

基于模糊粗糙集的案例推理 在发电设备 RCM 分析中的应用

董晓峰¹, 顾煜炯¹, 杨昆¹, 贺徙²

- (1. 电站设备状态监测与控制教育部重点实验室(华北电力大学), 北京市 昌平区 102206;
2. 中国电力投资集团公司, 北京市 西城区 100032)

Application of Case-based Reasoning for Power Plant Equipment RCM Analysis Based on Fuzzy Rough Set

DONG Xiao-feng¹, GU Yu-jiong¹, YANG Kun¹, HE Xi²

- (1. Key Laboratory of Condition Monitoring and Control for Power Plant Equipment(North China Electric Power University), Ministry of Education, Changping District, Beijing 102206, China;
2. China Power Investment Corporation, Xicheng District, Beijing 100032, China)

ABSTRACT: Case-based reasoning is introduced into power plant equipment RCM analysis process by amending similar equipment RCM analysis case to simplify the tasks of RCM analysis and shorten time which consumed in the analysis process. In order to use case-based reasoning for power plant equipment RCM analysis, case is represented with frame representation, case base is built according to power plant system structure, nearest neighbor algorithm is used as case retrieval method, case base is maintained with threshold. At last, a case inference process is implemented to steam feed pump RCM analysis case. At the same time, a fuzzy rough set algorithm is used in feature reduction and weight allocation for power plant equipment performance features to reduce case memory, simplify case retrieval complexity, avoid human's subjectivity influence on weight allocation and improve case retrieval efficiency, which is used to reduce steam feed pump performance features.

KEY WORDS: case-based reasoning; case retrieval; fuzzy rough set; feature reduction; weight allocation; reliability centered maintenance analysis

摘要: 将基于案例的推理应用于发电设备以可靠性为中心的维修(reliability centered maintenance, RCM)分析过程, 通过修改类似设备的RCM分析案例, 可以缩短RCM分析

时间, 简化 RCM 分析任务。在发电设备 RCM 分析案例推理应用过程中, 使用框架结构作为案例表示, 根据发电设备的系统划分建立案例库, 采用最相邻近法作为案例检索策略, 通过设定阈值实现案例库的维护, 最后对汽动给水泵 RCM 分析案例进行了案例推理。同时, 采用一种基于模糊粗糙集的属性约简方法, 对发电设备的性能特征进行约简和权重分配, 减少了案例存储空间, 简化了案例检索算法的复杂性, 避免了人的主观性对权重分配的影响, 提高了案例检索效率, 并将该方法应用于汽动给水泵的性能特征约简。

关键词: 案例推理; 案例检索; 模糊粗糙集; 属性约简; 特征权重; 以可靠性为中心的维修分析

0 引言

以可靠性为中心的维修(RCM)因其先进性和实用性在我国发电企业得到了广泛应用^[1], 但其复杂冗长的分析过程又妨碍了它的推广。近几年, 我国电力行业得到了迅猛发展, 发电机组向着高参数、大容量、高自动化发展, 随着生产系统的日益复杂化, 系统所包含设备的数量及类型亦日益增加, 其中一个小火电厂的设备就多达数千个, 若要进行 RCM 分析, 工作量非常大。因此很有必要寻找可节省实施 RCM 分析所需时间和精力有效方法对原有的 RCM 分析过程进行改进。

文献[2-3]分别考虑了发电设备的重要程度和维修/故障影响, 对电厂中的设备有选择的进行 RCM 分析, 把维修资源集中在能做得最好的地方,

基金项目: 国家 863 高技术基金项目(2006AA04Z426); 北京市优秀人才培养资助项目(20071D1600900433)。

The National High Technology Research and Development Program of China(863 Programme)(2006AA04Z426).

但是并没有改进重要设备的 RCM 分析过程。文献 [4-5] 把基于案例的推理(case-based reasoning, CBR) 这种人工智能技术应用于军用武器装备 RCM 分析过程, 通过对类似案例的修改, 重复的 RCM 分析任务被简化, 缩短了 RCM 分析时间。但是, 冗余的设备特征不仅大大增加了案例的存储空间, 而且使案例检索算法变得更加复杂。此外, 传统的权重分配方法增加了人的主观性对案例检索准确性的不良影响。

本文采用一种基于模糊粗糙集的属性约简方法, 对设备特征进行约简和权重分配, 去除了冗余的设备特征, 克服了主观分配权重存在的不足。并将该方法用于汽动给水泵的设备性能特征约简中。然后, 将基于案例的推理应用于发电设备 RCM 分析过程, 详细描述了发电设备 RCM 分析案例的案例表示、案例组织、案例检索、案例调整及案例维护, 并对汽动给水泵 RCM 分析案例进行了案例推理。

1 模糊粗糙集

1.1 粗糙集相关理论

1982 年波兰数学家 Pawlak 提出了粗糙集理论, 该理论是一种刻画不完整性和不确定性的数学工具, 能有效地分析和处理不精确、不一致和不完整等各种不完备信息, 并从中发现隐含的知识, 揭示潜在的规律^[6]。

粗糙集理论对知识的分类主要是通过知识的约简来实现的^[7-10]。由于信息系统的属性并非同等重要, 甚至一些属性是冗余的, 因此通过知识约简可在保持系统分类能力不变的条件下, 删除不相关或不重要的属性知识, 提取出最能反映系统特征和变化规律的属性, 使系统达到最简化。

以下为粗糙集理论中有关知识约简的相关定义。信息系统 $S = \langle U, C \cup D, V, f \rangle$, 其中, U 是一个非空对象集合; C 和 D 是对象的属性集合, 即条件属性集 C 和决策属性集 D ; V 是属性值的集合; f 是一个信息函数, 即 $f: U \times (C \cup D) \rightarrow V$, 它指定了 U 中每组对象的属性值。

给定知识库 $K = (U, S)$, 其中, U 为论域, S 表示论域 U 上的等价关系簇, 则对于每个子集 $X \subseteq U$ 和论域 U 上的一个等价关系 $R \in \text{IND}(K)$, 称

$$R_*(X) = \cup \{x \in U : [x]_R \subseteq X\} \quad (1)$$

为 X 的 R 下近似。

2 个属性集 C 与 D 之间的依赖程度定义为

$$\gamma_C(D) = |\text{POS}_C(D)|/|U| \quad (2)$$

式中: $\text{POS}_C(D) = C_*(D)$; $|U|$ 表示整个集合对象的个数。

对于条件属性 $a_i \in C$, a_i 关于 D 的重要程度定义为

$$\gamma_{a_i}(D) = \gamma_C(D) - \gamma_{C-\{a_i\}}(D) \quad (3)$$

式中, $\gamma_{C-\{a_i\}}(D)$ 表示在 C 中缺少属性 a_i 后, 条件属性相对于决策属性的依赖程度。如果 $\gamma_{a_i}(D) = 0$, 则属性 a_i 可以从属性集中去掉。对于保留下来的条件属性, 每个条件属性的权重被定义为

$$\omega_i = \gamma_{a_i}(D) / \sum_i \gamma_{a_i}(D) \quad (4)$$

由专家凭经验给出的权重虽然能在一定程度上反映实际情况, 而且评价结果与实际相近, 但其具有一定的主观性, 有时不能反映实际情况, 从而导致评价结果失真。而基于粗糙集理论得到的权重分配方法, 克服了主观分配权重存在的不足, 使权重分配结果符合实际。

1.2 模糊粗糙集

在经典粗糙集理论中, 信息系统中的各属性只能是离散的或者是符号型的, 然而, 在实际情况中出现在信息系统中的各属性经常是分明的或者是实值的。虽然连续的属性值可以通过属性离散化方法离散, 但势必会丢失一些重要信息, 从而对分类结果产生不良影响。近些年来许多研究者通过用模糊集来代替精确集、通过在论域上引入模糊相似关系代替等价关系, 把经典粗糙集理论扩展得到模糊粗糙集^[11-15]。

以下为模糊粗糙集中有关知识约简的相关定义。模糊关系 R_B 定义为

$$R_B = \{(x, y) \in U \times U : \mu_{R_B}(x, y)\} \quad (5)$$

式中: $R_B \in F(U \times U)$; $\mu_{R_B}(x, y) \in [0, 1]$, $\mu_{R_B}(x, y)$ 的大小反映了 (x, y) 隶属于 R_B 的程度。如果满足: 1) 对称性: $\mu_{R_B}(x, y) = \mu_{R_B}(y, x)$, $\forall x, y \in U$; 2) 自反性: $\mu_{R_B}(x, x) = 1$, $\forall x \in U$, 则称 R_B 是 U 上的模糊相似关系。

对于 $\forall x \in U$, 定义

$$R_B^\lambda = \{y \in U : \forall \mu_{R_B}(y, x) \geq \lambda, y R_B^\lambda x\} \quad (6)$$

式中: λ 为设定的阈值, $R_B^\lambda(x)$ 表示与 x 的相似度超过 λ 的对象集合。则基于模糊相似关系的 X 的 R_B^λ 下近似被定义为

$$R_{B^*}^\lambda(X) = \cup \{x \in X : R_B^\lambda(x) \subseteq X\} \quad (7)$$

X 的 R_B^λ 正域被定义为

$$\text{POS}_{R_B^\lambda}(X) = R_{B^*}^\lambda(X) \quad (8)$$

2 个属性集 C 与 D 之间的依赖程度 $\gamma_C(D)$ 定义为

$$\gamma_C(D) = \text{card}(\text{POS}_{R_C}(D)) / \text{card}(U) \quad (9)$$

1.3 模糊相似矩阵

信息系统 $S = \langle U, C \cup D, V, f \rangle$, U 是包含 n 个案例的对象集合, u_p 和 u_q 是 U 中的任意 2 个对象, $a_i(u_p)$ 是条件属性集 C 中条件属性 a_i 的条件属性值, u_p 与 u_q 的相似度被定义为

$$S_B(u_p, u_q) = 1 - \sqrt{\frac{1}{k} \sum_{i=1}^k [a_i(u_p) - a_i(u_q)]^2} \quad (10)$$

$R_B = S_B(u_p, u_q) (p=1 \sim n, q=1 \sim n)$, 可以表示为一个模糊相似矩阵 R_B , R_B 被定义为

$$\begin{cases} r_{pq} = 1, & p = q \\ r_{pq} = r_{qp}, & p \neq q \end{cases} \quad p, q = 1 \sim n \quad (11)$$

设定阈值 $\lambda \in [0, 1]$, 如果 $r_{pq} \geq \lambda$, 则 u_q 与 u_p 相似, u_q 属于 u_p 的相似类。如果 $r_{pq} < \lambda$, 则 u_q 与 u_p 不相似, u_q 不属于 u_p 的相似类。根据相似程度, R_B 被转化为

$$\begin{cases} r_{pq} = 1, & r_{pq} \geq \lambda \\ r_{pq} = 0, & r_{pq} < \lambda \end{cases} \quad p, q = 1 \sim n \quad (12)$$

2 汽动给水泵性能特征约简与权重分配

2.1 数据转化

以火电厂汽动给水泵额定工况下的性能特征

作为使用模糊粗糙集进行属性约简和权重分配的实例。表 1 显示了火电厂汽动给水泵的性能特征。

表 1 中包含 3 种类型的数据, “机组容量”特征是序数型数据, “锅炉压力”特征是名词型数据, 其余特征是实值型数据。名词型数据和序数型数据都将被转换为符号型数据: 机组容量: {300, 330, 350, 600} \rightarrow {0, 1, 2, 3}; 锅炉蒸汽压力: {亚临界, 超临界} \rightarrow {0, 1}。

为了消除各种量纲之间的差异, 对实值型数据使用极差标准化公式进行标准化:

$$V'_{a_{pi}} = [V_{a_{pi}} - \min_{p=1 \sim n}(V_{a_{pi}})] / [\max_{p=1 \sim n}(V_{a_{pi}}) - \min_{p=1 \sim n}(V_{a_{pi}})], \quad i=1 \sim n \quad (13)$$

式中: $V_{a_{pi}}$ 是案例 C_p 的属性 a_i 的属性值; $V'_{a_{pi}}$ 为标准化后的值; $\max_{p=1 \sim n}(V_{a_{pi}})$ 是属性 a_i 在所有案例中的最大值; $\min_{p=1 \sim n}(V_{a_{pi}})$ 是属性 a_i 在所有案例中的最小值。

对表 1 中的 3 种类型数据进行转化, 转化后的特征值见表 2。

按式(10)、式(11)计算表 2 中所有对象在全部条件属性 $a_1 \sim a_{10}$ 下的模糊相似矩阵 R_C 。

在全部条件属性 $a_1 \sim a_{10}$ 下的模糊相似矩阵为

表 1 汽动给水泵性能特征列表

Tab. 1 Performance features list of steam feed pump cases

序号	进水压 力 a_1/MPa	入口流 量 $a_2/(\text{t/h})$	出口流 量 $a_3/(\text{t/h})$	扬程 a_4/m	必需汽蚀 余量 a_5/m	转速 $a_6/$ (r/min)	出口压 力 a_7/MPa	轴功 率 a_8/kW	抽头压 力 a_9/MPa	抽头流 量 $a_{10}/(\text{t/h})$	机组容 量 d_1/MW	锅炉 压力 d_2
1	1.366	479.9	454.90	2 090.9	18.0	5 070	19.85	3 246.0	10.61	25.00	2×300	亚临界
2	1.850	565.0	538.00	2 318.5	33.8	5 388	21.00	4 220.1	11.76	27.00	2×330	亚临界
3	1.450	654.9	605.46	2 401.0	31.5	5 521	22.20	4 363.0	10.00	33.00	2×330	亚临界
4	1.700	634.0	604.00	2 448.3	33.9	5 324	23.27	5 052.6	12.42	30.00	2×350	亚临界
5	1.320	513.0	499.00	3 305.0	25.0	5 795	30.30	5 425.0	11.50	14.00	2×350	超临界
6	2.273	928.0	885.50	2 128.0	32.0	4 633	21.00	6 431.0	9.760	42.50	2×600	亚临界
7	2.238	932.5	890.00	2 258.0	23.0	5 390	22.11	6 657.0	11.27	42.50	2×600	亚临界
8	2.460	964.0	925.00	3 221.0	31.0	5 410	30.45	10 059.0	13.66	39.00	4×600	超临界
9	2.600	872.97	850.97	3 270.4	31.0	5 200	31.13	9 113.0	12.00	23.00	2×600	超临界
10	2.310	944.36	897.35	3 294.0	50.0	5 760	30.884	9 850.0	14.43	41.94	2×600	超临界

表 2 转化后的汽动给水泵性能特征列表

Tab. 2 Transformed performance features list of steam feed pump cases

序号	进水压 力 a_1/MPa	入口流 量 $a_2/(\text{t/h})$	出口流 量 $a_3/(\text{t/h})$	扬程 a_4/m	必需汽蚀 余量 a_5/m	转速 $a_6/$ (r/min)	出口压 力 a_7/MPa	轴功 率 a_8/kW	抽头压 力 a_9/MPa	抽头流 量 $a_{10}/(\text{t/h})$	机组容 量 d_1/MW	锅炉 压力 d_2
1	0.036	0.000	0.000	0.000	0.000	0.376	0.000	0.000	0.182	0.386	300	亚临界
2	0.414	0.176	0.177	0.187	0.494	0.65	0.102	0.143	0.428	0.456	330	亚临界
3	0.102	0.361	0.32	0.255	0.422	0.764	0.208	0.164	0.051	0.667	330	亚临界
4	0.297	0.318	0.317	0.294	0.497	0.595	0.303	0.265	0.57	0.561	350	亚临界
5	0.000	0.068	0.094	1.000	0.219	1.000	0.926	0.32	0.373	0.000	350	超临界
6	0.745	0.926	0.916	0.031	0.438	0.000	0.102	0.467	0.000	1.000	600	亚临界
7	0.717	0.935	0.926	0.138	0.156	0.651	0.200	0.501	0.323	1.000	600	亚临界
8	0.891	1.000	1.000	0.931	0.406	0.669	0.940	1.000	0.835	0.877	600	超临界
9	1.000	0.812	0.843	0.972	0.406	0.488	1.000	0.861	0.48	0.316	600	超临界
10	0.773	0.959	0.941	0.991	1.000	0.97	0.978	0.969	1.000	0.98	600	超临界

$$R_C = \begin{bmatrix} 1 & 0.744 & 0.718 & 0.685 & 0.490 & 0.437 & 0.452 & 0.200 & 0.275 & 0.128 \\ & 1 & 0.807 & 0.876 & 0.556 & 0.525 & 0.577 & 0.373 & 0.436 & 0.338 \\ & & 1 & 0.806 & 0.566 & 0.560 & 0.617 & 0.379 & 0.426 & 0.340 \\ & & & 1 & 0.590 & 0.558 & 0.627 & 0.469 & 0.510 & 0.433 \\ & & & & 1 & 0.239 & 0.337 & 0.361 & 0.469 & 0.329 \\ & & & & & 1 & 0.749 & 0.453 & 0.464 & 0.352 \\ & & & & & & 1 & 0.575 & 0.536 & 0.468 \\ & & & & & & & 1 & 0.761 & 0.775 \\ & & & & & & & & 1 & 0.627 \\ & & & & & & & & & 1 \end{bmatrix} \quad (14)$$

2.2 性能特征约简与权重分配

根据决策属性对所有对象进行分类：

$$U|(\{d_1, d_2\}) = \{\{1\}, \{2,3\}, \{4\}, \{5\}, \{6,7\}, \{8,9,10\}\}.$$

设定阈值 $\lambda=0.877$ ，按式(12)重新计算模糊相似矩阵 R_C ，得到式(15)：

$$R_C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & 1 & 0 & 0 & 0 & 0 & 0 \\ & & & & & 1 & 0 & 0 & 0 & 0 \\ & & & & & & 1 & 0 & 0 & 0 \\ & & & & & & & 1 & 0 & 0 \\ & & & & & & & & 1 & 0 \\ & & & & & & & & & 1 \end{bmatrix} \quad (15)$$

根据式(15)，把所有对象进行分类： $U|(C) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{9\}, \{10\}\}.$

按式(7)~(9)计算 $\gamma_C(D)$ ： $\gamma_C(D) = 1$ 。

按式(3)、(7)~(9)依次计算每个条件属性 a_i 关于 D 的重要程度： $\gamma_{a_1}(D) = 0$ ， $\gamma_{a_2}(D) = 1/5$ ， $\gamma_{a_3}(D) = 1/5$ ， $\gamma_{a_4}(D) = 0$ ， $\gamma_{a_5}(D) = 0$ ， $\gamma_{a_6}(D) = 0$ ， $\gamma_{a_7}(D) = 1/5$ ， $\gamma_{a_8}(D) = 0$ ， $\gamma_{a_9}(D) = 3/10$ ， $\gamma_{a_{10}}(D) = 0$ 。

经过计算性能特征“进水压力”、“扬程”、“必需汽蚀余量”、“转速”、“轴功率”、“抽头流量”被去除，“入口流量”、“出口流量”、“出口压力”、“抽头压力”4个性能特征被保留。

按式(4)计算“入口流量”、“出口流量”、“出口压力”、“抽头压力”4个性能特征的权重： $\omega_{a_2} = 0.222$ ， $\omega_{a_3} = 0.222$ ， $\omega_{a_7} = 0.222$ ， $\omega_{a_9} = 0.334$ 。

3 发电设备 RCM 分析案例推理

3.1 基于案例的推理

基于案例的推理是指借用旧的事例或经验来解决新的问题。它是伴随认知心理学的研究而发展起来的一种新的推理方法，属于类比推理，是机器

学习的一个分支^[16]。基于案例的推理由案例库和案例推理过程2部分组成。案例库包括案例表示和案例组织^[17]。案例推理过程由案例检索、案例调整和案例维护组成。

3.2 案例表示

案例表示就是对 RCM 分析案例尽可能进行详细地描述，由问题的描述部分和解决方案构成。RCM 分析案例的问题描述部分由设备的基本特征信息组成，如设备的名称、种类、性能参数等。RCM 分析案例的解决方案部分由 RCM 分析信息组成，主要包括故障模式影响分析、RCM 逻辑判断等。

案例的描述采用框架表示法^[18]。汽动给水泵的 RCM 分析案例的框架表示如表 3 所示。在设备结构特征中，由于现有案例库中汽动给水泵均为双壳体、筒形、内筒水平中分、叶轮对称布置的卧式离

表 3 汽动给水泵的 RCM 分析案例的框架表示

Tab. 3 Frame representation of steam feed pump RCM analysis case

框架名	某电厂汽动给水泵 RCM 分析案例
槽 1 设备类型描述	侧面 1. 设备所属系统：汽轮机系统
	侧面 2. 设备所属分系统：给水加热系统
	侧面 3. 设备所属于子系统：给水泵系统
槽 2 机组总体 性能指标	侧面 1. 电厂名称：A
	侧面 2. 机组容量：2×330MW
	侧面 3. 锅炉压力：亚临界
槽 3 设备特征	侧面 1. 设备生产商：A
	侧面 2. 设备型号：40CHTA/5SP
	侧面 3. 设备性能特征： 特征 1 入口流量为 479.9 t/h 特征 2 出口流量为 454.9 t/h 特征 3 出口压力为 19.85 MPa 特征 4 抽头压力为 10.61 MPa
	侧面 4. 设备结构特征： 特征 1 首级叶轮型式为单吸 特征 2 叶轮级数为 6 级 特征 3 密封形式为机械密封
槽 4 RCM 分析信息	侧面 1. 故障模式影响分析： 设备的功能、故障模式、故障原因等
	侧面 2. RCM 逻辑判断： 针对具体的故障采取的相应的维修策略

心泵, 轴承型式均为四油楔滑动轴承+瓦块式推力轴承, 其结构特征主要区别在首级叶轮型式(单吸或双吸)、叶轮级别(5级或6级)、密封型式(机械密封或迷宫密封)。所以并没有对设备结构特征进行约简, 而是直接采用了首级叶轮型式、叶轮级别、密封型式作为设备结构特征。

3.3 案例组织

案例组织就是在案例表示的基础上, 根据案例的特征和检索的需要, 对案例进行整理和归类。在发电设备RCM分析过程中, 首先要对发电设备进行系统划分^[19], 由于故障模式和原因离开具体的设备都难以把握其确切的含义, 本文主要研究基于设备相似的案例, 用系统、分系统、子系统或设备的名称作为案例的名称。因此, 案例可以分为系统案例、分系统案例、子系统案例和设备案例。系统根据种类不同, 可以分为锅炉系统案例、汽轮机系统案例、电动机系统案例, 同时根据层次关系, 系统案例又可以分解为分系统案例, 分系统案例又可以分解为子系统案例, 子系统案例又可以分解为设备案例, 这些案例形成树状结构, 各层次实际案例可分别存储或单独使用。案例树状结构反映了案例之间的关联关系和案例库结构。汽轮机系统案例库结构如图1所示。

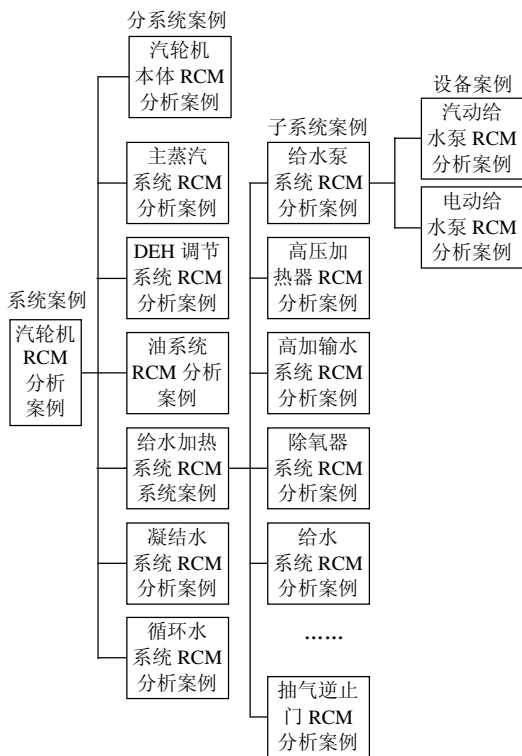


图1 汽轮机系统案例库结构

Fig. 1 System case base structure for steam turbine

3.4 案例检索

案例检索就是根据一定的检索策略, 在案例库中找到与待分析设备案例相似的案例。在RCM分析案例检索中采用最相近邻法(nearest neighbor, NN)^[20]作为案例检索策略, 通过最相邻近法可以得到待分析设备案例与案例库中被检索案例之间的相似度。

1) 2个案例的相似度被定义为

$$S(c_p, c_q) = 0.2S_s(c_p, c_q) + 0.2S_t(c_p, c_q) + 0.3S_p(c_p, c_q) + 0.3S_f(c_p, c_q) \quad (16)$$

式中: $S_s(c_p, c_q)$ 为设备生产商特征的相似度; $S_t(c_p, c_q)$ 为设备型号特征的相似度; $S_p(c_p, c_q)$ 表示设备性能特征的相似度; $S_f(c_p, c_q)$ 为设备结构特征的相似度。 $S(c_p, c_q)$ 的值越大, 表示 c_p 与 c_q 越相似。

2) 对于设备生产商和设备型号特征的相似度采用直接字符串匹配算法, 如果2个字符串属性相同, 则它们的相似度为1, 否则为0。

3) 设备性能特征相似度采用各属性相似度的加权:

$$S_p(c_p, c_q) = \sum_{j=1}^m \omega_j S_j(c_p, c_q) \quad (17)$$

式中: c_p 为待分析设备案例; c_q 为案例库中的被检索案例; ω_j 为第 j 个特征的权重, 权重采用根据模糊粗糙集约简后的计算结果; $S_j(c_p, c_q)$ 为 c_p 和 c_q 在第 j 个特征上的相似度, 其计算公式如下:

$$S_j(c_p, c_q) = 1 - \frac{|a_{pj} - a_{qj}|}{\max\{a_{pj}, a_{qj}\}} \quad (18)$$

式中: a_{pj} 为 c_p 第 j 个特征的特征值; a_{qj} 为 c_q 第 j 个特征的特征值。

4) 设备结构特征相似度采用各属性相似度的加权:

$$S_f(c_p, c_q) = \sum_{j=1}^m \omega_j S_j(c_p, c_q) \quad (19)$$

式中: c_p 为待分析设备案例; c_q 为案例库中的被检索案例; ω_j 为第 j 个特征的权重, 权重采用平均值, 均为0.333; $S_j(c_p, c_q)$ 为 c_p 和 c_q 在第 j 个特征上的相似度, 采用直接字符串匹配算法, 如果2个字符串属性相同, 则它们的相似度为1, 否则为0。

3.5 案例调整

通过案例检索, 得到了按相似度降序排列的候选案例集, RCM分析人员可以选择相似度的一个或几个案例作为参考。所选出的RCM分析案例与待分析设备案例之间往往存在或多或少的差异,

无法直接应用，需要适当的调整。RCM 分析人员根据所选出案例的 RCM 分析信息来修改和填写待分析设备案例的 RCM 分析信息。形成新的 RCM 分析案例，添加到案例库中，丰富案例库。

3.6 案例维护

随着新案例的加入，案例库不断完善和丰富，求解能力也逐渐加强。为避免案例库的无限膨胀，需要对案例库进行维护。本文中案例维护策略如下：

- 1) 不同生产商的设备案例，均予以保留。
- 2) 相同生产商的不同型号设备案例，均予以保留。

3) 对于相同生产商的相同型号设备案例，由案例的设备性能特征相似度 S_p 值决定新案例增删。设定相似度阈值为 0.9，如相似度大于 0.9，则删除新案例；否则予以保留。

4 汽动给水泵 RCM 分析案例推理

汽动给水泵 RCM 分析案例推理过程如下：

- 1) 形成待分析设备案例，见表 4。
- 2) 确定可检索的案例库，见表 5。
- 3) 进行案例检索，按式(16)~(19)计算待分析设备案例与案例库中案例的相似度，计算结果见表 6。

表 4 汽动给水泵 RCM 分析案例
Tab. 4 Steam feed pump RCM analysis case

序号	生产商	设备型号	设备性能特征				设备结构特征		
			入口流量/(t/h)	出口流量/(t/h)	出口压力/MPa	抽头压力/MPa	首级叶轮型式	叶轮级数	密封型式
1	B	FK6D32	555	531	22.14	8.45	单吸	6	迷宫密封

表 5 汽动给水泵案例库中的 RCM 分析案例
Tab. 5 RCM analysis cases of steam feed pump case base

案例号	生产商	设备型号	设备性能特征				设备结构特征		
			入口流量/(t/h)	出口流量/(t/h)	出口压力/MPa	抽头压力/MPa	首级叶轮型式	叶轮级数	密封型式
1	A	CHTC5/6SP-3	479.9	454.9	19.85	10.61	单吸	6	机械密封
2	B	FK6D32	565	538	21	11.76	单吸	6	迷宫密封
3	A	CHTC5/6-3P	654.9	605.46	22.2	10	单吸	6	机械密封
4	B	FK6F32M	634	604	23.27	12.42	单吸	6	机械密封
5	C	14X14X16A-5STGHDB	513	499	30.3	11.5	双吸	6	机械密封
6	D	CHTC6/5	928	885.5	21	9.76	单吸	5	机械密封
7	A	MDG346	932.5	890	22.11	11.27	双吸	5	迷宫密封
8	A	MDG364	964	925	30.45	13.66	双吸	5	迷宫密封
9	A	MDG366	872.97	850.97	31.13	12	单吸	6	机械密封
10	D	CHTD6/5	944.36	897.35	30.884	14.43	单吸	6	机械密封

表 6 汽动给水泵 RCM 分析案例检索结果
Tab. 6 Retrieval result of steam feed pump RCM analysis case

案例号	S	S_s	S_r	S_p	S_f
1	0.454	0	0	0.254	0.2
2	0.966	0.2	0.2	0.266	0.3
3	0.466	0	0	0.266	0.2
4	0.648	0.2	0	0.248	0.2
5	0.346	0	0	0.246	0.1
6	0.33	0	0	0.23	0.1
7	0.321	0	0	0.221	0.1
8	0.287	0	0	0.187	0.1
9	0.402	0	0	0.202	0.2
10	0.385	0	0	0.185	0.2

4) RCM 分析人员可以选择相似度高的 2 号案例或 4 号案例作为参考，形成新的案例。

5) 由于汽动给水泵案例库中已有同一生产商相同型号的案例，但 $S_p \leq 0.9$ ，新案例予以保留。

5 结论

在发电设备 RCM 分析案例推理中，不同的发

电设备具有不同的性能特征，有的发电设备具有很多的性能特征，本文采用的模糊粗糙集方法在避免了对连续的属性值进行属性离散化带来的信息损失的情况下，去除了冗余的设备性能特征，对特征权重进行了合理分配，改进了案例检索效果。把基于案例的推理应用于发电设备 RCM 分析，使用框架结构作为案例表示，根据发电设备的系统划分建立案例库，采用最相邻近法作为案例检索策略，通过设定阈值实现案例库的维护，并对汽动给水泵 RCM 分析案例进行了案例推理。使用案例推理可以通过修改类似设备的 RCM 分析案例，缩短发电设备 RCM 分析时间，简化 RCM 分析任务。

参考文献

[1] 曹先常, 史进渊, 蒋安众, 等. 基于模糊数学的电站设备故障风险定量研究[J]. 中国电机工程学报, 2005, 25(23): 119-123.
Cao Xianchang, Shi Jinyuan, Jiang Anzhong, et al. Based on fuzzy mathematics failure risk quantitative analysis for power station equipment[J]. Proceedings of the CSEE, 2005, 25(23): 119-123(in

- Chinese).
- [2] 董玉亮, 顾煜炯, 杨昆. 基于蒙特卡洛模拟的发电厂设备重要度分析[J]. 中国电机工程学报, 2003, 23(8): 201-205.
Dong Yuliang, Gu Yujiong, Yang Kun. Criticality analysis on equipment in power plant based on Monte Carlo simulation [J]. Proceedings of the CSEE, 2003, 23(8): 201-205(in Chinese).
- [3] 曹钟中, 杨昆, 顾煜炯, 等. 汽轮机及其辅助设备系统以可靠性为中心的维修(RCM)的技术分析原则[J]. 国际电力, 2002, 6(3): 30-35.
Cao Zhongzhong, Yang Kun, Gu Yujiong, et al. RCM technologic analyzing principles of a steam turbine and its auxiliary equipment & system[J]. International Electric Power for China, 2002, 6(3): 30-35(in Chinese).
- [4] 程中华, 贾希胜, 李震, 等. 基于案例的 RCM 分析系统案例库的设计[J]. 计算机工程与设计, 2005, 26(5): 1196-1202.
Cheng Zhonghua, Jia Xisheng, Li Zhen, et al. Design of case-base of case-based RCM analysis system[J]. Computer Engineering and Design, 2005, 26(5): 1196-1202(in Chinese).
- [5] Cheng Zhonghua, Jia Xisheng, Gao Ping, et al. A framework for intelligent reliability centered maintenance analysis[J]. Reliability Engineering & System Safety, 2008, 93(6): 806-814.
- [6] 王永强, 律方成, 李和明. 基于粗糙集理论和贝叶斯网络的电力变压器故障诊断方法[J]. 中国电机工程学报, 2006, 26(8): 137-141.
Wang Yongqiang, Lü Fangcheng, Li Heming. Synthetic fault diagnosis method of power transformer based on rough set theory and Bayesian network[J]. Proceedings of the CSEE, 2006, 26(8): 137-141(in Chinese).
- [7] 王志勇, 郭创新, 曹一家. 基于模糊粗糙集和神经网络的短期负荷预测方法[J]. 中国电机工程学报, 2005, 25(19): 7-11.
Wang Zhiyong, Guo Chuangxin, Cao Yijia. A method for short term load forecasting integrating fuzzy-rough set with artificial neural network[J]. Proceedings of the CSEE, 2005, 25(19): 7-11(in Chinese).
- [8] 刘思革, 程浩忠, 崔文佳. 基于粗糙集理论的多目标电网规划最优优化模型[J]. 中国电机工程学报, 2007, 27(7): 65-69.
Liu Sige, Cheng Haozhong, Cui Wenjia. Optimal model of multi-objective electric power network planning based on rough set theory[J]. Proceedings of the CSEE, 2007, 27(7): 65-69(in Chinese).
- [9] 朱永利, 吴立增, 李雪玉. 贝叶斯分类器与粗糙集相结合的变压器综合故障诊断[J]. 中国电机工程学报, 2005, 25(10): 159-165.
Zhu Yongli, Wu Lizeng, Li Xueyu. Synthesized diagnosis on transformer faults based on Bayesian classifier and rough set [J]. Proceedings of the CSEE, 2005, 25(10): 159-165(in Chinese).
- [10] 莫娟, 王雪, 董明, 等. 基于粗糙集理论的电力变压器故障诊断方法[J]. 中国电机工程学报, 2004, 24(7): 162-167.
Mo Juan, Wang Xue, Dong Ming, et al. Diagnostic model of insulation faults in power equipment based on rough set theory [J]. Proceedings of the CSEE, 2004, 24(7): 162-167(in Chinese).
- [11] Jensen R, Shen Q. Fuzzy-rough attribute reduction with application to web categorization[J]. Fuzzy Sets and Systems, 2004, 141(3): 469-485.
- [12] Shen Q, Jensen R. Selecting informative features with fuzzy-rough sets and its application for complex systems monitoring[J]. Pattern Recognition, 2004, 37(7): 1351-1363.
- [13] Jiang Yajun, Chen Jun, Ruan Xueyu. Fuzzy similarity-based rough set method for case-based reasoning and its application in tool selection [J]. International Journal of Machine Tools & Manufacture, 2006, 46(2): 107-113.
- [14] 熊浩, 李卫国, 畅广辉, 等. 模糊粗糙集理论在变压器故障诊断中的应用[J]. 中国电机工程学报, 2008, 28(7): 141-147.
Xiong Hao, Li Weiguo, Chang Guanghui, et al. Application of fuzzy rough set theory to power transformer faults diagnosis[J]. Proceedings of the CSEE, 2008, 28(7): 141-147(in Chinese).
- [15] 郭庆琳, 郑玲. 基于模糊粗糙集数据挖掘的汽轮机故障诊断研究[J]. 中国电机工程学报, 2007, 27(7): 81-87.
Guo Qinglin, Zheng Ling. A novel approach for fault diagnosis of steam turbine unit based on fuzzy rough set data mining theory [J]. Proceedings of the CSEE, 2007, 27(7): 81-87(in Chinese).
- [16] 余贻鑫, 吴建中. 基于事例推理模糊神经网络的中压配电网短期节点负荷预测[J]. 中国电机工程学报, 2005, 25(12): 18-23.
Yu Yixin, Wu Jianzhong. CBRFNN-based short-term nodal load forecasting for middle voltage distribution networks[J]. Proceedings of the CSEE, 2005, 25(12): 18-23(in Chinese).
- [17] 杨兵, 丁辉, 罗为民, 等. 基于知识库的变压器故障诊断专家系统[J]. 中国电机工程学报, 2002, 22(10): 121-124.
Yang Bing, Ding Hui, Luo Weimin, et al. Expert system of transformer fault diagnosis based on knowledge base[J]. Proceedings of the CSEE, 2002, 22(10): 121-124(in Chinese).
- [18] 年志刚, 梁式, 麻芳兰, 等. 知识表示方法研究与应用[J]. 计算机应用研究, 2007, 24(5): 234-236.
Nian Zhigang, Liang Shi, Ma Fanglan, et al. Study and application of knowledge expression[J]. Application Research of Computers, 2007, 24(5): 234-236(in Chinese).
- [19] 董玉亮. 发电设备运行与维修决策支持系统研究[D]. 北京: 华北电力大学, 2005.
- [20] 李青, 史雅琴, 周扬. 基于案例推理方法在飞机故障诊断中的应用[J]. 北京航空航天大学学报, 2007, 33(5): 622-626.
Li Qing, Shi Yaqin, Zhou Yang. CBR methodology application in fault diagnosis of aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(5): 622-626(in Chinese).



董晓峰

收稿日期: 2009-01-17。

作者简介:

董晓峰(1979—), 男, 博士研究生, 研究方向为电站设备状态监测、故障诊断和维修决策, dongxiaofengdx@163.com;

顾煜炯(1968—), 男, 工学博士, 教授, 研究方向为电站设备状态监测、故障诊断和维修管理;

杨昆(1957—), 男, 工学博士, 教授, 研究方向为电站设备状态监测、故障诊断和维修管理。

(责任编辑 车德竞)