

基于专家知识库属性重要度的故障诊断方法研究^{*})

黄大荣¹ 宋 军² 胡必锦¹

(重庆交通大学理学院 重庆 400074)¹ (重庆交通大学计算机学院 重庆 400074)²

摘 要 利用粗糙集理论中重要性度量的概念和专家知识库集成技术,研究了复杂系统故障诊断的一类新方法。首先,在给出多领域专家知识库集成系统定义和复杂系统状态赋值规则的基础上,利用粗糙集理论中的属性重要度的概念建立了专家知识库属性重要度模型;然后,提出了该模型运行机理的算法规则,并结合多领域专家知识库集成技术,构造了一种能对实时动态环境进行监控和故障诊断的网络化诊断系统;最后,仿真试验验证了该方法的有效性和合理性。此方法成功应用于复杂系统故障诊断的实际工程实例表明,这种诊断方法具有良好的更新能力和应用前景。

关键词 故障诊断,粗糙集理论,专家系统,属性重要度,故障预测

A New Method of Fault Diagnosis Based on Attribute Significance of Multi-field Experts Database

HUANG Da-Rong¹ SONG Jun² HU Bi-Jin¹

(College of Mathematic and Information, Chongqing Jiaotong University, Chongqing 400074)¹

(College of Computer Science, Chongqing Jiaotong University, Chongqing 400074)²

Abstract Based on attribute significance of rough set and integrated technology of experts database, a new method of fault diagnosis to complex systems is researched. First of all, the attribute significance model of experts database is presented by combining integrated system of multi-field experts and assignment rule of fault state with attribute significance concept in rough set theory. Secondly, a network model based on multi-field experts database is constructed, which can be used to monitor the dynamic systems in real-time and diagnosis fault. Finally, a simulated example has shown that the technique is effectively and exactly. The presented method has used in the fault diagnosis of complex systems and good results have been achieved, and the theoretical analysis indicates that the presented algorithm has a broad prospect for practical application.

Keywords Complicated system, Fault predict, Multi-agent integration, Confidence level, Tolerance error

随着现代工业及科学技术的迅速发展,生产设备日趋大型化、高速化、自动化和智能化,一旦发生故障,将产生难以估计的损失,为了避免产生很大的损失,产生了故障诊断这门学科^[1]。

故障诊断就是判断系统是否发生了故障以及检测故障发生的时刻,涉及到了故障分离、故障辨识等内容,包括分离出故障发生的部位(传感器、执行器或元部件)、判别故障的种类等等,并进行相应的评价和决策^[2]。

故障诊断技术最初采用的方式是硬件冗余技术^[3],但是该方式有个致命的缺点就是增加了更多的测试设备,使得系统结构更加复杂,这对于本来就复杂化的系统来说是非常不利的。而且由于该方法需要建立系统精确的模型,这在系统结构日趋复杂的情况下其实际应用就显得有些不足。为了解决这个问题,不少学者基于知识的形式研究了智能故障诊断系统^[1,4~8],其诊断的基本思想就是在实时故障信号处理的基础上,结合领域专家知识和人工智能的技术进行推理,得到诊断的结果。这类诊断系统由于不需要建立复杂系统的精确数学模型,弥补了传统故障技术所存在的不足,因而,在故障诊断技术中受到了极大的重视。然而,遗憾的是这些研究工作都是基于单一的专家系统知识库来作出的,而随着现代工业的不断发展,每个大型复杂系统都包含着不同的领域专家知识,单一的专家系统显然不能同时考虑各方面的因素,故而其故障诊断的结果不能很好地符合实际工程的需要,需要多个

领域的专家知识库来共同决策。与此同时,不同领域的专家系统知识库对于相同的复杂系统的故障诊断的结果又可能不一致,需要有选择地对各种专家知识库进行高效配置。为了解决这些问题,本文基于普适性的角度从单个专家诊断系统出发,结合计算机网络结构和粗糙集的属性重要度的概念研究了多领域专家集成系统在复杂系统故障诊断中的理论规则和应用前景,得到了较好的框架结构。

1 复杂系统故障状态的可诊断性

对于一个复杂系统的故障诊断技术来说,最为重要的就是能根据系统的状态信号来识别设备系统的状态,进而去调整系统的参数保持系统的稳定运行。由于专家在长期的领域研究和处理各种领域问题的过程中,对设备系统的状态有着强有力的知识领悟能力,可以根据设备系统的状态值去诊断故障,因而在此基础上产生了基于专家知识的故障诊断技术。与此同时,由于复杂系统是由一系列功能各异的子子系统组成,不同领域的专家可能会根据对子系统不同的偏好和系统的状态值得到不同的诊断结果;但由于子系统各部分之间相互联系,因此这些专家的领域知识应该是相互联系的,所以要进行准确的故障诊断就需要将它们有机地结合在一起,通过对系统的集成决策来进行故障诊断。

为了本文后面的研究,先给出多领域专家知识库集成系统的定义如下:

^{*})国家自然科学基金主任基金项目(60443006),重庆市教委科技计划项目(KJ060414),重庆交通学院青年科研基金项目(2004-08),重庆市自然科学基金项目(CSTC2006BB2413)。黄大荣 讲师,博士生,研究方向:随机控制理论、信息融合理论及故障预测研究等;宋 军 副教授,博士,研究领域:通信网络以及神经网络等;胡必锦 教授,博士生导师,研究领域:随机分析。

定义 1(多领域专家知识库集成系统) 假设存在 n 个专家知识库, 各领域专家知识库对复杂系统的某个状态的诊断结果记为 $Y_i (i=1, 2, \dots, n)$, 综合所有的专家知识库的决策结果记为 $Y = \sum_{i=1}^n w_i Y_i$ (其中 w_i 为第 i 个专家知识库的权重系数), 由 Y 来判定该复杂系统的状态。该系统称为多领域专家知识库集成系统。

多领域专家知识库集成系统是将各领域专家的知识综合到一起与计算机科学技术相结合的产物, 为复杂系统的故障诊断提供了一种新的算法和问题求解规范, 为实施强方法和弱方法的故障诊断技术提供了一种分布式的处理方法, 可以降低软件技术和硬件的费用。这对于复杂系统的故障状态来说, 恰好弥补了单一专家知识库不能满足各领域综合诊断需要的不足, 也显示了复杂系统故障状态的多领域专家知识库集成系统可诊断特性。这类多领域专家知识库集成系统本质上是从单个专家系统出发结合计算机网络结构来实现对复杂系统的故障诊断的, 增强了专家知识库的完备性。

2 基于专家知识库属性重要度的故障诊断模型及算法分析

2.1 模型参数结构的算法分析

从第 2 节的分析可以看出, 对于复杂系统的同一个状态, 不同的领域专家可能会给出不同的诊断结果。因而, 由定义 1 知道, 对于每一个领域专家知识库的诊断结果有着不同的权重系数 w_i 。同时又因为对于复杂系统的某一个状态来说, 其综合评判结果才是最终的系统的诊断结论, 称之为决策 Y , 即 $Y = \sum_{i=1}^n w_i Y_i$ 。这就是多领域专家知识库集成系统在复杂系统故障诊断中的应用的核心思想。很显然, 首先各领域的专家知识库会根据系统当前的各子系统的状态值确定系统的状态; 接着, 根据每个专家知识库的权重系数确定出系统状态的决策值为

$$Y = \sum_{i=1}^n w_i Y_i \quad (1)$$

由于设备系统的状态要么正常要么是故障情况, 对状态的赋值假定为如下规则:

规则 1 如果该系统为故障状态, 则给 Y 赋值为 1; 否则, 赋值为 0。

根据规则 1 可以得到各领域专家知识库给出的诊断结果要么是 1, 要么是 0, 可对于决策值来说, 由于所得到的结果为所有领域专家结果的权值相加, 不可能得到精确的 1 或者 0, 这样对于某个系统来说要进行故障诊断看起来似乎不是很方便, 为了克服这个问题, 使得故障诊断能顺利进行, 引入如下的定义。

定义 2^[9](就近原则) 在空间中存在有两个点 A, B , 如果存在一个点 C , 使得欧氏距离满足 $d(A, C) \leq d(B, C)$, 那么称点 C 更加接近于 A 。

根据就近原则, 对于复杂系统的决策状态 Y 来说, 一旦满足 $d(Y, 1) \leq d(Y, 0)$, 说明该系统正常; 否则存在故障, 需要对系统进行修正。很显然, 对于复杂系统的故障诊断技术来说, 确定各领域专家系统的权重系数 w_i 并成了多领域专家知识库集成诊断系统非常关键的问题。接下来将对这个进行详细的研究。

2.2 专家知识库属性重要度的确定方法

在复杂系统的运行过程中, 各领域专家和技术人员往往

可以凭视觉、听觉、嗅觉或者测量设备得到的一些客观事实, 并根据对系统结构和系统故障历史的了解很快作出判断。但是不同的领域专家知识库会根据不同的偏好对系统作出不一致的判断, 因而需要确立各领域专家知识库的权重系数 w_i , 一旦确立了 w_i , 就可以根据就近原则对系统的状态进行有效的判断。这说明权重系数的确立对复杂系统的故障诊断技术来说是非常重要的, 下面就来讨论该权重系数的确立方法。

假设对于复杂系统的故障诊断来说, 寻求 n 个领域的专家知识库对复杂系统得故障状态进行判断, 记为 $Y = \{Y_1, Y_2, \dots, Y_n\}$, 把复杂系统故障点记为 Y , 这样对于次历史的系统状态来说, 各领域专家有不同的历史判定结果 $Y_i = \{Y_{1i}, Y_{2i}, \dots, Y_{ni}\}$, ($i=1, 2, \dots, m$), 复杂系统故障点 $Y = \{Y_1, Y_2, \dots, Y_n\}$ 。这样在对每一个状态变量确定值域赋值后, 将全部的观测结果用一个 $n \times m$ 阶矩阵 K 表示, 得到下面的状态数据矩阵:

$$\begin{matrix} & Y_1 & Y_2 & \dots & Y_n & Y \\ \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ m \end{matrix} & \begin{bmatrix} Y_{11} & Y_{12} & \dots & Y_{1n} & Y_1 \\ Y_{21} & Y_{22} & \dots & Y_{2n} & Y_2 \\ Y_{31} & Y_{32} & \dots & Y_{3n} & Y_3 \\ \vdots & \vdots & & \vdots & \vdots \\ Y_{m1} & Y_{m2} & \dots & Y_{mn} & Y_m \end{bmatrix} \end{matrix} = K \quad (2)$$

根据粗糙集理论的相关知识, 可以把 $Y_i = \{Y_{1i}, Y_{2i}, \dots, Y_{ni}\}$ 看作是条件属性, 系统故障 $Y = \{Y_1, Y_2, \dots, Y_n\}$ 看作是结论属性。显然, 并不是所有的领域专家知识库都对复杂系统的正常运行起着同等重要的作用, 一些专家知识库的影响要大一些, 另外一些专家知识库的影响要小一些, 可以根据粗糙集理论中的属性重要度的概念对各个领域专家知识库的重要度进行排序。为了分析这个问题, 首先引入几个定义:

定义 3^[10, 11] 设 $K = (U, R)$ 为一个知识库, 定义 $ind(K)$ 为 K 中所有的等价关系的族, 且 $ind(K) = \{ind(P) | \emptyset \neq P \subseteq R\}$ 。

定义 4^[10, 11] P 和 Q 为 U 的等价关系, Q 的 P 正域记为 $Pos_P(Q) = \bigcup_{X \subseteq U/Q} PX$, 也即 U 中所有根据分类 U/P 的信息可以准确地划分到关系 Q 的等价类中去的对象的集合。

定义 5^[10, 11] 给定一知识库 $K = (U, R)$, 且 $P, Q \subseteq R$, 当 $k = \gamma_P(Q) = |Pos_P(Q)|/|U|$ 时, 称知识 Q 是 k ($0 \leq k \leq 1$) 度依赖于知识 P 的。

定义 6^[10, 11] 假设 C 和 D 分别为条件属性集和决策属性集, 属性子集 $C' \subseteq C$ 关于 D 的重要性定义为 $\sigma_{CD}(C') = \gamma_C(D) - \gamma_{C-C'}(D)$ 。

显然, 不同的领域专家知识库对于结论属性来说具有不同的重要性。为了确立其权重系数 w_i 的大小, 可以根据粗糙集理论中依赖度 $k = \gamma_P(Q)$ 和重要性量度 $\sigma_{CD}(C) = \gamma_C(D) - \gamma_{C-C'}(D)$ 的概念对各个领域专家知识库的重要程度进行排序, 得到如下结论。

$$\sigma_{CD}(Y_{j1}) \geq \sigma_{CD}(Y_{j2}) \geq \dots \geq \sigma_{CD}(Y_{jm}) \quad (3)$$

根据这个排序结果, 建立各领域专家知识库的权重系数确立模型如下:

$$w_i = \frac{\sigma_{CD}(Y_{ji})}{\sum_{i=1}^n \sigma_{CD}(Y_{ji})} \quad (4)$$

从分析中知道, 该权重系数量化指标完全反映了各领域专家知识库的重视程度。我们称之为专家知识库属性重要度模型。该模型为复杂系统的故障判定规则准备了基础条件。

2.3 多领域专家知识库集成系统的算法设计和网络拓扑构造

从上述的模型分析知道,一旦在线监测出复杂系统的各子系统的状态值,各领域专家知识库会根据自己的偏好给出

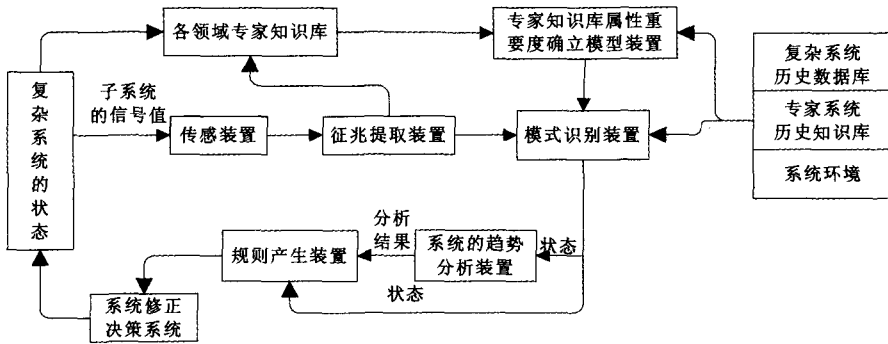


图1 多领域专家系统知识库故障诊断过程

该诊断系统集成多领域专家的知识,构成了一个多领域专家的团体联邦,综合了多领域的不同推理机制的优势,因而在复杂系统故障诊断的时候具有更加灵活的诊断规则和局部更新能力,在工程中能有效地进行在线实时控制。其运行的机理如下:

1)系统初始化(专家系统调研)。调入复杂系统的 m 次历史状态数据,寻求 n 个领域的专家知识库对复杂系统的故障的历史判定结果 $\bar{Y}_i = \{\bar{Y}_{i1}, \bar{Y}_{i2}, \dots, \bar{Y}_{in}\} (i=1, 2, \dots, m)$ 和复杂系统故障点的实际情形 $Y = \{Y_1, Y_2, \dots, Y_m\}$,根据规则1进行赋值,得到 $n \times m$ 阶观测矩阵 K ,论域 $U = \{Y, Y\}$,预存到计算机程序中。

2)粗糙分类。选择接受来自于计算机程序中的条件属性集合 $\bar{Y}$ 和结论属性集合 Y ,利用它们进行分类,得到 $U/\bar{Y}$ 和 U/Y ,由此得到 $\bar{Y}$ 的 Y 正域 $Pos_Y(\bar{Y})$,并得到知识 Y 关于 $\bar{Y}$ 的依赖度 $\gamma_Y(Y)$,输入到4)。

3)约简粗糙分类。对于条件属性集合 $\bar{Y}$ 中的属性进行约简计算,分别去掉属性 $\bar{Y}_j (j=1, 2, \dots, n)$ 后得到决策集合 $\bar{Y}_k = (k=1, 2, \dots, n-1)$,分别对这些决策集合进行粗糙分类,得到 $U/\bar{Y}_k (k=1, 2, \dots, n-1)$,计算知识 Y 关于 $\bar{Y}_k$ 的依赖度 $\gamma_{Y_k}(Y) (k=1, 2, \dots, n-1)$,输入到下一步。

4)知识库的简化。由2)和3)中传来的依赖度值计算系统状态变量重要性度量 $\sigma_{Y_j}(\bar{Y}_j) = \gamma_Y(Y) - \gamma_{Y-(Y_j)}(Y) (j=1, 2, \dots, n)$,根据公式(3)得到排序结果,输入到下一步。

5)通过公式(4)确立各领域专家知识库的属性重要度,输入到下一步。

6)故障诊断的实现。根据在线实时观测的系统状态数据,各领域专家知识库给出相应的诊断结果,调用程序包中的专家知识库属性重要度,结合公式(1)和规则1得出一定系统的状态值。根据定义2,判定系统状态是否满足 $d(Y, 1) \leq d(Y, 0)$,如果满足就修正系统参数,然后重新观测数据;否则表明系统的运行正常,可以继续观测系统的状态数据进行控制。

7)输出结果,得到复杂系统的运行情况的分析报告。

其相应的网络拓扑结构如下:

多领域专家集成故障诊断系统的上述运行机理,充分考虑了复杂系统在各种信息不完备情况下的运行情况,以此构建的系统结构具有高内聚、松耦合的特性,通过合理地考虑各种领域知识专家的知识能力,有着广阔的应用前景。

整个系统的状态情形,然后故障诊断技术员工就可以根据规则1和模型(1)得出整个系统的状态赋值,最后根据就近原则诊断系统的状态。该诊断系统的故障诊断的过程如图1所示。

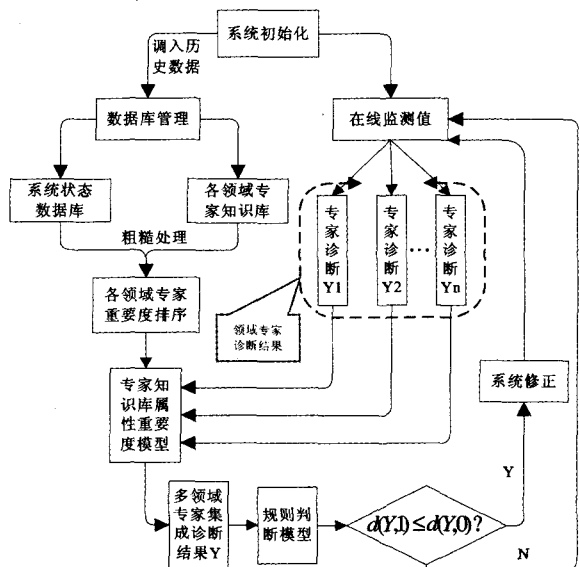


图2 多领域专家集成故障诊断系统的网络模型

3 实验论证

在重庆西蜀电力设备公司生产的SEG-III型励磁装置系统的故障诊断的实际应用中,我们将本文所设计的基于专家知识库属性重要度的故障诊断在线监控网络化系统嵌入到其中进行实际检验。该系统可以根据各领域专家知识对故障推理进行分析,进而去完善匹配的故障规则知识库。以机组起励失败为例,在SEG-III型励磁装置的技术指标中,调节器动态性能要求零起升压超调 M_p 不大于10%,调节时间 t_s 不大于5秒,振荡次数 N 不大于3次,当 M_p 大于10%,或 t_s 大于5秒,或 N 大于3次,励磁系统都被认为发生了起励失败的故障,这样就实现了故障的诊断。该智能故障预测系统适应了电厂状态检修的发展趋势,能够自动监测励磁系统的设备状态,预测和修正设备系统将来的工作情况。其故障诊断的曲线示意图如图3所示。

从系统运行情况的反映来看,以1000秒作为一个周期,在这个周期内,该励磁系统在第102、300、450、700、720、800秒等地方有异常现象发生,说明有故障了。实际应用表明,该诊断规则下的故障诊断结果和实际情形之间的拟合曲线如图4所示。

(下转封三)

细定义且标准的过程;TSP 将综合的产品组织原理与 CMMI 和 PSP 方法结合以产生高效的小组,帮助他们更有效地开发并维护软件。实质上,CMM 和 PSP 为高质量的工程提供了环境和技能,而 TSP 指导开发人员怎样切实地工作,并在具体指导工作的同时,利用了 PSP 和 CMMI 所作的准备。因此为了有效地进行软件过程改进,应将实施 CMMI 与实施 PSP 和 TSP 有机结合起来。

组织在具体实施过程中,可以以 CMMI 为基础架构,从 PSP 先做起,然后在此基础上逐渐转换到 TSP,循序渐进以确保基础的稳定,以助于 CMMI 高成熟度的提升。

总结 总之,目前我国绝大部分小型软件企业的软件过程能力薄弱,需要进行软件过程改进,CMMI 将是软件过程改

进的最佳选择,但由于我国小型企业的规模小、资金不足等特点及所处的文化背景,CMMI 的有关内容不适合于小型企业和项目的实际情况,需对 CMMI 模型进行剪裁和改进,使之适合于小型企业和项目的实际需求。

参考文献

- 1 Ahern D M, et al. CMMI 精粹—集成化过程改进实用导论. 周伯生,等译. 北京:机械工业出版社, 2002
- 2 龚波,于自跃,何新贵,等. 小型软件企业实施 CMMI 过程改进研究和分析. 计算机应用研究, 2004, 8: 64~67
- 3 CMU/SEL. Standard CMMI appraisal method for process improvement version 1. 1[EB/OL]. 2001

(上接第 233 页)

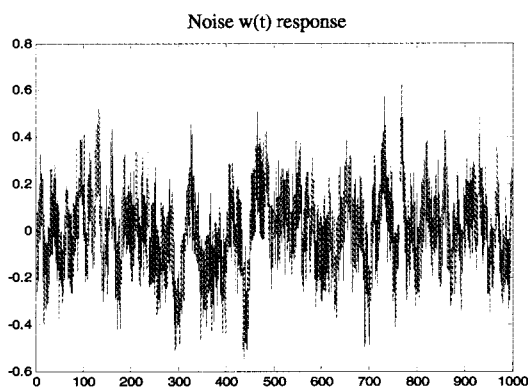


图3 励磁系统故障反应曲线

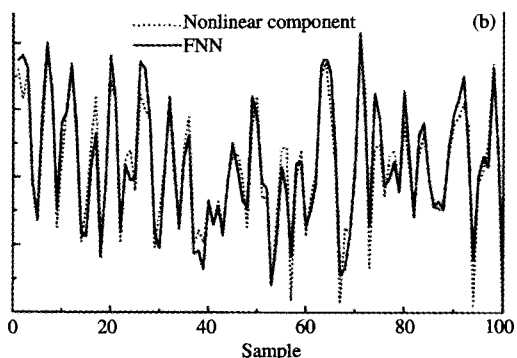


图4 故障诊断系统的数据拟合分析图

同时该系统诊断误差的稳定性分析如图 5 所示。

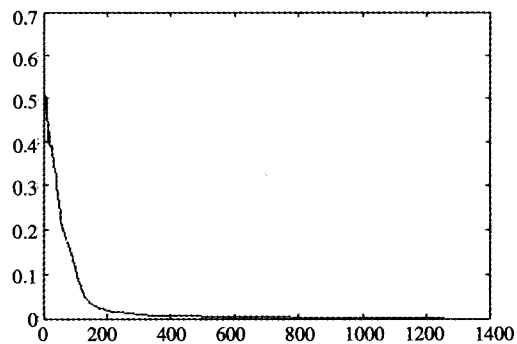


图5 集成故障诊断系统的诊断误差

可以看出,该集成故障系统的诊断误差越来越小,相应的准确性就越来越高。该故障监控系统在电厂励磁装置系统的

有效应用表明,该故障诊断方法的应用价值和前景非常广泛。

结论 复杂系统是个包含多领域专家知识领域的大系统,实施故障诊断的时候需要考虑系统中各领域信息的完备性和协同性。本文所提出的这类基于多领域专家知识库属性重要度的集成诊断系统由于充分考虑了复杂系统各种领域系统的协同一致性,将十分有助于解决实际故障诊断问题,实践证明该系统在复杂系统故障诊断技术中是一种非常有效的方法。

根据作者研究和开发多领域专家知识库集成诊断系统的实践,可以得出如下结论:

1)复杂系统的故障诊断是一个非常庞大的系统,包含有不用领域专家知识,但是并不是所有的领域专家知识对故障诊断具有同等重要的作用,基于多领域专家知识库属性重要度的集成诊断系统可以使复杂系统故障诊断的准确率得到极大提高。

2)在利用多领域专家知识库故障诊断集成系统进行故障诊断的时候,由于各领域专家可以根据相应的检测值判断到底是哪个部位出现了故障,因而该系统可以精确地对故障进行定位。

3)由于各领域专家可以对故障状态和时机进行相应的预测,该系统可以运用到复杂系统的故障预报技术中去。

4)该系统不仅可以运用到复杂系统的故障诊断,还可以扩展其应用领域,特别是在金融工程、政府决策以及企业规划等决策部门,实现一些示范性的应用工程。

参考文献

- 1 吴今培,肖建华. 智能故障诊断与专家系统[M]. 北京:科学出版社, 1997. 1~343
- 2 杨迎化,唐大全,卢建华. 神经网络在智能故障诊断技术中的应用及其发展趋势[J]. 测控技术, 2003, 22(9): 1~5
- 3 周鸣岐,徐军. 控制系统故障诊断[J]. 计算机自动测量与控制, 2000, 8(3): 5~8
- 4 Hayes-Roth. Building expert systems [M]. Addison- Wesley, 1983
- 5 胡寿松,王源. 基于支持向量机的非线性系统故障诊断[J]. 控制与决策, 2001, 16(5): 617~620
- 6 Handelman D A, Stengel R F. Combining expert system and analytical redundancy concepts for fault tolerant flight control [J]. Guide, Control, Dynam, 1989, 12(1): 39~45
- 7 杨天社,杨开忠,李怀祖. 基于知识的卫星故障诊断与预测方法[J]. 中国工程科学, 2003, 5(6): 63~67
- 8 Teo C Y. Application of knowledge-based system for diagnosis and supply restoration [J]. Engineering Application, 1997 (6): 631~638
- 9 黄大荣,黄席樾,范敏,等. 基于卡尔曼滤波和专家系统的混合故障预测研究[J]. 计算机仿真, 2005, 22(9): 150~152
- 10 Pawk Z. Rough set theory and its applications to data analysis [J]. Cybernetics and systems, 1998, 29: 661~668
- 11 张文修,等. 粗糙集理论与方法[M]. 北京:科学出版社, 2003. 1~222