

文章编号: 1001-0920(2013)04-0511-06

基于粗糙集和图论的电力系统故障诊断方法

卢 鹏, 王锡淮, 肖健梅

(上海海事大学 物流工程学院, 上海 200135)

摘 要: 将粗糙集与图论相结合处理电力系统故障诊断, 提出了故障决策表图的新概念, 得到一种基于粗糙集和图论的电力系统故障诊断方法, 并进一步提出了故障信息覆盖度和故障诊断规则分级的概念. 利用故障决策表图及其邻接矩阵, 得到了快速识别决策表核属性和属性约简的方法, 并将规则分级应用于故障规则提取. 利用所提出的方法对具体实例进行处理, 仿真结果表明, 该方法能有效地减少时间和空间复杂度, 可根据设定的阈值提取诊断规则.

关键词: 故障决策表图; 邻接矩阵; 属性约简; 故障信息覆盖度

中图分类号: TP18

文献标志码: A

Method of fault diagnosis in power system based on rough set theory and graph theory

LU Peng, WANG Xi-huai, XIAO Jian-mei

(Logistic Engineering College, Shanghai Maritime University, Shanghai 200135, China. Correspondent: LU Peng, E-mail: lplplp20051@163.com)

Abstract: The new concept of the fault decision table graph is proposed based on the idea of processing for fault diagnosis in power system which combines rough set with graph theory. Furthermore, the concepts of fault information coverage and fault diagnosis rule classification are presented by using the fault decision table graph and its adjacency matrix, and the method of fast recognition of core attributes of a decision table and the method of attribute reduction are obtained, and the concept of fault diagnosis rule classification is applied to fault rules extraction. An example is processed by using the proposed method, and the simulation results show that, the method can effectively reduce the complexity of time and space, and extract diagnosis rules according to the set threshold value.

Key words: fault decision table graph; adjacency matrix; attribute reduction; fault information coverage

0 引 言

随着风电、水电等清洁能源规模的不断扩大, 电力系统的结构也日趋复杂. 为了适应各种复杂情况下故障的快速、准确识别和判定故障元件的位置、类型, 国内外学者和工程技术人员提出了许多电力系统故障诊断的技术和方法, 其中主要有专家系统^[1-2]、优化算法^[3]、模糊集理论^[4]和多代理技术^[5]等.

粗糙集理论以其强大的处理不确定性信息的能力在人工智能、决策支持等领域获得了广泛的应用. 电力系统故障诊断要根据保护和断路器的动作信号判断出故障元件或区域, 可以将故障现象及元件状态表述为模式分类问题, 适合用粗糙集理论的决策表方法进行处理^[6]. 而在进行电力系统的故障诊断时, 电力系统的故障诊断规则对应于粗糙集理论中的属性

约简, 所以使用粗糙集方法进行电力系统故障诊断时需要取粗糙集决策表中的所有约简, 这是一个 NP 完全问题^[7].

目前将粗糙集应用于故障诊断的领域有很多. 如文献 [8-9] 将粗糙集理论与支持向量机相结合; 文献 [10] 将沃尔什变换与粗糙集相结合; 文献 [11] 结合了粗糙集理论与克隆算法; 文献 [12] 将自适应遗传算法用于粗糙集属性约简.

上述文献主要关注于应用粗糙集理论对故障诊断决策表进行约简后的规则提取问题, 而粗糙集属性约简本身即是 NP 问题. 对此, 本文将图论知识和粗糙集理论相结合, 对粗糙集中的决策表进行处理: 首先根据决策表构造决策表图, 推导出相应的决策表图的邻接矩阵; 然后提出了快速识别决策表核属性的方法

收稿日期: 2011-12-29; 修回日期: 2012-04-17.

基金项目: 上海市教委科研创新重点项目(12ZZ158); 上海市教委重点学科建设项目(J50602).

作者简介: 卢鹏(1981—), 男, 博士生, 从事粗糙集理论、智能信息处理的研究; 王锡淮(1961—), 男, 教授, 博士生导师, 从事粗糙集理论、复杂系统建模与控制等研究.

以及对邻接矩阵进行处理并得出属性约简的方法,该方法可根据核属性得到全面的属性约简.在规则提取方面,针对约简后的决策表,提出了故障信息覆盖度和故障诊断规则分级的概念,可根据设定的阈值,按照一定的置信度和规则覆盖度提取诊断规则.

1 基于图论的粗糙集属性约简方法

1.1 决策表信息系统与加权多部决策表图

定义 1 设给出的决策表系统

$$S = (U, C, D, V, f). \quad (1)$$

其中: $U = \{x_1, x_2, \dots, x_n\}$ 为样本空间, $C = \{c_1, c_2, \dots, c_r\}$ 为条件属性集合, $D = \{1, 2, \dots, k\}$ 为决策属性集合. S 中不存在不确定信息和冲突信息,将相同的样本按照一个样本来处理,得到新的样本空间 U (以下所提到的 U 都是经过处理后的样本空间).以 U 中的样本为点,任意两个样本点之间都连 $|C \cup D|$ 条属性边,构成的经典多重完全图称为决策表图.

定义 2 在定义 1 给出的决策表图中,以决策属性值作为条件,对图中的样本点进行分部,决策属性值相同的点在同一部,任意不在同一部的两点连 $|C|$ 条边.这样得到的图称为分部(或多部)决策表图.

定义 3 在定义 2 给出的多部决策表图中,如果两样本点关于某个条件属性的值是相同的,则将这两个样本点间所连的这条属性边删除,再将两样本点之间条件属性值不同的每条边用所代表的条件属性作为边的权进行标注,并将两样本点之间的所有边合并为一条新边,标上它们的全部权的并集.这样得到的图称为加权多部决策表图.

例 1 点 X_1, X_2 为加权多部决策表图中不同部

表 1 例 1 的决策表

U	条件属性		
	a	b	d
X_1	1	0	1
X_2	2	2	2
X_3	1	1	0
X_4	2	0	1
X_5	0	2	0
X_6	0	1	1
X_7	2	1	1

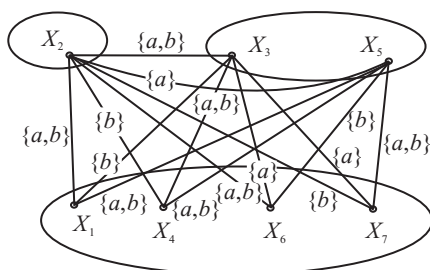


图 1 依据决策表 1 构造加权多部决策图

的两点,两点间有边 e_1, e_2 和 e_3 , 权值分别为 i, j 和 k , 于是 X_1, X_2 之间的边合并后为一条多权边 e_{12} , 权值为 $\{i, j, k\}$, 也可以表示为 $i \cup j \cup k$. 进而构造由决策表 1 中的决策属性作为分部条件, 条件属性作为加权条件所确定的加权多部决策表图, 如图 1 所示.

1.2 利用加权多部决策表图判断核属性

定义 4 边 e 称为加权多部决策表图的单边, 当且仅当与边 e 所关联的点 u, v 之间仅连接有 e 这一条边.

定义 5 在加权多部决策表图 G 中, 如果边 e 满足条件: $G - e$ 是完全图的充分必要条件为 G 是完全图, 则称 e 是加权多部决策表图 G 的可去边; 否则是 G 的不可去边.

本文将具有性质 a 的所有边形成的边类记为 E_a , 则有如下定义:

定义 6 在加权多部决策表图 G 中, 如果边类 E_a 满足条件: $G - E_a$ 是完全图的充分必要条件为 G 是完全图, 则称 E_a 是加权多部决策表图 G 的可去边类; 否则称 E_a 是 G 的不可去边类.

定理 1 加权多部决策表图 G 中, e 是 G 的不可去边的充分必要条件为 e 是 G 的单边.

证明 e 是 G 的不可去边 $\Leftrightarrow G - e$ 不是完全图 $\Leftrightarrow e$ 是单边. \square

定理 2 加权多部决策表图 G 中, E_s 为可去边类的充分必要条件是: $\forall V_i, V_j \in G (i \neq j)$, 使得 $E_{ij} \not\subseteq E_s$. 其中 $E_{ij} = \{e | e \text{ 是连接 } V_i \text{ 和 } V_j \text{ 间的边}\}$.

证明 E_s 为可去边类 $\Leftrightarrow G - E_s$ 是完全图 \Leftrightarrow 任意两点 V_i 和 $V_j (i \neq j)$ 间仍有边相连 $\Leftrightarrow E_{ij} \not\subseteq E_s$.

推论 1 加权多部决策表图 G 中, E_s 为不可去边类的充分必要条件是: $\exists V_i, V_j \in G$, 使得 $E_{ij} \subseteq E_s$.

结合粗糙集知识, 可得定理 3.

定理 3 如果决策表信息系统可约简, 则加权多部决策表图中的单边所表示的条件属性为核属性之一; 如果所有的属性边类均存在单边, 则该决策表信息系统在绝对意义上不可约简, 即没有核属性.

证明 如果决策表信息系统可约简, 而且图中存在单边(单边在加权多部决策表图中表现为权为单元素集的边), 则由定理 1 和定理 2 知该边为不可去边, 且该边所属的一类边为不可去边类. 由于边类与条件属性一一对应, 该边所表示的属性为核属性. 如果所有的属性边类均存在单边, 则由前面的证明可知决策表中所有的条件属性均不可约简, 即没有核属性. \square

1.3 基于属性置信度的近似计算方法

定义 7 属性 C_i 的置信度定义为: 属性 C_i 的单

边个数与所有属性单边个数之和的比值,即

$$\text{Confidence}(C_i) = \text{NUM}_i / \sum_{j=1,2,\dots,k} \text{NUM}_j. \quad (2)$$

当样本空间很大且决策表中所有条件属性都绝对不可约简时,利用定义7进行近似约简.若置信度条件为 $(100 - a)\%$,则可将所有属性置信度小于 $a\%$ 的条件属性约去.

2 基于粗糙集和图论的故障诊断方法

2.1 故障诊断决策表图的构造

故障诊断决策表图的产生主要有以下几个步骤:

- 1) 依照故障现象提取故障诊断决策表;
- 2) 用定义1的方法依照故障诊断决策表构造决策表图;
- 3) 用定义2的方法依照决策属性作为分部条件构造分部决策表图;
- 4) 用定义3的方法依照条件属性作为边的权构造加权多部决策表图.

具体应用如前面的例1所示.

2.2 对故障诊断决策表图进行属性约简

2.2.1 故障诊断决策表图的邻接矩阵

在2.1节得到的加权多部决策表图 $k(n_1, n_2, \dots, n_k)$ 中, n_i 代表决策表中属于决策属性值 i 的样本点有 n_i 个,一共有 k 个决策属性值, $n_1 + n_2 + \dots + n_k = n$. 可以得到加权多部决策表图的邻接矩阵为

$$\begin{bmatrix} \phi_{n_1 \times n_1} & A_{n_1 \times n_2} & \cdots & A_{n_1 \times n_k} \\ A_{n_1 \times n_2}^T & \phi_{n_2 \times n_2} & \cdots & A_{n_2 \times n_k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n_1 \times n_k}^T & A_{n_2 \times n_k}^T & \cdots & \phi_{n_k \times n_k} \end{bmatrix}. \quad (3)$$

该矩阵为分块对称阵,矩阵 $A_{n_i \times n_j}$ 中的元素是第 i 部与第 j 部的边的权集合,表示决策表中各个样本点之间的差异.

定义8 在由决策表信息系统构成的加权多部决策表图 G 的邻接矩阵中,如果存在着非空元素是某个集合 M 的子集合,则称集合 M 对这个邻接矩阵是“包容的”;否则称为“不包容的”.

定理4 在决策表信息系统中,某些属性的集合不可以被约简掉的充分必要条件是:该属性集合对决策表信息系统构成的加权多部决策表图 G 的邻接矩阵是包容的.

证明 充分性:如果对于属性集合 $C_k (\exists A_{ij} \neq \emptyset) \wedge (A_{ij} \subseteq C_k)$,则由定理2的推论知属性集合 C_k 所代表的边类为不可去边类,属性集合 C_k 不可以被约简掉.

必要性(用反证法):若该属性集合对邻接矩阵

是“不包容的”,则邻接矩阵中不存在非空元素是其子集合,由定理2知该属性所代表的边类为可去边类,该属性集合是可以被约简的,这与已知条件矛盾. \square

2.2.2 基于邻接矩阵的属性约简

基于加权多部决策表图的邻接矩阵的元素,通过对加权多部决策表图进行点的收缩和边的合并,实现粗糙集决策表的属性约简.

点收缩的方法是将与要收缩的两点相关联的所有多权边进行合并,具体做法如下.

- 1) 当要收缩的两点 A 和 B 都相邻于同一个点 C 时,即 AC 间有多权边 e_1 , BC 间有多权边 e_2 时,将边 e_1 、 e_2 合并成一条新的多权边 e_{12} . 新边 e_{12} 的权值按以下规则得到:将 e_1 和 e_2 两条多权边的权值集合都化为范式形式,进行 \cap 操作,其中“;”表示 \cup ,“*”表示 \cap .

例如, $\{a_1, \{a_2 * a_3\}\}$ 可表示为 $a_1 \cup (a_2 \cap a_3)$, $a_1 \cap (a_2 \cup a_3)$ 可表示为 $\{a_1 * \{a_2, a_3\}\}$.

- 2) A 、 B 两点之间的关联边保持原来的权不变,成为 A 、 B 收缩合并后的新点上的多权环.

- 3) 与 A 或 B 关联的边,均与 A 、 B 收缩合并后的新点关联权不变.

- 4) 其余边保留不变.

- 5) 对加权多部决策表图的所有点依次进行步骤1)~步骤4)的操作点的收缩,最后将图收缩为两个复合点.

最后得到的两点间的权值集合即为粗糙集决策表的属性约简.

2.2.3 属性约简的算法描述

输入信息:决策系统(1).

- 1) 遍历样本空间,按照决策属性值将样本点进行分属,决策属性值相同的样本点属于同一类,一共分为 k 类.设每类包含样本点的个数分别为 n_1, n_2, \dots, n_k ,于是得到加权多部决策表图 G 的点集合 $V(n_1, n_2, \dots, n_k)$.

- 2) 依照本文2.1节的方法构造加权多部决策表图 G ;再按照2.2.1节的方法对样本点进行处理,得到决策表图 G 的分块邻接矩阵 A_k .

- 3) 判断加权多部决策表图 G 是否符合1.3节中的条件,如果符合,则按照属性置信度进行近似计算;否则,依照2.2.2节的方法进行点的收缩,直至收缩为两个样本点.在此过程中,随时将 A_k 中表示单边的元素记入核属性集合.

- 4) 收缩完毕,计算两个样本点之间的权值便可得到粗糙集的属性约简.

2.2.4 属性约简方法的复杂度分析

设 $M = \sum_{i=1}^k N_i$ 表示样本空间, 根据决策属性值将样本空间分为 k 块, 每一块的样本个数为 N_i ($i = 1, 2, \dots, k$), 则本文算法的空间和时间复杂度为 $|C| \sum_{i,j=1,2,\dots,k,i \neq j} |N_i| \times |N_j|$; 特别地, 当 N_i 中有一项特别大时, 复杂度近似为 $k|C||M|$, 与传统粗糙集方法的复杂度 $|C||M|^2$ 相比可以发现, 本文方法大幅降低了算法的复杂度。

当存在置信度约束时, 时间复杂度进一步降为 $\sum_{i,j=1,2,\dots,k,i \neq j} |N_i| \times |N_j|$. 实际应用中, 可以根据故障的重要程度在各个属性上进行分配, 并与置信度结合, 按照置信度进行属性约简。

2.3 决策规则的分级与故障信息覆盖度

设故障决策信息表中有 N 个条件属性, 即 N 个故障, 每个条件属性值为 0 或 1, 则所有的故障信息总数为

$$\text{NUM} = C_N^1 \times 2^1 + C_N^2 \times 2^2 + \dots + C_N^N \times 2^N. \quad (4)$$

2.3.1 决策规则的分级

如果提取的决策规则中, 仅有一个条件属性值确定, 其他可以取任意值, 则定义此类决策规则为一级规则; 有两个条件属性值确定, 其他可取任意值的, 为二级规则; 以此类推. 例如: $1, *, *, \dots, *$ 为一级规则; $1, 0, *, *, \dots, *$ 为二级规则。

2.3.2 故障信息覆盖度

设故障诊断决策表有 N 个故障, 即 N 个条件属性. 对于一级规则而言, 仅有一个条件属性值确定, 其他可取 0 或者 1, 因此这个一级规则包含 $C_{N-1}^0 \times 2^0$ 个 (即 1 个) 一级规则 (它本身); $C_{N-1}^1 \times 2^1$ 个二级规则, \dots , $C_{N-1}^{N-1} \times 2^{N-1}$ 个 N 级规则. 所以一级规则的故障信息覆盖数为

$$\text{NUM}_1 = C_{N-1}^0 \times 2^0 + C_{N-1}^1 \times 2^1 + \dots + C_{N-1}^{N-1} \times 2^{N-1}. \quad (5)$$

对于二级规则而言, 有两个条件属性值确定, 其他取 0 或者 1, 包含 $C_{N-2}^0 \times 2^0$ 个 (即 1 个) 二级规则, $C_{N-2}^1 \times 2^1$ 个三级规则, \dots , $C_{N-2}^{N-2} \times 2^{N-2}$ 个 N 级规则. 所以二级规则的故障信息覆盖数为

$$\text{NUM}_2 = C_{N-2}^0 \times 2^0 + C_{N-2}^1 \times 2^1 + \dots + C_{N-2}^{N-2} \times 2^{N-2}. \quad (6)$$

以此类推, 可得 m 级规则 ($m \leq n$) 的故障信息覆盖数为

$$\text{NUM}_m = C_{N-m}^0 \times 2^0 + C_{N-m}^1 \times 2^1 + \dots +$$

$$C_{N-m}^{N-m} \times 2^{N-m}. \quad (7)$$

定义 9 任一 q 级决策规则的信息覆盖度定义为

$$\text{coverage rate}(q) = \text{NUM}_q / \text{NUM}. \quad (8)$$

2.3.3 故障信息覆盖度的计算

设一个决策规则系统中, 包含一个 a 级规则和一个 b 级规则, 则此决策规则系统的故障信息覆盖度有以下两种情况。

1) 两个规则为真包含关系。

例如, 其中一个为一级规则, 另一个为由此一级规则衍生出的低级规则. 设规则 a 为 $1, *, *, \dots, *$; 规则 b 为 $1, 0, *, *, \dots, *$. 规则 a 真包含规则 b . 可知两个规则为真包含关系情况下故障信息覆盖度为

$$\text{coverage rate}(a + b) = \text{coverage rate}(a). \quad (9)$$

2) 两个规则为相交关系。

例如, 规则 a 为 $1, 1, *, *, \dots, *$; 规则 b 为 $*, 1, 1, *, \dots, *$. 相交后的规则为 $1, 1, 1, *, *, \dots, *$. 又如, 两规则为 $1, 1, *, *, \dots, *$ 和 $*, *, 1, 1, *, \dots, *$, 则相交后的规则为 $1, 1, 1, 1, *, \dots, *$. 可知两个规则为相交关系情况下故障信息覆盖度为

$$\begin{aligned} \text{coverage rate}(a + b) = \\ \text{coverage rate}(a) + \text{coverage rate}(b) - \\ \text{coverage rate}(a \cap b). \end{aligned} \quad (10)$$

2.4 决策规则的样本覆盖度

对于决策系统 $S = (U, A, V, f)$, $X_i \in U/\text{IND}(C)$ ($i = 1, 2, \dots, m$) 和 $Y_j \in U/\text{IND}(D)$ ($j = 1, 2, \dots, n$) 分别代表 U/C 和 U/D 中的各个等价类, 置信度为 $\mu(X_i, Y_j) = |X_i \cap Y_j| / |X_i|$, 则决策规则的样本覆盖度为

$$\alpha(X_i, Y_j) = |X_i \cap Y_j| / |Y_j|. \quad (11)$$

2.5 基于粗糙集和图论的故障诊断算法

算法步骤如下:

1) 按照 2.2.2 节的方法对原始决策表进行约简, 得到核属性和相应的属性约简。

2) 从核属性所能代表的决策规则开始, 按照故障信息覆盖度、置信度和样本覆盖度的顺序由大到小将决策规则加入规则集合, 已经被规则集合中的已有规则包含的规则不再计算. 可以事先规定各个数值的阈值, 根据阈值进行选择. 如果不设阈值, 规则集合将比较杂乱, 提取规则时很难做到规则提取的完备性. 设置阈值, 可以得到有规律的规则集合. 设置的阈值越高, 得到的规则越少, 越具有代表性. 规则集合里面的

下限, 如设置30%, 将表示规则集合里的每一条规则至少能代表30%的覆盖面. 阈值越小, 规则越细, 越不具有代表性. 最低阈值可以提取所有的规则, 但其中大部分规则要求的条件过高, 因而不具有代表性.

3) 规则提取完毕后, 在进行故障诊断时, 根据故障诊断样本的实际情况选择对应级别的故障规则, 应选取置信度大的故障规则.

3 算法应用实例

为了检验本文算法的正确性和性能, 选取文献[13]给出的配电系统(如图2所示), 按照本文所给出的方法与文献[13]进行对比实验.

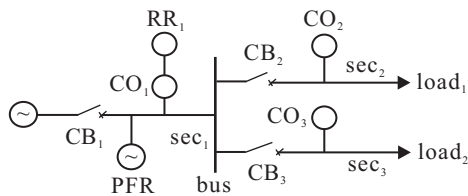


图2 简单配电系统

在如图2所示的简单配电系统中, 3个区域sec₁、sec₂和sec₃分别配有过流保护CO₁、CO₂和CO₃, sec₁配有距离保护RR₁, 为sec₂和sec₃提供后备保护. CB₁、CB₂和CB₃是断路器, PFR是电力故障录波器, 记录sec₁、sec₂和sec₃区域故障前后的电气模拟量.

根据保护动作原理, 可以得到表2所示的由7个电网故障样本组成的决策表, 其条件属性分别为: 断路器CB₁、CB₂、CB₃, 过流保护CO₁、CO₂、CO₃, 距离保护RR₁. 条件属性取值为0或者1, 其中: 1表示断路器由闭合变位为断开或保护动作, 0表示断路器未变位或保护未动作. 决策属性为故障区域, 分别为线路sec₁、sec₂、sec₃和sec_{2/3}(表示故障发生在sec₂或者sec₃). NO表示无故障.

表2 电网故障诊断数字化后的决策表

样本	CB ₁	CB ₂	CB ₃	CO ₁	CO ₂	CO ₃	RR ₁	故障
1	1	0	0	1	0	0	0	sec ₁
2	0	1	0	0	1	0	0	sec ₂
3	0	0	1	0	0	1	0	sec ₃
4	1	0	0	0	1	0	1	sec ₂
5	1	0	0	0	0	1	1	sec ₃
6	1	0	0	0	0	0	1	sec _{2/3}
7	0	0	0	0	0	0	0	NO

3.1 构造决策表图及其邻接矩阵

根据本文中2.5节的方法对表2进行处理. 记CB₁为a, CB₂为b, CB₃为c, CO₁为d, CO₂为e, CO₃为f, RR₁为g, 可以得到如下的加权多部决策表图和邻接矩阵.

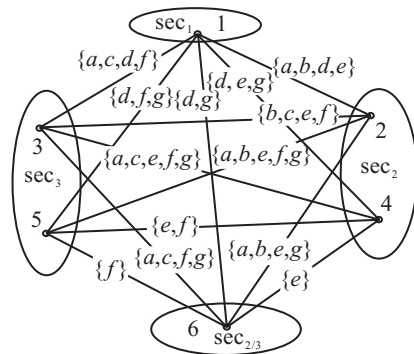


图3 依据表2构造的加权多部决策表图

图3的邻接矩阵为

$$A = \begin{bmatrix} \phi & \{a, b, c, d\} & \{d, e, g\} \\ \{a, b, d, e\} & \phi & \phi \\ \{d, e, g\} & \phi & \phi \\ \{a, c, d, f\} & \{b, c, e, f\} & \{a, c, e, f, g\} \\ \{d, f, g\} & \{a, b, e, f, g\} & \{e, f\} \\ \{d, g\} & \{a, b, e, g\} & \{e\} \\ \{a, c, d, f\} & \{d, f, g\} & \{d, g\} \\ \{b, c, e, f\} & \{a, b, e, f, g\} & \{a, b, e, g\} \\ \{a, c, e, f, g\} & \{e, f\} & \{e\} \\ \phi & \phi & \{a, c, f, g\} \\ \phi & \phi & \{f\} \\ \{a, c, f, g\} & \{f\} & \phi \end{bmatrix}$$

3.2 依据故障信息覆盖度提取规则集合

从图3的邻接矩阵可以得到核属性{CO₂, CO₃}, 同时约简掉了条件属性CB₂和CB₃, 得到的属性约简为{CB₁, CO₁, CO₂, CO₃}, {CB₁, RR₁, CO₂, CO₃}, {RR₁, CO₁, CO₂, CO₃}. 从核属性开始提取规则并得到规则的各个参数, 因为表2中没有重复样本, 所以规则的样本覆盖度基本无意义, 可以不考虑. CB₂和CB₃为可约简属性, 在实际计算规则参数时不考虑. 于是可得规则CO₂ = 1 → sec₃的故障信息覆盖度为33.5%, 置信度为100%.

在故障信息覆盖度为33.5%, 置信度为100%的阈值下可得规则集合为

$$\{CO_2 = 1 \rightarrow sec_2, CO_3 = 1 \rightarrow sec_3, CO_1 = 1 \rightarrow sec_1\}.$$

设置故障信息覆盖度的阈值为11.2%, 置信度阈值为100%, 可得规则集合为

$$\{CO_2 = 1 \rightarrow sec_2, CO_3 = 1 \rightarrow sec_3, CO_1 = 1 \rightarrow sec_1, CB_1 = 1, RR_1 = 0 \rightarrow sec_1\}.$$

设置故障信息覆盖度的阈值为3.7%, 置信度阈值为100%, 可得规则集合为

$$\{CO_2 = 1 \rightarrow sec_2, CO_3 = 1 \rightarrow sec_3, CO_1 = 1 \rightarrow sec_1, CB_1 = 1, RR_1 = 0 \rightarrow sec_1\}.$$

设置故障信息覆盖度的阈值为 1.2%，置信度阈值为 100%，得规则集合为

$$\{CO_2 = 1 \rightarrow sec_2, CO_3 = 1 \rightarrow sec_3, CO_1 = 1 \rightarrow sec_1, \\ CB_1 = 1, RR_1 = 0 \rightarrow sec_1, \\ CB_1 = 0, CO_1 = 0, CO_2 = 0, CO_3 = 0 \rightarrow NO, \\ CB_1 = 1, CO_1 = 0, CO_2 = 0, CO_3 = 0 \rightarrow sec_{2/3}, \\ CB_1 = 1, RR_1 = 1, CO_2 = 0, CO_3 = 0 \rightarrow sec_{2/3}, \\ CO_1 = 0, RR_1 = 1, CO_2 = 0, CO_3 = 0 \rightarrow sec_{2/3}, \\ CB_1 = 0, RR_1 = 0, CO_2 = 0, CO_3 = 0 \rightarrow NO, \\ CO_1 = 0, RR_1 = 0, CO_2 = 0, CO_3 = 0 \rightarrow NO\}.$$

由此可知，故障信息覆盖度选择的阈值越小，得到的规则越细，规则的分级越低。在实际计算中，按照由高到低的顺序依次得出规则。文献 [13] 中最终得到的决策规则集合为

$$\{CO_2 = 1 \rightarrow sec_2, CO_3 = 1 \rightarrow sec_3, CO_1 = 1 \rightarrow sec_1, \\ CB_1 = 0, CO_2 = 0, CO_3 = 0 \rightarrow NO, \\ CB_1 = 1, CO_1 = 0, CO_2 = 0, CO_3 = 0 \rightarrow sec_{2/3}\}.$$

3.3 与已有文献的结果对比

比较本文方法与文献 [13] 可以发现，按照本文的设置阈值方法得到的规则集合比较整齐有规律。依据不同的阈值设置得到不同层次的规则集合，阈值为 $x\%$ 的规则集合中包含了所有具有更高代表性的高级规则，舍弃的是影响较小的不能满足阈值为 $x\%$ (代表性低于 $x\%$) 的低级规则；而文献 [13] 舍弃的规则中既有二级规则，又有三、四级规则，且所得到的规则集合中的三级规则 $CB_1 = 0, CO_2 = 0, CO_3 = 0 \rightarrow NO$ ，有可能与一级规则 $CO_1 = 1 \rightarrow sec_1$ 冲突。

对比结果可以发现，本文算法中给出的依照阈值分级进行规则提取的分类方法层次清晰，避免了上述冲突，得到了更加完备的规则集合。

4 结 论

本文采用基于图论的粗糙集算法，对由实际故障所得到的故障诊断决策表进行属性约简，可以方便得出核属性，并根据核属性得出相应的属性约简。在故障规则提取方面，给出了决策规则的故障信息覆盖度定义和决策规则分级，按照级别由高到低从核属性依次提取规则，在避免重复计算的同时，可以根据设置的阈值进行规则的选择。实验表明，本文算法条理清晰，计算简便，能有效地减少时间和空间复杂度。

参考文献(References)

[1] Febriyanto A, Saba T. Oil-immersed power transformers condition diagnosis with limited dissolved gas analysis (DGA) data[C]. Universities Power Engineering Conf.

Sydney, 2008: 1-5.

- [2] Ahfaz Khan M, Sharma A K, Saxena R. Expert system for power transformer condition monitoring and diagnosis[C]. IEEE Power Electronics Drives and Energy Systems. New Delhi, 2006: 1-6.
- [3] 郑蕊蕊, 赵继印, 赵婷婷, 等. 基于遗传支持向量机和灰色人工免疫算法的电力变压器故障诊断[J]. 中国电机工程学报, 2011, 31(7): 56-63.
(Zheng R R, Zhao J Y, Zhao T T, et al. Power transformer fault diagnosis based on genetic support vector machine and gray artificial immune algorithm[J]. Proc of the CSEE, 2011, 31(7): 56-63.)
- [4] Monsef H, Ranjbar A M, Jadid S. Fuzzy rule-based expert system for power system fault diagnosis[J]. IEEE Proc of Generation, Transmissinn and Distribution, 1997, 144(2): 186-192.
- [5] 朱永利, 宋少群. 基于广域网和多智能体的自适应协调保护系统的研究[J]. 中国电机工程学报, 2006, 26(16): 15-20.
(Zhu Y L, Song S Q. Study on multi-agent and WAN based adaptive coordinated protection system[J]. Proc of the CSEE, 2006, 26(16): 15-20.)
- [6] 张腾飞, 王锡淮, 叶银忠, 等. 粗糙集理论在故障诊断中的应用综述[J]. 上海海事大学学报, 2005, 26(4): 20-25.
(Zhang T F, Wang X H, Ye Y Z, et al. Survey on application of rough set theory in fault diagnosis[J]. J of Shanghai Maritime University, 2005, 26(4): 20-25.)
- [7] Skowron A, Rauszer C. The discern ability matrices and functions in information systems[C]. Intelligent Decision Support Handbook of Applications and Advances of the Rough Sets Theory. Kluwer: Academic Publishers, 1992: 331-362.
- [8] 彭文季, 罗兴铨. 基于粗糙集和支持向量机的水电机组振动故障诊断[J]. 电工技术学报, 2006, 21(10): 117-122.
(Peng W J, Luo X Q. Vibration fault diagnosis of hydro-rurbine generating unit based on rough sets and support vector machine[J]. Trans of China Electrotechnical Society, 2006, 21(10): 117-122.)
- [9] 徐金良, 陈五星, 唐耀阳, 等. 基于粗糙集理论和支持向量机算法的核电厂故障诊断方法[J]. 核动力工程, 2009, 30(4): 52-54.
(Xu J L, Chen W X, Tang Y Y, et al. Study on fault diagnosis in nuclear power plant based on rough sets and support vector machine[J]. Nuclear Power Engineering, 2009, 30(4): 52-54.)
- [10] Xiang Xiuqiao, Zhou Jianzhong, Li Chaoshun, et al. Fault diagnosis based on Walsh transform and rough sets[J]. Mechanical Systems and Signal Processing, 2009, 23(4): 1313-1326.