

基于专家知识和模糊推理的大科学装置故障诊断方法

员天佑, 王晓丽, 安宝冉, 向昭东

(中国工程物理研究院 计算机应用研究所, 四川 绵阳 621999)

摘要: 针对大科学装置技术综合、结构复杂、系统庞大, 在故障诊断方面面临的故障机理不清楚, 难以建立精确的数学模型; 诊断信息不完整、不精确, 难以进行确定性推理; 诊断数据受限, 无法实现数据驱动等诸多问题; 提出了基于专家知识和模糊推理相结合的故障诊断方法和模式匹配算法, 通过模糊因子的引入和基于数据库的模糊诊断知识可视化建模方法的使用, 解决了故障诊断环节的诸多不确定性问题, 形成了面向用户的模糊专家系统故障诊断基础平台, 并在某大型激光驱动装置测试验证平台中得到初步应用, 实现了电气驱动及控制系统故障的智能诊断。

关键词: 故障诊断; 专家系统; 模糊推理; 大科学装置

Fault Diagnosis Method of Big Scientific Device Based on Expert Knowledge and Fuzzy Reasoning

Yun Tianyou, Wang Xiaoli, An Baoran, Xiang Zhaodong

(Institute of Computer Application, China Academy of Engineering Physics, Mianyang 621999, China)

Abstract: A fault diagnosis method and pattern matching algorithm based on combination of expert knowledge and fuzzy reasoning is proposed. Aiming at many problems faced by large scientific installations in fault diagnosis, such as it is difficult to establish an accurate mathematical model because of the failure mechanism is not clear, the diagnostic information is incomplete and inaccurate, so it is difficult to carry out deterministic inference and diagnostic data limitations do not enable data-driven self-learning. The method solves many uncertainty issues in the fault diagnosis process, and forms user-oriented fuzzy expert system fault diagnosis basic platform by introducing fuzzy factors in diagnostic knowledge and the use of database-based fuzzy knowledge modeling methods, which has been initially applied in a large-scale laser drive device test platform and achieved intelligent diagnosis of system failure.

Keywords: fault diagnosis; expert system; fuzzy reasoning; big scientific device

0 引言

大科学装置具有技术综合、结构复杂和系统庞大等特点, 且工作在极端环境下, 在成本、技术和工艺受限的情况下, 系统可靠性难以大幅提升, 必须通过先进的智能故障诊断技术来提升系统的可用性和可维护性^[1]。然而, 目前这方面的研究还相对滞后, 诊断环节人工参与较多, 诊断效率和准确性难以提高, 影响了大科学装置的可用性和可维护性。在技术实现层面也面临不少问题, 如: 大科学装置系统结构复杂, 故障机理不清楚, 难以建立精确的数学模型; 故障征兆与故障原因之间的关联关系带有较强的不确定性, 难以进行确定性推理; 诊断数据积累不足无法实现数据驱动和自主学习等。针对上述问题本文提出了基于专家知识和模糊推理的故障诊断方法^[2-3], 充分利用系统研制和日常维护中形成的先验知识, 通过模糊推理解决因果关系不确定、诊断信息不完备条件下的故障诊断问题。

在知识建模方面, 针对专家系统语言在诊断规则描述方面存在的专业性强, 无法实现面向用户的规则录入问题, 提出了基于数据库的故障诊断模糊规则的可视化建模与结构化存储方法, 构建面向用户的故障诊断专家系统构建框架。

1 基本结构

模糊故障诊断专家系统围绕知识库和模糊推理引擎展开设计, 其具体结构和各组件功能如下:

1) 数据采集与预处理组件: 该组件通过网络化控制系统提供的设备服务接口和通讯中间件自动获取与故障诊断相关的设备状态信息、传感器数据和控制软件的报错信息等, 同时接收工作人员输入的与故障诊断相关的证据。这些综合信息通过清洗、筛选和特征提取, 形成有效的故障征兆信息传递给模糊化组件。

2) 模糊化组件: 该组件对故障征兆进行模糊化处理, 得到故障征兆的隶属度值。然后将故障征兆及其隶属度传递给模糊推理引擎, 前者用于规则前件的匹配, 后者用于规则匹配度计算。

3) 模糊推理引擎: 推理引擎根据实际输入的故障征兆信息和模糊隶属度, 与规则库中的每一条规则进行匹配, 利用模糊推理算法计算其匹配度。如果匹配度大于规则的

收稿日期: 2018-07-09; 修回日期: 2018-08-13。

基金项目: 国防基础科研计划(JCKY2017212C005)。

作者简介: 员天佑(1977-), 男, 大学本科, 高级工程师, 主要从事自动控制、故障诊断和工控安全相关技术方向的研究。

激活阈值, 则触发该规则, 并将该规则对应的诊断结论从知识库中提取出来作为诊断结果输出。

4) 诊断知识建模及管理组件: 该组件向用户提供可视化模糊规则录入界面和规则的浏览、添加、删除、修改工具。

5) 基础库: 基础库包括知识库、算法库和诊断数据库, 分别用于模糊规则、相关算法和过程数据的结构化存储。

存储, 通过结论号进行关联 (如表 2 所示)。具体结构见表 1。

3 模糊推理算法

模糊推理算法主要分为故障征兆模糊化处置, 模糊故障征兆注入, 模糊推理与模式匹配, 诊断结果综合评判及可视化等 4 大步骤, 具体实现过程及算法如下^[2-11]:

表 1 模糊规则

字段名	字段类型	说明
规则编号	整数	用于规则的标记和识别
规则可信度	浮点数	对支撑该规则的多个专家知识进行综合评判给出
前件 1	字符串	对应故障征兆的论域
前件 1 权值	浮点数	故障征兆对诊断结论的支撑度
...		其它前件
诊断结论号	整数	诊断结论的编号, 用于与诊断结论知识库关联

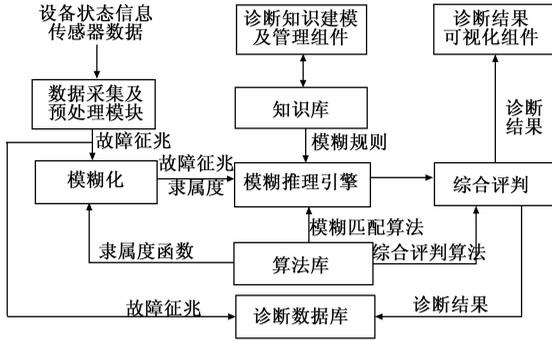


图 1 模糊故障诊断系统结构

2 模糊知识表述与数据库实现

知识的表示方法有产生式规则、框架、语义网、范例等多种形式, 其中产生式规则由于粒度细、灵活性强, 符合人的判断思维而被广泛采用。本方案也采取基于产生式规则的知识表示方式, 通过引入模糊因子形成以下模糊规则^[4-5]:

$$\begin{aligned}
 & \text{IF } P_1(\lambda_1, W_1) \text{ AND } P_2(\lambda_2, W_2) \text{ AND} \dots \\
 & \quad \text{AND } P_n(\lambda_n, W_n) \\
 & \text{THEN } R (CF, \tau)
 \end{aligned}$$

P_i 为规则前件, 即, 故障征兆; λ_i 为前件的隶属度, 即, 故障征兆的可信度 $0 \leq \lambda_i \leq 1$; W_i 为前件的权值, 即, 故障征兆对结论的支持度; R 为规则结论, 即, 故障诊断结论; CF 为规则可信度; τ 为规则激活阈值。

诊断规则的实现方法一般有两种: 第一种是通过编程语言的“IF THEN”判定函数进行描述, 但这种方法无法实现规则 (知识) 与程序的分离, 灵活性极差; 第二种也是最常用的方法, 通过专家系统开发语言进行描述, 将规则编写成结构化文本文件, 由专家系统推理引擎进行解析和执行, 这种方式虽然实现了知识与程序的分离, 但规则文件编写难度极大, 需要掌握专门的专家系统开发语言, 且只能由知识工程师完成, 用户难以建立诊断规则。因此, 本方案拟采取基于数据库的规则建模方法, 通过对规则标准结构的分解, 将前件、结论和各种模糊因子等相关要素用数据库字段加以表示, 形成表 1 所示的结构化模糊诊断规则。用户可以通过可视化知识建模工具方便、快捷的建立、修改、删除和浏览模糊诊断规则。由于模式匹配过程仅对规则的前件进行匹配, 且规则结论包涵信息较多, 为了便于处理, 将规则结论分离出来作为诊断结论知识单独

表 2 诊断结论

字段名	字段类型	说明
诊断结论号	整数	用于诊断结论的标记和识别
故障模式	字符串	故障的具体类型。
故障原因	字符串	导致故障的原因。原因可能有多种, 如果出现多种原因要给出每种原因的隶属度。
故障影响	字符串	故障对系统造成的影响, 要给出影响范围。
严酷度	整数	故障对系统产生的危害程度。
处置措施	字符串	故障的排查、处理方法。包括临时措施和永久措施。

1) 第一步: 输入向量模糊化

诊断系统采集的故障征兆信息按其特征可分为离散量和连续量两种, 前者以状态位和状态字为主, 一般代表某种明确的故障征兆, 后者以传感器测量数据为主, 需要进行预处理才能提取出故障征兆信息。对于这两种故障征兆信息需采取不同的模糊化方法。离散量信息只有两种状态, “真”则隶属度为“1”, “假”则隶属度为“0”。连续量需要用模糊术语进行描述, 并根据隶属度函数计算其隶属度值, 处理结果包括故障征兆论域和隶属度 (可信度), 如“轴温过高 0.8”。

为了便于论述, 假设: 某设备表现出 n 个故障征兆, 可能出现的故障原因有 m 种, 则:

$$\text{故障征兆论域表示为: } U = \{U_1, U_2, \dots, U_n\}$$

$$\text{故障征兆模糊向量表示为: } x = \{x_1, x_2, \dots, x_n, x_i \in [0, 1] (1 \leq i \leq n)\}$$

$$\text{故障原因论域为: } V = \{V_1, V_2, \dots, V_m\}$$

$$\text{故障原因模糊向量表示为: } y = \{y_1, y_2, \dots, y_m\}, y_j \in [0, 1] (1 \leq j \leq m)$$

其中： x_i 为 U_i 的隶属度， y_j 为 V_j 的隶属度。

选择经典的升半梯形、梯形和降半梯形隶属度函数进行计算，则其广义表示表达式为^[3]：

$$\mu(x) = \begin{cases} I(x), x \in [a, b) \\ 1, x \in [b, c] \\ D(x), x \in (c, d] \\ 0, x \notin [a, d] \end{cases} \quad (a \leq b \leq c \leq d) \quad (1)$$

式中， $I(x)$ 为 $[a, b]$ 上的单调递增函数， $D(x)$ 为 $[c, d]$ 上的单调递减函数，其典型分布为：

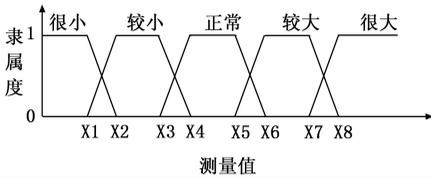


图 2 模糊域划分图

本方案规定数值型测量值的模糊集合为： $A = [很大, 较大, 正常, 较小, 很小]$ ，所有来自传感器的精确测量都用上述 5 种模糊术语表述，并附带其隶属度。

2) 第二步：确定模糊关系矩阵 R

根据多个专家的经验，通过综合评估确定模糊关系矩阵的值，即，确定故障征兆与故障原因的关联强度，该值对应模糊规则中子前件的权值。关系矩阵单独从一个状态原因出发进行评判，不考虑多个原因的综合影响。矩阵的元素值后期可通过诊断历史数据的统计分析结果进行自动修正。

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nm} \end{bmatrix} \quad \sum_{i=1}^n r_{ij} = 1 (j = 1, 2, \dots, m)$$

$r_{ij} \in [0, 1]$ 表示故障征兆 x_i 与故障原因 y_j 的关系强度 (权值)， $r_{ij} = 0$ 表示没有相关性。

3) 第三步：模式匹配

模糊算子采用加权平均型模型则有：

$$Y = X^{\circ}R \Leftrightarrow y_j = \sum x_i \cdot r_{ij} \cdot (j = 1, 2, \dots, m), \quad y_j \in [0, 1] (1 \leq j \leq m) \quad (2)$$

4) 第四步：故障原因模糊综合评判

设 U 为故障原因的集合，则：

采用最大隶属度法有： $\max [y_i] \in U$ (适合求解主要故障原因)

采用阈值法有： $y_i \geq \lambda_i, y_i \in U \lambda_i$ (适合求解多种并发故障原因)

4 故障诊断流程

基于专家知识与模糊推理的故障诊断流程，主要分为数据采集、征兆提取与模糊处理、模式匹配、结果输出和诊断数据归档等 6 大环节，其中模式匹配环节还可进一步细化为规则提取、模式匹配、阈值判断和结论置信度计算

等环节，具体流程如下：

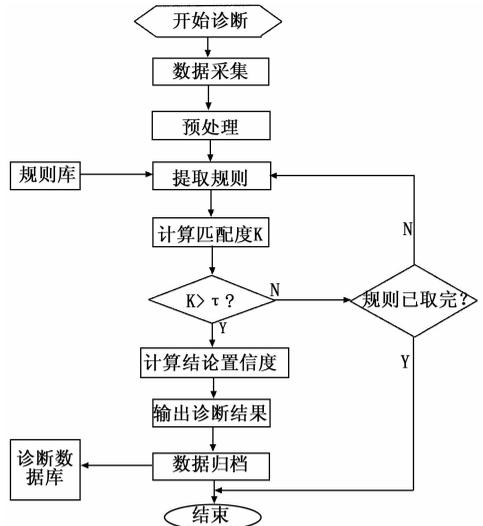


图 3 故障诊断流程

1) 数据采集：数据采集是获取故障诊断信息的过程。故障诊断信息主要包括传感器监测数据、设备运行状态和报警信息等。传感器数据通过数据采集器获取和转换，设备运行状态信息和报警信息通过 OPC (OLE for Process Control) 接口从监控系统直接读取。

2) 预处理：预处理是诊断信息的筛选、清洗和模糊化、归一化过程。由于采集到的原始诊断信息涵盖了装备所有设备及部组建的状态信息和监测数据，数据量大，复杂度高，在利用之前，首先需要根据专家经验进行初步筛选，提取出与某个具体故障模式可能相关的征兆信息，然后用模糊术语加以表述，同时计算出每个故障征兆对应的隶属度。将故障征兆的模糊论域和其隶属度构成的模糊向量作为故障推理的有效数据输入。

3) 模糊推理：模糊推理是模糊规则的遍历与匹配过程。该过程从规则提取开始，推理引擎从规则库中逐条提取诊断规则，将实际获取的故障征兆信息与规则的前件进行匹配，利用公式 (1) 计算每条规则的匹配度，激活匹配度超过阈值的规则。为了提高匹配效率，规则库可以按设备类型进行分类设计，匹配时只遍历与诊断对象对应的规则库。

4) 诊断结果输出：模式匹配成功后，将被激活的规则的后件作为故障诊断结论输出。如果没有匹配成功说明该故障模式没有发生或规则库不完善。由于采用阈值法，允许多个规则同时激活，因此可以诊断多种并发故障。

5) 诊断数据归档：诊断数据归档是诊断结果及过程数据的结构化存储过程。诊断数据主要包括故障征兆 (输入)、诊断结果 (输出)、处置结果 (最终结论) 和时间、人员等附加信息。这些信息以故障编号作为关键字进行关联，形成结构化数据表单，用于后期的诊断规则完善和模糊因子修正，也可作为训练样本用于神经网络训练。