叶斯网络, 根据实际现场统计的各个功能单元的故障率, 反推系统状态, 实现样本扩充<sup>[22]</sup>. 基于该网络产生不同训练参数集, 然后随机选择贝叶斯网络的初始参数值, 利用上述训练数据分别对车载子系统结点的参数进行预测, 并与真实值对比, 实现车载子系统状态预测.

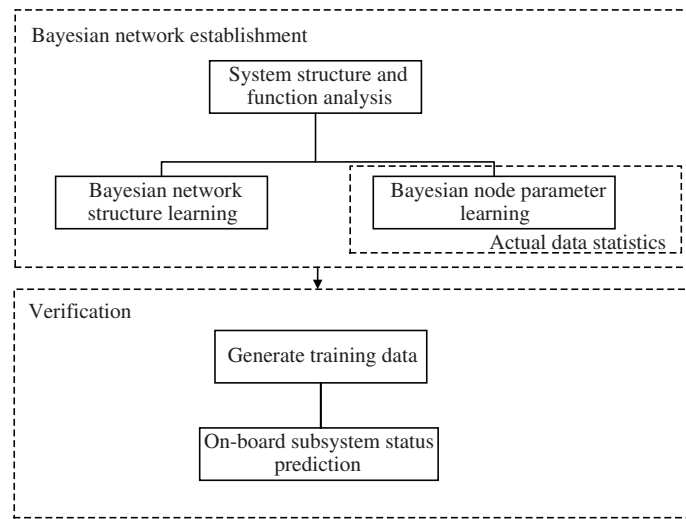


图 2 基于贝叶斯网络的故障预测步骤

Figure 2 BN-based fault prediction

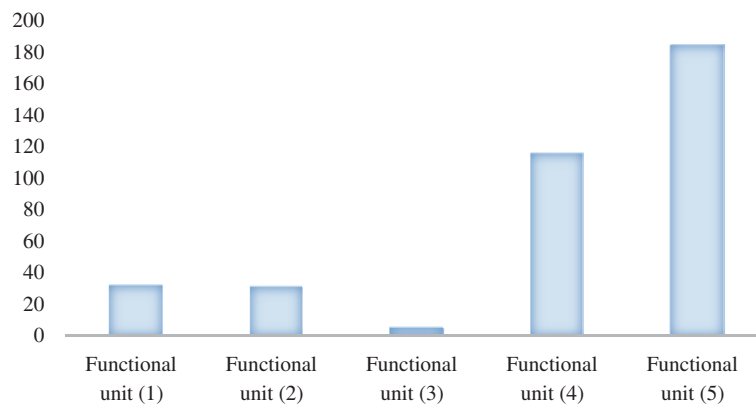


图 3 (网络版彩图) 2015 年车载子系统功能模块故障数

Figure 3 (Color online) Faults produced in 2015 by different modules of on-board subsystem in train control system

### 3.2.1 结构学习

车载子系统功能实现依赖于硬件、软件的正常工作及电缆的正确连接,因此故障分类也可以分为硬件故障、软件故障、电缆连接故障 3 类. 本文不研究软件故障,假设软件版本都能正常执行原有功能. 先不考虑系统的冗余结构,根据功能将车载子系统分为如图 1 所示的 6 个部分. (1) 应答器信息接收功能单元. (2) 测速测距功能单元. (3) 轨道电路信息接收功能单元. (4) 紧急制动输出功能单元. (5) 无线信息接收功能单元. 由于 ATPCU, C2CU, SDP, TSG 单元的硬件部分完全相同,只是软件不同从而执行不同的功能,同时 4 个模块均作为信息最终的接收单元对数据进行分析处理,并通过 PROFIBUS 总线向 JRU 发送数据,因此划分 (6) 为核心功能单元.

将各功能单元看作黑箱,只关注功能单元的供电电压、输入及输出. 图 3 为某铁路局动车组故障数据分析工区 2015 年的全年覆盖表数据统计结果,横坐标表示功能单元,纵坐标表示故障数目,该数据已经剔除一些后期观察误报的故障.

表 1 功能单元信号测试点  
Table 1 Test points in functional units

Functional unit	Test point	Parameter
(1)	$t_1$	BTM working voltage
	$t_2$	Energy carrier signal
	$t_3$	Up-link signal
	$t_4$	Power sense signal
(2)	$t_5$	SDU working voltage
	$t_6$	ODO working voltage
	$t_7$	Radar working voltage
(3)	$t_8$	TCR working voltage
	$t_9$	Antenna frequency
(4)	$t_{10}$	VDX working voltage
	$t_{11}$	Mining voltage
(5)	$t_{12}$	STU-V working voltage
	$t_{13}$	Radio operating frequency

由图 3 可知, 车载子系统功能单元中, 容易发生硬件故障的功能单元为应答器信息接收单元、无线信息接收功能单元、轨道电路信息接收功能单元、测速测距功能单元、紧急制动输入输出单元, 因此本文主要研究以上 5 个功能单元的故障。

表 1 给出了不同功能单元的故障征兆参数样本集信号测试点. 5 个功能单元各自独立工作, 以并行工作的方式向功能单元 (6) 提供所需信息. 分别选取不同功能单元的故障征兆参数, 故障征兆参数分为正常、偏离、失效 3 类, 表 1 共有 13 个信号测试点, 用  $t_1, \dots, t_{13}$  表示. 信号测试点的选取原则为, 在分析不同功能单元结构的基础上, 分别选取可以描述功能单元中不同部件故障的性能参数作为信号测试点.

如图 4 所示, 将车载子系统分为功能单元层和模块层, 建立 13 个信号测试点与车载子系统之间的关系.

本文研究功能单元层次, 将各功能单元看作黑箱, 因此故障模型简化, 如图 5 所示.

因为有一些测试点表征同一功能单元, 功能单元之间相互独立, 但是在功能单元内部, 部件之间结构和功能都互相关联, 例如功能单元 (1) 中, 部件之间结构上互相连接, 如果  $t_1, t_3, t_4$  出现偏离、异常等情况, 均会导致  $t_2$  出现偏离、异常. 换言之, 如果  $t_2$  出现异常, 可以断定功能单元 (1) 功能异常,  $t_1, t_3, t_4$  可以用来进行故障的模块级定位, 因此  $t_2$  可以作为功能单元的故障征兆参数. 其他功能单元同理, 因此保留  $t_2, t_5, t_8, t_{10}, t_{13}$ , 利用  $F_1$  表示车载子系统, 因此贝叶斯网络中包括 6 个结点. 由此可建立车载子系统贝叶斯网络, 功能单元中任何一个功能单元出现故障, 都会导致车载子系统故障.

在离散贝叶斯网络中, 各节点的值均应该为离散值. 功能单元的故障分为完全性失效和功能性故障, 完全性失效是指功能单元已经无法完成正常功能, 功能性故障指性能指标偏离正常范围. 本文研究功能性故障, 从《CRH3, CRH380B, BL 型车 CTCS3-300T 列控车载设备维护说明书 V2.3.0》中获取参数规定值, 建立贝叶斯结点与离散值的对应关系, 如表 2 所示.

由于车载子系统可获得的实际测量数据很少, 因此利用概率抽样工具产生 2, 20, 200, 2000 组数据, 利用贝叶斯结构学习中的 K2 结构学习算法, 建立车载子系统的贝叶斯网络结构. 结果发现, 当样



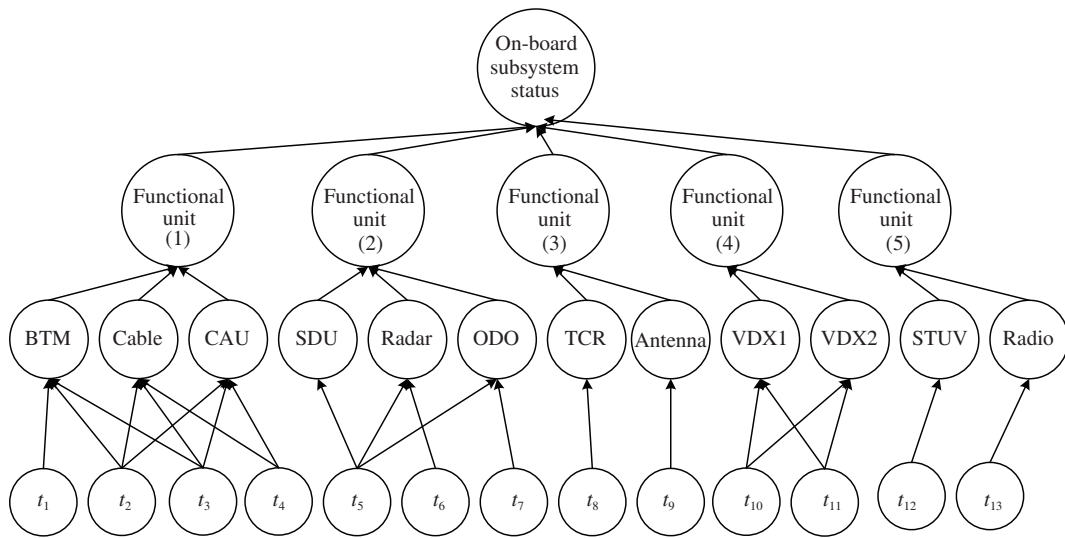


图 4 车载子系统故障模型

Figure 4 Fault model of on-board subsystem

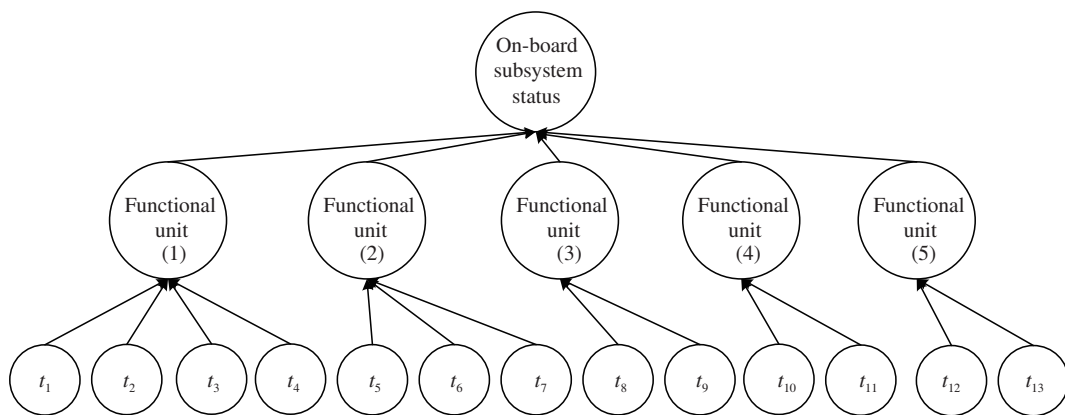


图 5 简化的车载子系统故障模型

Figure 5 Simplified fault model of on-board subsystem

表 2 贝叶斯结点与离散值

Table 2 Bayesian nodes and discrete values

Discrete value	State	$t_2$	$t_5$	$t_8$	$t_{10}$	$t_{13}$	$F_1$
1	Normal	27.095 (MHz)	18 (V)	110 (V)	24~120 (V)	915 (MHz)	1
2	Deviation	[-5%, +5%]	[-5%, +5%]	[-30%, -25%]	[-25%, +30%]	[-5%, +5%]	-
3	Abnormal	< -5%, > 5%	< -5%, > 5%	< -30%, > 25%	< -25%, > 30%	< -5%, > 5%	2

本数据量达到 2000 组时, K2 算法的结构学习效果最好, 当数据量继续增加时, 网络结构与数据量为 2000 组时保持一致. 20 组数据时, 结点之间的关系弧存在错误, 200 组数据时存在少边情况, 结构学习的结果如图 6 所示.

图 6 中, 1, 2, 3, 4, 5, 6 分别对应能量载波信号, SDU 工作电压、TCR 工作电压、VDX 工作电压、

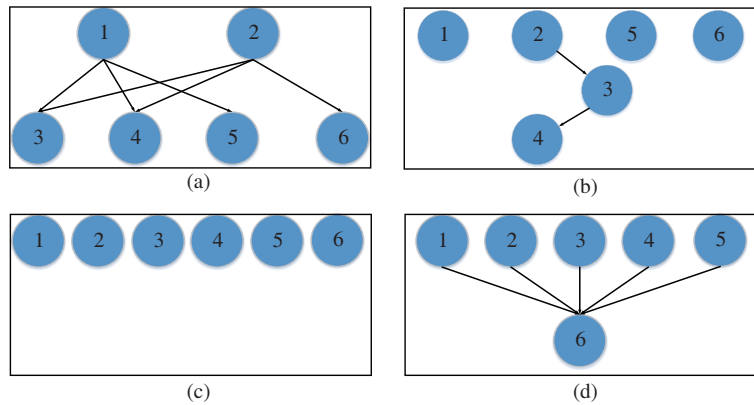


图 6 (网络版彩图) 贝叶斯网络结构学习结果

Figure 6 (Color online) Result of BN structure learning. (a)–(d) Sample = 2, 20, 200, 2000, respectively

天线电台工作频率、车载子系统状态. 基于该学习结构, 在数据完备情况下利用 MLE 算法, 并与神经网络及隐马尔科夫模型方法进行对比, 比较其预测准确率.

### 3.2.2 参数学习

基于现场运营单位获取的实际列控车载子系统运行统计数据, 构造车载子系统条件概率表, 各结点的参数来自真实数据的统计. 首先, 选取 2000 组样本数据, 间隔 100 分为 20 组, 统计参数学习前后信号测试点  $t_2, t_5, t_8, t_{10}, t_{13}$  的参数值, 比较在正常、偏离、异常情况下的参数学习结果, 如图 7 所示.

可以看出, 信号测试点参数在训练数据量为 200 时已经均趋于稳定, 并满足真实值的误差范围, 证明基于该信号测试点的预测具有合理性与可行性.

## 4 验证分析

### 4.1 车载子系统状态预测

基于上述参数信号测试点的贝叶斯网络, 利用概率抽样分别产生 20, 200, 2000, 20000 组离散化后的数据作为训练数据. 随机选择贝叶斯网络初始参数值, 利用训练数据分别对车载子系统结点的参数进行预测, 并与真实值对比, 车载子系统有 243 个参数需要预测, 将预测的参数与真实参数做差值, 误差值越趋近于 0, 预测越准确. 结果如图 8 所示, 可知, 随着样本数的增加, 趋近于 0 的参数点越来越多, 说明预测越准确. 经计算, 样本值为 20, 200, 2000, 20000 组时, 参数预测准确率分别为 5%, 27%, 92%, 96.3%.

### 4.2 对比验证

为了验证贝叶斯网络的预测效果, 利用隐马尔科夫模型和神经网络两种在工业领域应用比较成熟的故障预测方法作为对比验证. 隐马尔科夫模型中, 选取已知参数贝叶斯网络随机抽样产生的离散化后的 2000 组数据作为训练数据, 利用 Baum-Welch 算法<sup>[23]</sup> 训练 HMM 参数, 得到正常、故障两种状态的 HMM 模型, 然后随机产生 200 组观测数据, 利用前向 - 后向算法, 得到其与 HMM 模型的匹配值, 选取最大值即为当前状态, 并与真实状态对比, 实现故障预测.

由结果可知, 正常训练数据迭代到第 23 步时已经收敛, 异常训练数据迭代到第 54 步时达到收敛,

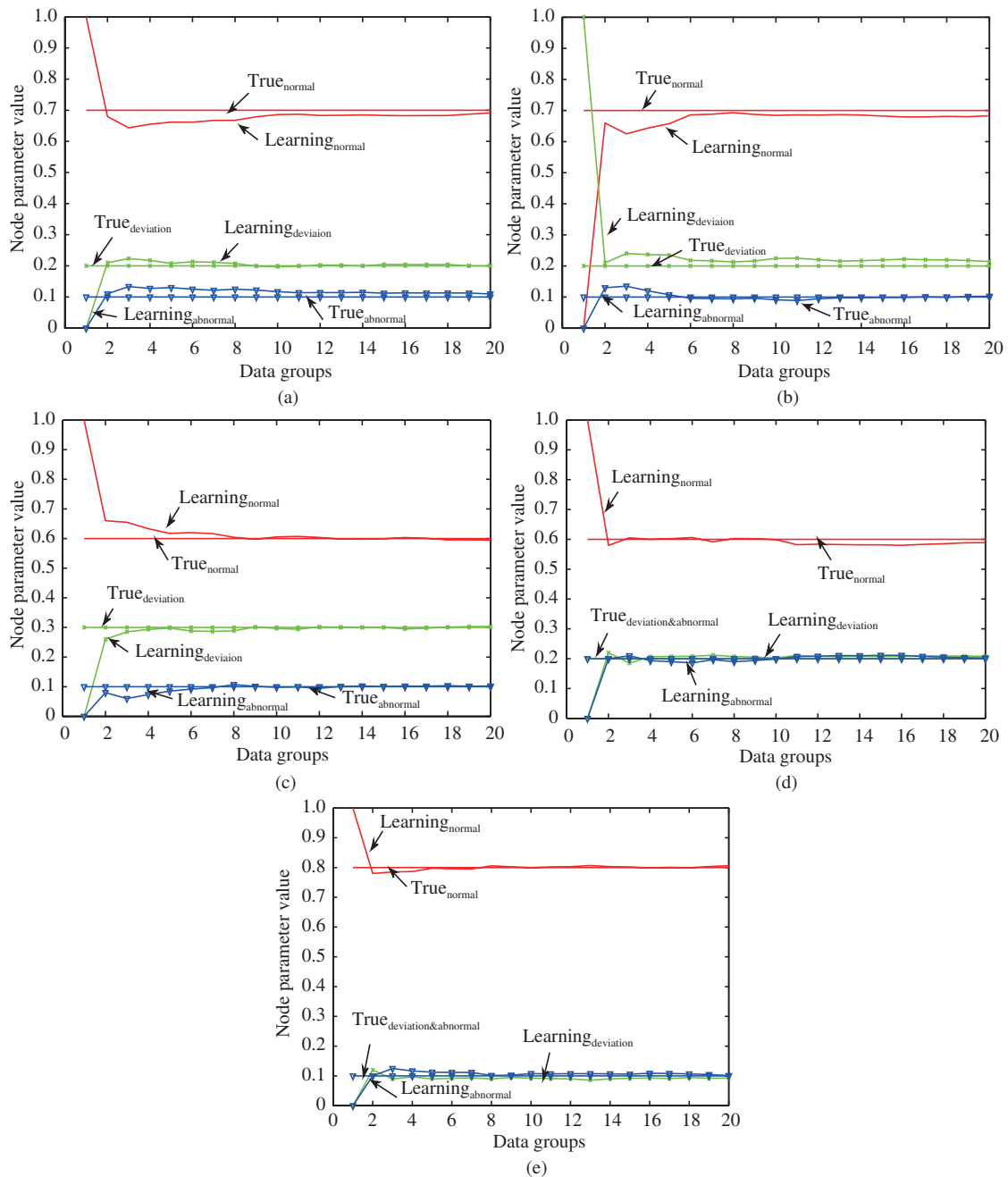


图 7 (网络版彩图) 信号测试点参数学习结果

**Figure 7** (Color online) Results of signal test point parameter learning. (a) Energy carrier signal; (b) TCR working voltage; (c) SDU working voltage; (d) VDX working voltage; (e) radio operating frequency

证明模型构建成功. 然后输入 200 组观测数据, 如图 9 所示. 200 组观测数据中包含 95 组系统正常数据, 105 组异常数据. HMM 预测结果为 77 组车载子系统正常, 123 组车载子系统异常, 故障预测准确率为 80.1%. 采用神经网络实施预测过程中, 利用二进制代码 01, 10 分别表示车载子系统的状态为正常 1, 异常 2. 为了对比 3 种算法的预测正确率, 同样利用贝叶斯网络最终进行参数训练的 2000 组

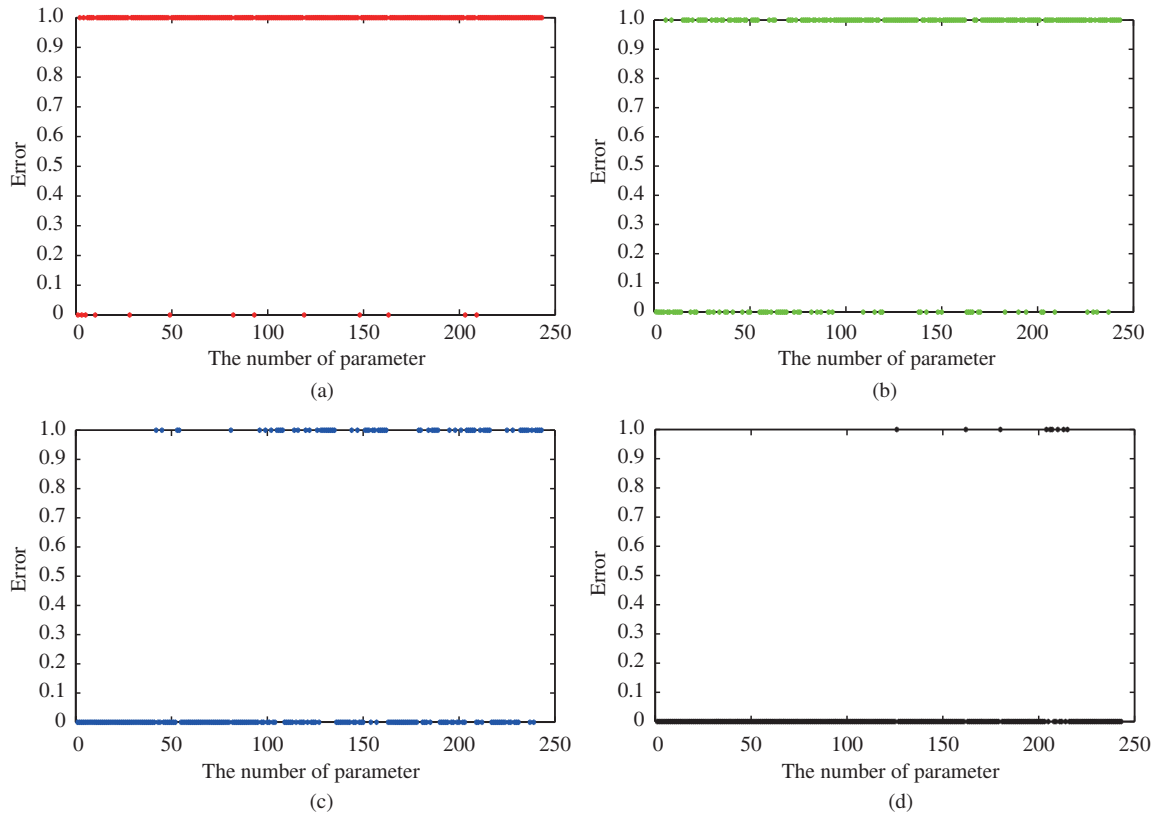


图 8 (网络版彩图) 车载子系统贝叶斯网络故障预测结果

Figure 8 (Color online) Results of fault prediction based on the BN. (a)–(d) Sample = 20, 200, 2000, 20000, respectively

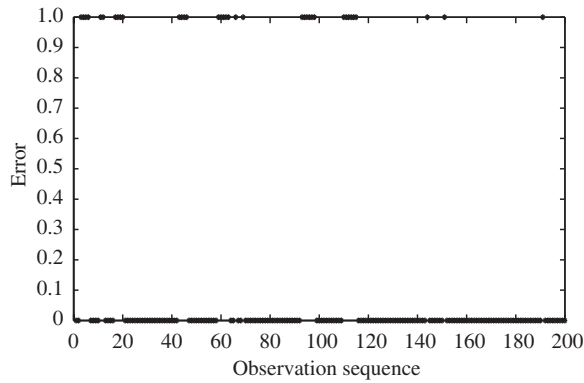


图 9 隐马尔科夫模型预测准确率

Figure 9 Result of fault prediction based on the HMM

数据作为训练数据, 隐马尔科夫方法中的 200 组数据作为观测数据, 利用神经网络算法进行故障预测, 仿真结果如图 10 所示.

由图 10 可知, 绿色为各阶段预测准确率, 红色为预测错误率, 可见基于神经网络算法的故障预测在训练阶段的准确率为 67.8%, 验证阶段的准确率为 71.8%, 测试阶段的准确率为 61.8%, 利用神经网络算法进行故障预测, 最终预测率为 67.4%. 可见, 神经网络算法在处理该问题上故障预测能力较差,

Output	1	666 35.6%	280 15.0%	70.4% 29.6%
	2	322 17.2%	602 32.2%	65.2% 34.8%
		67.4% 32.6%	68.3% 31.7%	67.8% 32.2%
		1	2	Target
(a)				
Output	1	49 44.5%	14 12.7%	77.8% 22.2%
	2	17 15.5%	30 27.3%	63.8% 36.2%
		74.2% 25.8%	68.2% 31.8%	71.8% 28.2%
		1	2	Target
(b)				
Output	1	72 32.7%	41 18.6%	63.7% 36.3%
	2	43 19.5%	64 29.1%	59.8% 40.2%
		62.6% 37.4%	61.0% 39.0%	61.8% 38.2%
		1	2	Target
(c)				
Output	1	787 35.8%	335 15.2%	70.1% 29.9%
	2	382 17.4%	696 31.6%	64.6% 35.4%
		67.3% 32.7%	67.5% 32.5%	67.4% 32.6%
		1	2	Target
(d)				

图 10 (网络版彩图) 神经网络故障预测准确率

**Figure 10** (Color online) Result of fault prediction based on the NN. (a) Training confusion matrix; (b) validation confusion matrix; (c) testing confusion matrix; (d) final confusion matrix

表 3 不同算法预测准确率

Table 3 Predict accuracy under different algorithms

Algorithm	Prediction accuracy (2000 set) (%)
Hidden Markov model	80.1
Neural network	67.4
Bayesian network	92

计算性能与上述两种算法存在一定差距, 结果对比见表 3.

上述仿真验证过程中, 基于真实数据构造贝叶斯的概率条件表, 得到 20000 组训练数据时贝叶斯网络的预测性能达到很好的效果为 96.3%, 为了对比验证贝叶斯网络方法的故障预测效果, 利用贝叶斯网络预测的 2000 组数据训练隐马尔科夫模型及神经网络, 利用随机产生的 200 组数据验证模型的预测性能, 得到的结果是隐马尔科夫模型为 80.1%, 神经网络为 67.4%. 该仿真结果也表明了故障预测方面贝叶斯网络更具优势, 文献 [22] 将贝叶斯网络算法应用于电源系统进行故障预测, 其预测结果是 6000 组数据下为 89.17%, 这证明贝叶斯网络算法应用于车载子系统进行故障预测效果显著. 高铁车载子系统目前仅限于故障诊断, 尚未发现系统级故障预测研究, 本文可用作后续高铁车载子系统故障预测算法的参考.

## 5 结语

本文构建了列控车载子系统级贝叶斯网络框架, 提出了基于贝叶斯网络的系统级故障预测方法,

进行不同数据集下的验证, 并与其他算法进行对比验证. 结果表明, 贝叶斯网络算法相对于隐马尔科夫算法和神经网络算法有明显提升, 实际运用中, 能够发挥贝叶斯网络故障预测算法效益. 后续工作可在横向和纵向维度展开, 横向上, 首先考虑加入准确性验证和先验知识未知或不准确等情况; 其次加入时间因素, 拓展仅根据故障关联关系建立故障模型预测车载子系统状态的局限; 最后考虑车载子系统的冗余部件, 把冗余部件考虑在内重新建立故障模型. 纵向上, 继续定义深层次模块级、板卡级甚至元件级数据需求, 构建多层次贝叶斯故障预测模型. 根据本文所衍生的研究内容对于解决现场车载子系统级故障预测问题具有重要的指导意义.

结合本文所做工作可以明确看出, 采用主动维修、视情维护模式的智能高铁是将来的重要发展方向, 故障预测是其中的关键问题, 还需要在全生命周期性能参数获取、性能退化模型等方面进行深入探索, 并解决设备的剩余寿命预测问题, 从而使主动维护的优势得到更好的发挥和体现, 为新型高速列车关键装备故障预测及健康管理的发展作出重要贡献.

## 参考文献

- 1 Wang H S, Wang J F. Key technology of train control system of 200~350 km/h trains. *J China Railway Soc*, 2009, 31: 107-110 [王化深, 王俊峰. 200~350 km/h 列车运行控制系统关键技术研究. *铁道学报*, 2009, 31: 107-110]
- 2 He R, Ai B, Wang G, et al. High-speed railway communications: from GSM-R to LTE-R. *IEEE Veh Technol Mag*, 2016, 11: 49-58
- 3 Han X, Tang T, Lv J D. Analysis of requirement-errors-caused failure of on-board subsystem of CTCS-3 train control system based on failure logs. *J China Railway Soc*, 2017, 39: 59-70 [韩笑, 唐涛, 吕继东. 基于失效日志的 CTCS-3 级列车车载子系统需求错误所致的失效致因分析. *铁道学报*, 2017, 39: 59-70]
- 4 Pecht M G, Jaai R. A prognostics and health management roadmap for information and electronics-rich systems. *Microelectron Reliabil*, 2010, 50: 317-323
- 5 Ray A, Tangirala S. Stochastic modeling of fatigue crack dynamics for on-line failure prognostics. *IEEE Trans Contr Syst Technol*, 1996, 4: 443-451
- 6 Li Y, Kurfess T R, Liang S Y. Stochastic prognostics for rolling element bearings. *Mech Syst Signal Process*, 2000, 14: 747-762
- 7 Zhang S, Ganesan R. Multivariable trend analysis using neural networks for intelligent diagnostics of rotating machinery. *J Eng Gas Turbines Power*, 1997, 119: 378-384
- 8 Ali J B, Chebel M B, Saidi L, et al. Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. *Mech Syst Signal Process*, 2015, 56: 150-172
- 9 Liang X, Wang H F, Guo J. Bayesian network based fault diagnosis method for on-board equipment of train control system. *J China Railway Soc*, 2017, 39: 93-100 [梁潇, 王海峰, 郭进. 基于贝叶斯网络的列控车载设备故障诊断方法. *铁道学报*, 2017, 39: 93-100]
- 10 Zhao Y, Xu T H, Zhou Y P. Bayesian network based fault diagnosis system for vehicle on-board equipment of high-speed railway. *J China Railway Soc*, 2014, 11: 48-53 [赵阳, 徐田华, 周玉平. 基于贝叶斯网络的高铁信号系统车载设备故障诊断方法的研究. *铁道学报*, 2014, 11: 48-53]
- 11 Zhao Y, Xu T H. Text mining based fault diagnosis for vehicle on-board equipment of high speed railway signal system. *J China Railway Soc*, 2015, 8: 53-59 [赵阳, 徐田华. 基于文本挖掘的高铁信号系统车载设备故障诊断. *铁道学报*, 2015, 8: 53-59]
- 12 Wang Z X. CBR-based fault diagnosis system for vehicle equipment of high-speed railway. Dissertation for Master's Degree. Beijing: Beijing Jiaotong University, 2015 [王振显. 基于案例推理的高铁信号系统车载设备故障诊断. 硕士学位论文. 北京: 北京交通大学, 2015]
- 13 Chen X. The theory and method of fault discovery based on text information extraction for on-board vehicle equipment of high speed railway. Dissertation for Master's Degree. Beijing: Beijing Jiaotong University, 2017 [陈曦. 基于文本信息抽取的高铁车载设备故障发现的理论与方法. 硕士学位论文. 北京: 北京交通大学, 2017]
- 14 Li Y. Application research on fault diagnosis of on-board equipment of train control system based on fuzzy theory and Bayesian network. Dissertation for Master's Degree. Lanzhou: Lanzhou Jiaotong University, 2015 [李逸. 模糊贝叶斯

- 决策在列控车载设备故障诊断中的应用研究. 硕士学位论文. 兰州: 兰州交通大学, 2015]
- 15 Marcot B G. Common quandaries and their practical solutions in Bayesian network modeling. *Ecol Modell*, 2017, 358: 1–9
  - 16 Petitjean F, Buntine W, Webb G I, et al. Accurate parameter estimation for Bayesian network classifiers using hierarchical Dirichlet processes. *Mach Learn*, 2018, 107: 1303–1331
  - 17 Hasan S, Ukkusuri S V. Reconstructing activity location sequences from incomplete check-in data: a semi-Markov continuous-time Bayesian network model. *IEEE Trans Intell Transport Syst*, 2017, 19: 687–698
  - 18 Cai B G, Liu Y, Xie M. A dynamic-Bayesian-network-based fault diagnosis methodology considering transient and intermittent faults. *IEEE Trans Automat Sci Eng*, 2017, 14: 276–285
  - 19 Sun S, Zhang C, Yu G. A Bayesian network approach to traffic flow forecasting. *IEEE Trans Intell Transp Syst*, 2006, 7: 124–132
  - 20 Chen X W, Anantha G, Lin X T. Improving Bayesian network structure learning with mutual information-based node ordering in the K2 algorithm. *IEEE Trans Knowl Data Eng*, 2008, 20: 628–640
  - 21 Bañbura M, Modugno M. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *J Appl Econ*, 2014, 29: 133–160
  - 22 Xu L J, Wang H J, Long B. Fault prediction of complex systems based on Bayesian network. *Syst Eng Electron*, 2008, 30: 780–784 [许丽佳, 王厚军, 龙兵. 基于贝叶斯网络的复杂系统故障预测. *系统工程与电子技术*, 2008, 30: 780–784]
  - 23 Wu G, Qiu Y J, Wang G R. Map matching algorithm based on hidden Markov model and genetic algorithm. *J Northeastern Univ (Nat Sci)*, 2017, 38: 472–475 [吴刚, 邱煜晶, 王国仁. 基于隐马尔可夫模型和遗传算法的地图匹配算法. *东北大学学报 (自然科学版)*, 2017, 38: 472–475]

## Research on the fault prediction method of an on-board subsystem in high-speed train control systems

Yu ZANG<sup>1</sup>, Baigen CAI<sup>1,2\*</sup>, Wei SHANGGUAN<sup>1</sup> & Huashen WANG<sup>1</sup>

1. *School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044, China;*

2. *State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China*

\* Corresponding author. E-mail: bgcai@bjtu.edu.cn

**Abstract** The train control system is the “nerve center” of the high-speed train information system, which is large-scale and comprises various components. The on-board subsystem is the core of the train control system and is key to ensuring traffic safety and improving operating efficiency. Currently, the fault data processing methods of the on-board subsystem remain manual, which primarily realizes the fault detection and location. It is difficult to achieve the fault mechanism level, and fault prediction cannot be realized effectively. In this paper, the on-board subsystem structure and the fault disposal status were analyzed. The existing problems have been summarized, and some concepts and algorithms to predict faults were introduced. Based on the on-board subsystem structure and each module’s performance parameters, the system-level fault prediction model was established. Based on the practical operational data sets, the fault prediction based on the Bayesian network was carried out and verified under 20, 200, 2000 and 20000 sets, respectively. The prediction accuracy was 5%, 27%, 92% and 96.3%, respectively, under the condition of 2000 data sets. The hidden Markov model and neural network-based fault prediction solutions were compared with the proposed method. The results demonstrate the advanced performance of the Bayesian network-based solution in system-level fault prediction.

**Keywords** high-speed train, train control system, on-board subsystem, fault prediction, Bayesian network



Her research interests include system reliability evaluation, prognostics, and health management of complex systems.

**Yu ZANG** was born in 1994. She received the B.S. degree in automation from Beijing Jiaotong University, Beijing, China, in 2014. She was a visiting student at the University of Maryland, College Park, USA, from 2017 to 2018. Currently, she is working toward a Ph.D. degree at the School of Electronic and Information Engineering, Beijing Jiaotong University.



include the train control system, system modeling, simulation and testing, GNSS navigation, and the cooperative vehicle infrastructure system of China (CVIS-C).

**Wei SHANGGUAN** received his B.S, M.S and Ph.D. degrees from Harbin Engineering University, in 2002, 2005, and 2008, respectively. He was a professor at the School of Electronic and Information Engineering, Beijing Jiaotong University from 2008 to 2011. He was an academic visitor to the University College London from 2013 to 2014. Currently, he is a professor at Beijing Jiaotong University. His research interests



His research interests include railway train control systems, intelligent transportation systems, GNSS navigation, multi-sensor fusion, and intelligent traffic control.

**Baigen CAI** received his B.S., M.S., and Ph.D. degrees in traffic information engineering and control from Beijing Jiaotong University in 1987, 1990, and 2010, respectively. Since 1990, he has been a faculty member of the School of Electronic and Information Engineering in Beijing Jiaotong University. He was also a visiting scholar at Ohio State University from 1998 to 1999. Currently, he is a professor at Beijing Jiaotong University.



and control technology for intelligent transportation systems, and electromagnetic compatibility.

**Huashen WANG** received his B.S., M.S., and Ph.D. degrees in traffic information engineering and control from Beijing Jiaotong University in 1982, 1985, and 2009, respectively. Currently, he is the director of the National Rail Transit Safety Assessment Research Center. His research interests include rail transit safety assessment theory and technology, rail transit automation and control, information