

DOI: 10.7652/xjtuxb201501012

高速铁路信号系统异构数据融合和智能维护决策

徐田华¹, 杨连报¹, 胡红利², 王小鑫²

(1. 北京交通大学轨道交通控制与安全国家重点实验室, 100044, 北京;
2. 西安交通大学电力设备电气绝缘国家重点实验室, 710049, 西安)

摘要: 针对高速铁路智能维护决策中的信息多源异构问题,提出了高铁信号系统异构数据融合和智能维护决策架构。通过本地数据库资源描述框架(resource description framework (schema), RDF(S))到全局 RDF(S)的转换和基于 RDF(S)图的本体合并,实现了多源异构信息的融合。利用适合缺失数据的结构期望最大化(SEM)算法,结合专家知识,构建了高铁信号系统的智能维护贝叶斯网络(BN)决策模型。最后,利用高铁武广线 2011~2012 年的监控数据,分别对基于 RDF(S)图的全局 RDF(S)合并算法性能和故障诊断结果的准确性进行了分析对比,实验结果表明所提出的本体融合算法具有多项式级的计算复杂度,同时融合专家知识和 SEM 算法的智能维护 BN 决策模型的一级故障诊断准确率为 92.4%。因此,所提出的异构数据融合和智能维护架构可以有效提高高铁信号系统维护决策的准确性和有效性。

关键词: 高速铁路;信号系统;异构数据;本体融合

中图分类号: TM934 **文献标志码:** A **文章编号:** 0253-987X(2015)01-0072-07

Heterogeneous Data Fusion and Intelligent Maintenance Decision for High Speed Railway Signaling Systems

XU Tianhua¹, YANG Lianbao¹, HU Hongli², WANG Xiaoxin²

(1. State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China;
2. State Key Laboratory of Electrical Insulation and Power Equipment, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: A framework of integrated heterogeneous data and intelligent maintenance decision was proposed aiming at the multi-source heterogeneous data in intelligent maintenance decision for railway signaling systems. By means of transformation and fusion from local resource description framework (schema)(RDF(S)) to global RDF(S), the fusion of heterogeneous data was realized. In addition, a Bayesian net (BN) based intelligent maintenance decision model was constructed by combing the structural expectation maximum (SEM) algorithm for missing data with the expert knowledge. The correctness and efficiency of the proposed framework and the RDF(S) fusion algorithm were verified by the maintenance data from Wuhan-Guangzhou high-speed railway signaling systems in 2011-2012. The experimental results show that the computational complexity of the proposed ontology fusion algorithm is polynomial, and the average accuracy of the fault diagnosis for the first level reaches 92.4%. Therefore, the proposed framework may improve the accuracy and efficiency of the intelligent maintenance decision of high-speed railway systems.

Keywords: high-speed railway; signaling system; heterogeneous data; ontology fusion

收稿日期: 2014-07-21。 作者简介: 徐田华(1971—),男,副教授。 基金项目: 国家自然科学基金资助项目(51305021); 铁路总公司重点资助项目(2013X015-B);轨道交通控制与安全国家重点实验室自主研究课题(RCS2012ZT005)。
网络出版时间: 2014-10-15 网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20141015.1752.006.html>

近年来,我国高速铁路正处于快速发展时期,逐步进入网络化运营阶段,在社会经济发展中起到了重要作用。高铁信号系统是高速铁路的大脑和神经中枢,对保障高速铁路安全、高效的运营起关键作用。

与传统的信号系统相比,高速铁路信号系统不断采用新技术和新设备。设备的先进性一方面保证了系统运行的高效性,同时也增加了故障的隐蔽性和突发性,为设备的故障定位和智能维护设置了障碍。另一方面,支撑高铁信号系统可靠、高效运营的运营维护记录和监控数据仍然存在多源异构和信息非完整性等突出问题。造成这些问题的主要原因是维护记录信息来源复杂,从司机、调度员、铁路专家和维护人员的记录日志到微机监测系统、司法记录仪(JRU)和列车动态监控系统(DMS)的设备记录信息等组成了多源的维护记录信息。更为重要的是,多源的维修记录信息存在概念术语不统一、缺乏统一的格式和标准,维护数据语义不一致甚至冲突、维护记录信息不完整等问题,因而无法形成一种有效的信息共享机制,最终形成了大量的信息孤岛,不能有效地实现知识的共享和互操作,这将影响高铁的智能运营维护决策和行车效率。

因此,如何构建高铁信号系统规范统一的知识体系,在多源异构数据融合的基础上实现高铁信号系统的智能维护决策,以适应高铁信号设备技术现代化发展需要,成为亟待解决的问题。

目前,国内外学者在异构信息的融合和智能维护决策等领域开展了卓有成效的探索。文献[1-3]应用本体论解决了欧洲铁路领域内由于运营商和开发商不同而造成维护信息不统一的问题,实现了不同国家、不同格式铁路基础设施、旅客信息和设备管理信息的互操作和共享;文献[4]主要应用本体论对我国高铁领域中各专业领域本体的融合方法进行了研究;文献[5-6]应用本体论解决了铁路选线时数据不统一、共享率低的问题;欧洲 EuRoMain FP5-IST Project 提出了通过建立维护数据的标准化文档,实现不同来源的维护数据的共享,进而搭建支持欧洲铁路维护系统的解决方案^[7]。但是,该项目还是建立在 XML、XSLT 等数据格式的基础上。由于 XML 缺乏语义描述,因此很难实现在语义层次上的机器理解 and 处理。欧洲的 InterGRail 项目提出铁路领域的信息共享,使得维护决策优化得以实现^[8]。在该项目参与者中,英国伯明翰大学的 Cliver 提出了使用本体论解决大规模轨道交通维护数据融合方法。在故障诊断和决策支持方面,Ferreiro 和 Vong 等分

别给出了基于案例的推理和基于贝叶斯网络(BN)的运输系统故障诊断和决策维护^[9-10],但其研究并未涉及异构数据的融合和语义冲突的消除。

上述研究对推动轨道交通领域的异构数据融合和智能决策支持的发展起着重要作用,但针对高速铁路信号数据记录存在的多源异构、非确定和不完整数据等突出问题,没有提出统一的异构数据融合和智能维护决策架构。同时,对于异构数据的属性抽取、转换和融合缺乏严格的形式化定义,这将对后续的严格规范推理和智能维护决策造成不利影响。

针对上述问题,本文提出了高铁信号系统异构数据融合和智能维护决策的统一架构,给出了局部关系数据库到局部本体数据库资源描述框架(resource description framework (schema),RDF(S))转换的形式化定义,提出了匹配数据记录时间和地点模式的本体合并方法,并给出了适合非确定和不完整数据的 BN 维护决策方法,最后利用武广高铁维护记录数据验证了所提架构和算法的可行性和正确性。

1 高铁信号系统异构数据融合和智能维护架构

高铁信号系统异构数据融合和智能维护架构参见图 1,整个架构分为 3 个层次。

最底层是实际的物理数据库,分布在不同的电务段、车站、分局等数据库中,内容包括司机、调度员、铁路专家和维护人员的记录和信号监控设备的监控信息,如微机监测、JRU 记录信息、AElog 日志记录、Abis 协议记录和 DMS 等。其突出的特点是:记录缺乏统一的概念、术语及关系的准确定义,存在概念歧义和语义冲突等问题,阻碍了高铁信号系统维护信息的逻辑关系的建立,如故障模式、故障原因、故障征兆、测试手段、维护手段等。同时,维护数据呈现明显的不完整性和非确定性。高铁信号系统维护数据来源多源化,分布在运营维护人员和不同的监控设备中,加之专家经验和历史数据,其完整性、一致性和确定性都难以得到保证。

中间层为本体合并层。首先构建对应分布式数据库的局部 RDF(S)本体,抽取分布式数据库的概念、属性及其关联关系。接着借助于 RDF(S)图的合成思路,定义关系数据库到 RDF(S)的形式化转换方法,为下一步本体转化的完备性和覆盖性等性能验证提供依据。在模型转换的支持下,实现局部本体到全局本体的合成,同时建立全局本体到局部本体概念、属性的映射关系,为后续的分布式数据库

数据检索和查询提供依据。

框架的最上层对应智能决策层。利用全局本体和局部本体的映射关系,实现故障原因、故障征兆、故障现象和故障类别、设备结构及其维护措施的数据训练样本的抽取,采用结构和参数学习,结合专家知识,形成信号系统维护决策 BN 模型,构造故障诊断和维护的智能维护支持系统。

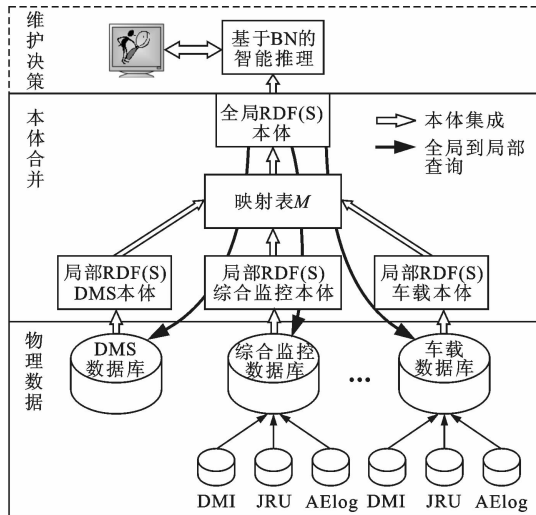


图1 高铁信号系统异构数据融合和智能维护架构

2 高铁信号系统异构维护信息融合

通过异构信息的融合,实现高铁信号系统物理数据到全局本体的转换,并在全局本体的指导下,实现概念数据的统一、语义冲突的消除和多源信息的统一集成。其主要步骤包括本地关系数据库到局部RDF(S)的转换,通过基于图论的方法实现局部RDF(S)本体到全局本体RDF(S)的融合。

2.1 本地关系数据库到局部RDF(S)的转换

构建局部本体到全局本体的形式化定义是本体转换的严格推理和逻辑验证的保证,同时也是实现本体合并的语义一致性、完备性证明中的重要步骤。

下面首先给出关系数据库、本体图和关系数据库的形式化定义。

定义1 一个关系数据库模式 $D_{BS} = (L, P_{Key}, F_{Key}, A_{tr})$ 是一个四元组,其中:

(1) $L = R \cup D$ 为名称集合,这里 $R = E_T \cup R_T$ 为表名集合, E_T 和 R_T 分别为实体表和关系表, D 为数据类型名称的集合,表示关系数据库中预定义的类型名称;

(2) A_{tr} 为属性集合, $\forall T \in R, T$ 有一个非空的列集合 $A_{tr}(T)$,且 $\forall A \in A_{tr}(T)$ 有一个预先定义的数据类型 $T_{type}(A) \in D$;

(3) $\forall T \in R, T$ 有且仅有一个主键 $P_{Key}(T)$ (非空列集合), $P_{Key}(T) \in A_{tr}(T)$ (单键,并且 $T \in E_T$),或者 $P_{Key}(T) \in A_{tr}$ (复合键,并且 $T \in R_T$);

(4) $\forall T \in R, T$ 可能有一个或多个引用其他实体表 $G (G \in E_T)$ 主键的外键 $F_{Key}(T, G)$ 的列的集合 $F_{Key}(T, G) \in A_{tr}(T)$ 。

定义2 建立在字母表 A_R 上的RDF模式图 R 是一个有向标记树 $R = (V, E, \lambda)$,这里 V 是带有标记的节点集合,节点的标记包括类 C 、特性 P 和数据类型 $L, E = \{(v_i, v_j) \mid v_i, v_j \in V\}$ 为带有标记的有向边集合,为标记函数 $\lambda: V \cup E \rightarrow A_R$,满足:

(1) $\forall v \in P, \text{domain}(v) \in C, \text{range}(v) \in C \cup L$,
 $\lambda(v, \text{domain}(v)) = \text{"rdfs:domain"}$,
 $\lambda(v, \text{range}(v)) = \text{"rdfs:range"}$;

(2) $\forall e \in (v_i, v_j) \in E$,在 $v_i, v_j \in C$ 情况下, $\lambda(e) = \text{"rdfs:subPropertyOf"}$,或者在 $v_i, v_j \in P$ 的情况下, $\lambda(e) = \text{"rdfs:subPropertyOf"}$ 。

下面给出关系数据库到局部本体RDF(S)的映射规则。

规则1 $\forall T \in E_R \mapsto \{V_T \mid \text{label}(V_T) = \text{ns}:T_{Name}\}$ 。

对于每一个实体类,映射为图中的一个节点,对应的标记为命名空间前缀和表的名称组合。

规则2 $\forall T, G \in E_R \mapsto \{V_T \mid \text{label}(V_T) = \text{ns}:T_{Name}\}$

$\wedge \{V_G \mid \text{label}(V_G) = \text{ns}:G_{Name}\} \wedge \lambda(V_T, V_G) = \text{rdfs:subClass}$,其中 T 是 G 的子表。关系数据库中的子表映射为RDF(S)图中的类节点,并且其有向边为 rdfs:subClass 关系。

规则3 $\forall T \in E_R, A \in A_{tr}(T)/F_{Key}(T) \mapsto (V_T, T_{type}(A)) \in E \wedge \lambda(V_T, T_{type}(A)) = \text{ns}:A_{Name}$ 。数据表 T 中的非主键属性 A 映射为RDF(S)边,其定义域为表 T 映射后的节点 V_T ,值域为其数据类型节点。

规则4 $T, G \in E_R, A \in F_{Key}(T, G) \mapsto (V_T, V_G) \in E \wedge \lambda(V_T, V_G) = \text{ns}:A_{Name}$ 。数据表 T 中的外键 A 关联数据表 G 映射为RDF(S)边,其定义域和值域分别为表 T 和 G 映射后的节点 V_T, V_G ,对应的标记为命名空间前缀和属性名称的组合。

2.2 局部RDF(S)本体到全局本体RDF(S)的融合

分布式数据库的数据字段由于在设计初期没有考虑数据融合问题,导致局部RDF(S)存在类名和属性命名冲突。例如,在信号系统车载维护数据记录中,频繁出现“ATP故障”、“ATP故障停车”、“ATPCU故障”、“ATPCU相关”等数据来刻画ATPCU这一概念。同时,不同的监控设备(如

JRU、AElog 和 DMS 等)记录 ATPCU 的数据字段存在不一致的现象,这直接导致 RDF(S)中的术语不统一甚至冲突。

本文针对上述问题,借鉴文献[11-12]提出的 RDF(S)合并思路,利用霍尔(Horn)逻辑来刻画概念和属性间的等价、包含或者冲突关系。例如,将表达同一概念的“ATP 故障”、“ATP 故障停车”、“ATPCU 故障”之间的关系用下述 Horn 等价逻辑刻画:

ATP 故障 $\Leftrightarrow$ ATP 故障停车 $\Leftrightarrow$ ATPCU 故障

同时考虑到高铁信号系统数据记录的时空特征,提出了改进的 RDF(S)合成算法。算法中考虑了信号系统维护记录中的时间地点匹配原则,即 RDF(S)图的合并三元组必须满足时空匹配,也就是在相同的时间、地点(如线路站点、电务段)的前提下进行融合。具体的算法如表 1 所示。

所提算法利用 Horn 逻辑刻画异构数据源的概念和属性的关系(参见算法第 3 行),其中的 timeAddrMode 表示时间地点匹配模式,保证所有融合的节点需要满足相同的时空约束。算法第 10~13 行根据 Horn 逻辑描述的不同本体中概念的等价关系,计算强连通图 SCC,完成等价节点的合并,实现包括同物异名节点等语义冲突的消除,第 14~15 行消除 RDF(S)合并过程中的冗余边。本体融合算法伪代码见表 1。

表 1 本体融合算法伪代码

算法. OntologyMerge($O1,O2,H,NC,timeAddrMode$)
输入:本体 $O1,O2$,Horn 逻辑表达式,负定条件 NC,时间地点模式 timeAddrMode。
输出: $G=(C,E)$,其中 C,E 分别为合并后的节点和边。
1. $C \leftarrow V_1 \cup V_2$
2. $E \leftarrow E_1 \cup E_2$
3. for all $r_1 \wedge r_2 \wedge \cdots \wedge r_n \wedge timeAddrMode \rightarrow r$
4. $V \leftarrow V \cup \{r_1,r_2,\cdots,r_n,timeAddrMode\}$
5. $E \leftarrow E \cup \{(r_1,r_2,\cdots,r_n,timeAddrMode),r\}$
6. end
7. for all $n \in V$
8. $E \leftarrow E \cup \{(r_1,r_2,\cdots,r_n,timeAddrMode),n\}$
9. end
10. for all p in Property
11. if p is transitive
12. collapseSCC(C,E)
13. end
14. EliminateRedundantEdge(C,E)
15. end
16. return G

与经典的基于闭包算法的本体合并相比,算法的计算复杂度得到有效降低。经典的基于 RDF(S)的闭包算法主要是根据 RDF(S)本体的推理规则如类的传递性、自反性等将 RDF(S)本体图的隐含知识表示出来,闭包算法要对 RDF(S)本体每个节点和属性进行推理计算,之后再进行本体相似度的计算、推理及剪枝,主要的时间复杂度为 $T((v_i + e_i)^3)$,其中 v_i,e_i 表示第 i 个本体的节点数和边数。本文给出的改进算法不需要去构建 RDF(S)图的闭包,而是通过人为输入约束集合 H 负定条件 NC 以及时空模式 timeAddrMode 实现融合,时间复杂度为 $T(\sum_{i=1}^2 (v_i + e_i)v_k)$, v_k 表示霍尔约束的条目数。

3 基于全局 RDF(S)检索数据的 BN 诊断模型

3.1 故障数据训练矢量的获取

根据生成的全局 RDF(S)本体和局部 RDF(S)本体的映射关系,构建包含故障现象和原因的矢量样本,其对应同一时刻、同一地点和设备故障的所有数据库的监控设备记录、人员记录以及维修日志。在信号系统的故障诊断中,矢量空间中的变量定义如下:DMIMasterFail 表示 DMI 与主机通信中断;DMIEquipFail 表示 DMI 报地面设备故障;StartRBC 表示启动 RBC;MT 表示车载电台;SafeTrans 表示安全传输模块;Stop 表示停车。

需要注意的是:由于记录设备或人员的限制,导致数据库中的故障存在记录缺失的现状(见表 2),这对后续的完全数据的结构和参数学习设置了极大的障碍。

表 2 高铁信号设备故障记录表单

故障 标号	DMIMasterFail	SafeTrans	MT	Start RBC	...	Stop
1	是	否	否	否	...	是
2	否	缺失	是	是	...	是
3	缺失	是	否	缺失	...	否
⋮	⋮	⋮	⋮	⋮	⋮	⋮
N	是	缺失	是	否	...	是

注:“缺失”表示没有对应的数据记录。

3.2 缺失数据的智能维护 BN 模型

鉴于故障数据记录缺失的现状和目前存在大量的高铁信号系统故障传播机理先验知识,本文中采用以专家经验 BN 模型为主,利用结构和参数学习

挖掘隐含 BN 关系为辅的方式构建智能维护 BN 模型。下面给出适应缺失数据下的结构期望最大化 (SEM) 算法^[13-14] 的描述。

为了适应训练样本空间 $U^N = \{X_i^j; 1 \leq i \leq n, 1 \leq j \leq N\}$ (X_i^j 表示 j 个样本数据中变量 X_i 的赋值) 中的不可观测变量 $H = U^N - O$ (O 为可观测变量), 在模型选择的评分标准中使用期望代替完全数据的统计参数, 即

$$Q(G; \theta; G^*, \theta^*) = \sum_i Q_i(\prod X_i, \theta_i; G^*, \theta^*) = \sum_i (\sum_{x_i, \prod x_i} (E[N(x_i, \prod x_i) | O] \log(\theta_{x_i | \prod x_i}) - \frac{\log N}{2} \#(X_i, \prod X_i))) \quad (1)$$

式中: X_i 为网络中节点变量, $\prod X_i$ 为其父变量, $\#(X_i, \prod X_i)$ 为贝叶斯网络所需要的参数个数。这里 BN 定义为 $B = \langle G, \theta \rangle$, G 为网络结构, θ 为网络参数。其基本思想: 从某一初始贝叶斯网络 $B_0 = \langle G_0, \theta_0 \rangle$ 出发开始迭代, 在进行了 t 次迭代得到当前最佳网络 $B_t = \langle G_t, \theta_t \rangle$ 后, 第 $t+1$ 次迭代由以下两步骤组成。

(1) 基于当前最佳网络 $B_t = \langle G_t, \theta_t \rangle$ 对数据集 D 利用 EM 算法进行补全, 使之完整并得到完整数据集 D_t 。

(2) 基于数据集 D 对模型及参数进一步优化, 得到 $B_t = \langle G_t, \theta_t \rangle$, 其中初始网络的选择一般是: 隐变量和所有的已观测变量都是相连的, 并且是它们的父节点。如果一个网络结构中有很多隐变量, 那么就从中随机地选择边, 以确保不会超过父节点集的大小, 然后按照上述思想进行迭代。

参照全局 RDF(S) 和局部 RDF(S) 的映射关系, 提取武广高铁信号系统车载设备、地面监控设备、DMS 设备等多源异构信息数据库中数据用于 BN 的结构和参数学习, 由 SEM 算法得到的 BN 结果如图 2 所示。

由于数据结构量较小, 所以结构学习的结果并不十分理想, 但是在有限的学习过程中, 我们也可以从中提取到有用的信息。例如: 18、19、20 三个节点的关联 (参见图 2 中的粗箭头线) 在基于专家知识的 BN 故障模型中是缺失的, 但是在 SEM 结构学习中挖掘出了专家知识忽视的隐含关联关系。

图 2 中 1~34 节点的具体含义如表 3 所示。按照故障传播机理, 完整的故障因果关系包括故障现象、故障征兆、故障传播环节, 这里可以根据

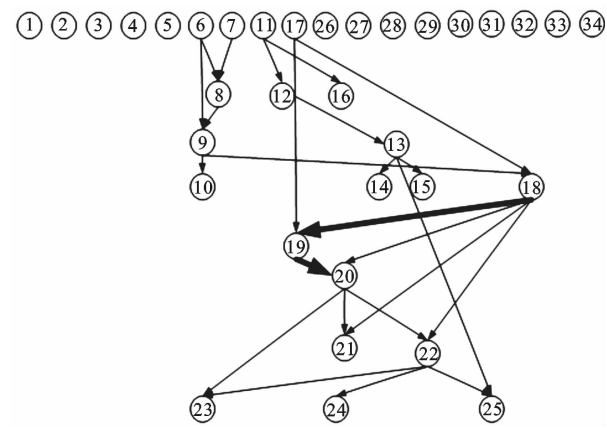


图 2 基于 SEM 算法的 BN 故障模型

表 3 BN 中各节点所代表的含义

节点	故障类型	节点	故障类型
1	BSA	18	BP(旁路)继电器
2	全零应答器	19	制动测试不成功
3	BTM 相关	20	列车接口
4	A-kernel 模式转换无效	21	DMI 与主机通信中断
5	V_MRSP Mismatch	22	安全软件故障
6	MA 起点 AB 不一致	23	DMI 地面设备
7	ATPCU 失败	24	车载底层拆链
8	信息中断	25	MT 模块
9	启动失败	26	RBC、ISDN
10	短链接出窗	27	GCD 和 COMC
11	地面设备	28	无线超时
12	ATPCU 相关	29	启动呼叫 RBC 失败
13	ATP 故障	30	车地通信
14	BTM 故障	31	无法连接 RBC
15	应答器报文错误	32	制动测试失败
16	冗余制动继电器(RBR)	33	C3 转 C2
17	EBR2 继电器	34	停车

高铁信号系统的子系统、设备和具体部件分为多级故障, 即系统级故障、子系统级故障、器件级故障和最终的故障源。

大量的故障分析工具如故障树、FEMA 等方法为建立信号系统的故障因果关系的 BN 提供了支持。鉴于篇幅的限制, 本文不再赘述。

3.3 融合先验知识的故障诊断 BN 结构

在实际应用中, 考虑到先验知识构建的 BN 的稳定性和可靠性, 确定以其为主、结构学习为辅的策略。利用结构学习中发掘的隐含联系对先验知识构建的 BN 进行补充, 当出现二者之间的冲突联系, 应该以先验知识的 BN 为主。同时, 充分考虑结构学习的“过拟合”问题, 即结构过于复杂, 对训练样本达

到非常高的精度,但样本产生的噪声会导致网络的泛化能力降低。

图 3 为执行上述融合原则所得到的最优 BN 结构,其中细线表示基于专家知识的节点关联关系,粗线表示 SEM 算法挖掘的隐含关系,因此融合后的模型综合考虑了先验的故障传播机理和经过数据挖掘后先验知识无法揭示的隐含故障传播关系。

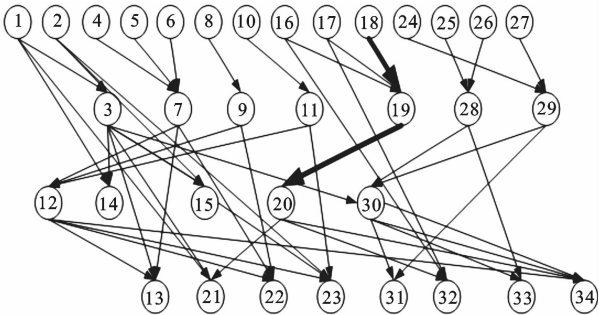


图 3 融合先验知识和 SEM 算法的 BN 故障诊断模型

4 实验结果

本文所使用的故障数据来自武广线 2011 年 8 月至 2012 年 12 月的信号系统监控数据。分别对基于 RDF(S)图的全局 RDF(S)的合并算法性能和故障诊断结果的准确性进行了分析对比。

4.1 RDF(S)合并算法性能比较

为了说明本文提出的改进的 RDF(S)合并算法的有效性,将提出的算法与经典的闭包本体融合算法进行比较。图 4 给出了计算时间的比较结果。图 4 中计算了本体节点在 20~980 范围内的计算时间,可以看出,在较小的节点数量下,两算法具有基本相似的计算时间,但随着节点数量的增加,基于闭包算法的计算时间远高于提出的改进算法。这主要是由于前者的计算复杂度为指数级,而提出的改进本体融合算法的时间复杂度为多项式。

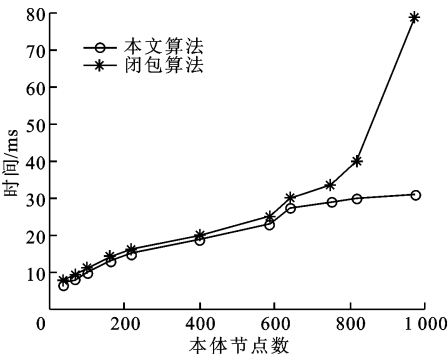


图 4 RDF(S)本体融合运行时间

4.2 基于 BN 的智能维护决策模型的性能

根据全局 RDF(S)和局部 RDF(S)的映射关系,提取武广高铁信号系统车载设备、地面监控设备、DMS 设备等多源异构信息数据库中 608 条数据用于 BN 的结构和参数学习,剩余 100 条用来验证 BN 的正确性。图 5 为故障诊断系统测试结果。

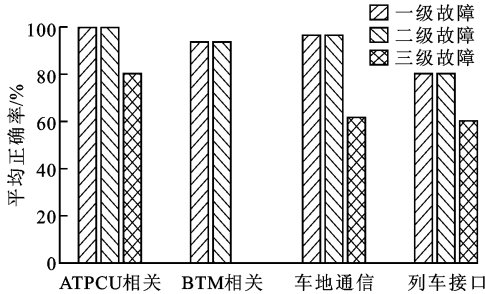


图 5 故障诊断系统测试结果

由图 5 可以看出,一级故障和二级故障诊断的平均正确率均为 92.4%,三级故障为 67.2%。

在故障诊断过程中,三级故障中误诊的概率比较大,主要原因是监控数据缺少大量的细节证据来支持第三级的故障诊断。如果结合维修检测证据和日常巡检和第三级故障相关的证据,则诊断的准确性会得到较大的提升。

5 结 论

针对高速铁路信号系统的多源异构监控信息,本文建立了融合异构设备监控数据的智能维护决策架构。在架构中的 RDF(S)本体合并部分,根据监控数据的时空配准原则,提出了改进的 RDF(S)融合算法,算法具有多项式的计算复杂度。针对高铁信号监控数据的缺失现状,在智能维护 BN 模型部分,结合结构化期望最大 BN 学习算法和专家知识,提出优化的 BN 维护决策模型。利用实际的高铁监控数据对决策模型的测试,证明了其对于第一级和第二级故障诊断的准确性。对于更为细粒度的判定准确性,需要进一步加强相关证据的抽取和测试,这将是下一步需要解决的关键问题。

参考文献:

[1] ABANDA H, NG'OMBE A, TAH J, et al. An ontology-driven decision support system for land delivery in Zambia [J]. Expert Systems with Applications, 2011, 38(9): 10896-10905.
[2] VERSTICHEL S, ONGENAE F, LOEVE L, et al. Efficient data integration in the railway domain through an ontology-based methodology [J]. Trans-

- portation Research: Part C, 2011, 19: 617-643.
- [3] SAA R, GARCIA A, GOMEZ C, et al. An ontology-driven decision support system for high-performance and cost-optimized design of complex railway portal frames [J]. Expert Systems with Applications, 2012, 39: 8784-8792.
 - [4] LIU Z, HUANG L, XU D. Research on semantic retrieval system for high-speed railway knowledge based on ontology [C]//International Colloquium on Computing, Communication, Control, and Management. Piscataway, USA: IEEE, 2008: 303-307.
 - [5] 韩春华, 易思蓉, 吕希奎. 基于 GIS 的铁路选线智能环境及领域本体建模方法 [J]. 中国铁道科学, 2006, 27(6): 84-88.
HAN Chunhua, YI Sirong, LV Xikui. GIS based railway location intelligent environment and domain ontology modeling method [J]. China Railway Science, 2006, 27(6): 84-90.
 - [6] CHOUGULE R, RAJPATHAK D, BANDYOPADHYAY P. An integrated framework for effective service and repair in the automotive domain: an application of association mining and case-based-reasoning [J]. Computers in Industry, 2011, 62: 742-754.
 - [7] European Commission. European railway open maintenance system. [EB/OL]. [2014-06-30]. http://www.transport-research.info/web/projects/project_details.cfm?id=15048.
 - [8] European Commission. Intelligent integration of railway systems. [DB/OL]. [2014-06-30]. <http://www.integrail.info/>.
 - [9] FERREIRO S, ARNAIZ A, SIERRA B, et al. Application of Bayesian networks in prognostics for a new Integrated Vehicle Health Management concept [J]. Expert Systems with Applications, 2012, 39(7): 6402-6418.
 - [10] VONG C, WONG P, IP W. Case-based expert system using wavelet packet transform and kernel-based feature manipulation for engine ignition system [J]. Engineering Applications of Artificial Intelligence, 2011, 24(7): 1281-1294.
 - [11] UDREA O, DENG Y, RUCKHAUS E, et al. A graph theoretical foundation for integrating RDF ontologies [C]//Proceedings of the Twentieth National Conference on Artificial Intelligence. Pittsburgh, Pennsylvania, USA: AAAI, 2005: 1442-1447.
 - [12] HOU X, OON S K, NEE A Y C, et al. A graph-based approach for automatic construction of domain ontology [J]. Expert Systems with Applications, 2011, 38: 11958-11975.
 - [13] FRIEDMAN N. Learning belief networks in the presence of missing values and hidden variables [C]//Proceedings of the Fourteenth International Conference on Machine Learning. San Francisco, USA: ADM, 1997: 125-133.
 - [14] NIELSEN J D, RUMÍ R, SALMERÓN A. Structural-EM for learning PDG models from incomplete data [J]. International Journal of Approximate Reasoning, 2010, 51: 515-530.
- [本刊相关文献链接]**
- 杜晓林, 苏涛, 朱文涛, 等. 一种多输入多输出雷达相位编码信号的设计方法. 2014, 48(12): 93-99. [doi:10.7652/xjtuxb201412015]
- 尚佳栋, 王祖林, 周丽娜, 等. 采用随机共振增强的混合扩频信号跳频参数估计. 2014, 48(10): 42-48. [doi:10.7652/xjtuxb201410007]
- 李实锋, 王玉林, 华宇, 等. 罗兰拟 C 信号的抗干扰快速检测方法. 2013, 47(10): 91-96. [doi:10.7652/xjtuxb201310016]
- 陈晓光, 徐光华, 梁霖, 等. 一种电流弱信号特征提取方法. 2013, 47(9): 83-87. [doi:10.7652/xjtuxb201309014]
- 刁瑞朋, 孟庆丰. 一种衰减信号加窗频域插值算法. 2013, 47(7): 85-90. [doi:10.7652/xjtuxb201307016]
- 齐巍, 陆明泉, 李强, 等. 高动态测量系统中的自适应信号成形滤波器设计. 2013, 47(6): 124-129. [doi:10.7652/xjtuxb201306021]
- 谢桢, 付立军, 肖飞, 等. 三电平直流变换器全模式小信号建模和控制设计. 2013, 47(6): 109-116. [doi:10.7652/xjtuxb201306019]
- 郑纪彬, 任爱锋, 苏涛, 等. 多分量 Chirp 信号相位参数的精确估计算法. 2013, 47(2): 69-74. [doi:10.7652/xjtuxb201302012]
- 谢航, 廖与禾, 林京. 一种含噪声多模态调频调幅信号的提取方法. 2012, 46(7): 93-97. [doi:10.7652/xjtuxb201207017]
- 朱军涛, 廖红建, 谢勇勇, 等. 采用短时傅里叶变换的铁路车载探地雷达数据解译方法. 2012, 46(7): 108-114. [doi:10.7652/xjtuxb201207020]

(编辑 荆树蓉)