

文章编号: 1001-0920(2011)08-1273-04

## 基于二分图的故障规则匹配优化算法

牛强, 夏士雄, 胡祖辉

(中国矿业大学 计算机学院, 江苏 徐州 221116)

**摘要:** 针对传统的基于相似度的故障规则匹配方法中未考虑输入条件与规则前件的整体匹配程度问题, 采用二分图最优匹配方法对匹配过程进行优化, 提出一种基于二分图的故障规则匹配优化算法, 并将其应用于故障诊断推理. 实例分析表明, 与其他相似度匹配算法相比, 所提出的方法有效提高了规则匹配的准确率, 而且降低了时间消耗.

**关键词:** 二分图; 相似度; 模糊推理; 规则匹配

中图分类号: TP18

文献标识码: A

### Fault rule matching optimization algorithm based on bipartite graph

NIU Qiang, XIA Shi-xiong, HU Zu-hui

(School of Computer Science Technology, China University of Mining Technology, Xuzhou 221116, China.  
Correspondent: NIU Qiang, E-mail: niuqiang@vip.163.com)

**Abstract:** According to traditional fault rule matching method based on similarity taking no account of overall matching degree between input conditions and rule antecedents, bipartite graph optimal matching method is adopted to optimize the matching process in this paper. Then it is applied to fault diagnosis reasoning, and fault rule matching optimization algorithm based on bipartite graph is proposed. The example analysis shows that, compared with other similarity matching algorithms, the proposed algorithm effectively improves the accuracy of matching rules and reduces the matching time meanwhile.

**Key words:** bipartite graph; similarity; fuzzy reasoning; rule matching

## 1 引言

模糊推理最早由 Zadeh 提出, 随着模糊控制在越来越多领域获得成功的应用, 关于模糊推理和模糊系统的研究取得了迅速发展. 1973年, Zadeh 首先给出了模糊推理理论中最基本的推理规则——模糊分离规则(FMP), 并提出了求解 FMP 问题的合成推理方法(CRI)<sup>[1]</sup>. 在模糊推理应用中, CRI 方法是最基本的方法, 影响非常广泛, 很多模糊推理方法都是基于 CRI 方法的, 如 Mamdani 方法, Mizumoto 方法等.

基于相似度的模糊推理方法采用与 CRI 方法完全不同的推理机制, 利用相似度来表示模糊命题的接近程度. 只有当输入条件与模糊规则前件的相似程度大于阈值时, 规则才会被激活. 这种推理思路符合领域专家运用知识分析解决问题的一般思维方式, 具有理论与现实的合理性. 基于相似度的模糊推理方法比 CRI 方法更容易理解, 计算更方便, 且满足还原性. 将基于相似度的模糊推理方法应用于故障诊断推理, 可以通过故障规则匹配得到推理结果. 本文对基于相似

度的故障规则匹配方法进行研究, 提出一种基于二分图的故障规则匹配优化算法. 实例分析表明了该方法的有效性.

## 2 故障规则匹配

故障规则采用产生式方法统一表示成“**If** (前件) **Then** (后件)”的形式. 这里的规则前件可以是单个命题, 也可以是多个命题组成的逻辑表达式, 多个命题之间用逻辑与“**and**”(或合取符号  $\wedge$ )、逻辑或“**or**”(或析取符号  $\vee$ ) 连接.

**定义 1** (故障规则)

$$R: \text{If } A \text{ Then } B, \text{ CF} = \mu, \lambda.$$

其中:  $A$  是前件命题 ( $A$  可以是单个命题, 也可以是复合命题),  $B$  是后件命题,  $\mu \in [0, 1]$  是规则  $R$  的可信度因子 (CF),  $\lambda \in [0, 1]$  是规则的阈值.

根据基于相似度的模糊推理方法的原理, 只有当输入条件与模糊规则前件的相似程度大于阈值时, 规则才会被激活. 对于基于相似度的故障诊断推理, 首

收稿日期: 2010-05-13; 修回日期: 2010-06-29.

基金项目: 国家自然科学基金项目(50674086).

作者简介: 牛强(1974—), 男, 副教授, 博士, 从事数据挖掘与故障诊断等研究; 夏士雄(1961—), 男, 教授, 博士生导师, 从事数据处理与信息融合等研究.

先给出下面两个定义:

**定义 2 (匹配操作)** 给定一个输入条件  $X$  和一条规则  $R = \{A, B, \mu, \lambda\}$  ( $A$  为规则前件,  $B$  为规则后件,  $\mu$  为可信度,  $\lambda$  为阈值),  $X$  与  $A$  之间的相似度记为  $\text{Sim}(X, A)$ . 如果满足  $\text{Sim}(X, A) \geq \lambda$ , 则激活规则  $R$ , 输出规则后件  $B$ , 这样的操作称为匹配操作.

**定义 3 (匹配度)** 给定一个输入条件  $X$  和一条规则  $R = \{A, B, \mu, \lambda\}$ , 定义  $X$  与  $R$  的匹配度为  $X$  与  $A$  之间的相似度, 匹配度的值为  $\text{Sim}(X, A)$ .

基于相似度的故障诊断推理的总体思路是: 首先计算故障征兆与故障规则前件之间的匹配度, 只有当故障征兆与故障规则前件的匹配度大于阈值时, 规则才会被激活, 输出故障原因, 从而实现故障诊断推理.

根据以上分析, 基于相似度的故障诊断推理的关键是相似度计算方法和故障规则匹配方法. 传统的方法往往忽略了规则匹配的方法, 仅简单地计算集合之间的相似度. 典型的方法如: 海明距离相似度, 近似类比推理模式 (AARS), 最大模糊 (MF) 相似度等<sup>[2-4]</sup>, 均未考虑输入条件与模糊规则前件的整体匹配程度. 对此, 本文引入二分图来计算两个集合之间的整体匹配程度, 以优化故障规则匹配策略.

### 3 基于二分图的故障规则匹配方法

#### 3.1 二分图概述

下面给出二分图相关概念的定义<sup>[5]</sup>:

**定义 4 (二分图)** 如果无向图  $G = (V, E)$  的顶点集合  $V$  可分为两个集合  $V_1$  和  $V_2$ , 满足  $V_1 \cup V_2 = V$ ,  $V_1 \cap V_2 = \emptyset$ , 且对于任意一条边  $e = (v_i, v_j) \in E$ , 均有  $v_i \in V_1, v_j \in V_2$ , 则称  $G$  为二分图, 并称  $V_1$  和  $V_2$  为  $G$  的互补顶点子集.

**定义 5 (完全二分图)** 如果二分图  $G$  的互补顶点子集  $V_1$  的每一个顶点都与  $V_2$  中的所有顶点邻接, 则称  $G$  为完全二分图<sup>[6]</sup>.

常用  $K_{m,n}$  表示一个完全二分图, 其中:  $m = |V_1|, n = |V_2|$ , 显然完全二分图有  $m \times n$  条边. 定义故障征兆概念集合为  $C = \{c_1, c_2, \dots, c_m\}$ , 故障规则前件概念集合为  $V = \{v_1, v_2, \dots, v_n\}$ . 构造  $C$  与  $V$  之间的匹配关系为一个无向完全图, 两组概念结点集合之间相互两两连接, 概念结点集合内部任意两个结点都不相连. 显然,  $C$  与  $V$  之间的匹配关系图是一个完全二分图.

#### 3.2 故障规则匹配优化方法

根据以上分析, 可将故障规则匹配问题建模成一个带权二分图  $G = (C, V, E)$ . 其中:  $C$  和  $V$  为  $G$  的两个互补的顶点子集, 分别对应上述故障征兆概念集合  $C = \{c_1, c_2, \dots, c_m\}$  和规则前件概念集合  $V =$

$\{v_1, v_2, \dots, v_n\}$ ; 边集  $E$  可按如下规则构造: 对于  $\forall c_i \in C, \forall v_j \in V$ , 若  $\text{Sim}(c_i, v_j) \geq 0$ , 则在二分图  $G$  中  $c_i$  和  $v_j$  对应的两个顶点之间连一条边  $(c_i, v_j)$ , 并给该边赋一个权重  $\omega_{i,j} = \text{Sim}(c_i, v_j)$ .

通过二分图建模之后, 故障规则匹配问题便转化为在二分图  $G$  上求解从顶点集合  $C$  到  $V$  的一个最优匹配  $M$ , 使得  $M$  能覆盖集合  $C$  中所有节点, 同时要求匹配  $M$  的边权之和最大.

对于带权二分图的最优匹配问题, 本文采用求解带权二分图最优匹配的经典算法 Kuhn-Munkres (KM) 来求解<sup>[7]</sup>. KM 算法要求输入是两个互补顶点子集顶点数目相等的完全二分图, 即要求输入不仅是完全二分图, 而且还要满足条件  $|C| = |V|$  (即  $m = n$ ). 然而在故障规则匹配过程中, 故障征兆概念的个数与故障规则前件概念的个数通常是不等的, 因此有必要对二分图进行转换. 不妨设  $|C| < |V|$ , 此时在  $X$  集合中增加了  $|V| - |C|$  个虚拟顶点  $Z = \{z_1, z_2, \dots, z_k\}, k = |V| - |C|$ , 并且在新增加的每一个顶点与集合  $V$  中的每一个顶点之间增加一条权重为 0 的边. 二分图的转换过程如图 1 所示, 图中空心顶点表示虚拟顶点, 虚线表示权重为 0 的边.

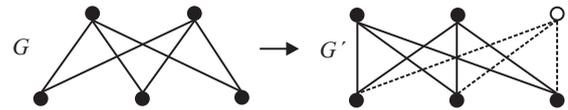


图 1 二分图转换过程

由图 1 可见, 二分图  $G$  经过转化之后变成  $G' = (C \cup Z, V, E')$ . 由于  $|C \cup Z| = |V|$ , 此时可使用 KM 算法为  $G'$  求得最优匹配  $M'$ , 之后再通过去除  $M'$  中覆盖了虚拟顶点的所有边而将  $M'$  转化成  $M$ . 对此, 文献 [7] 提出了扩展最优匹配的概念, 并证明了通过这种转换求出的最优匹配即为原图的最大权匹配. 因此, 对于故障规则匹配问题也可以理解为最大相似度匹配, 将最大相似度匹配中边的权重求和并进行归一化处理之后, 作为故障征兆概念集合与故障规则前件概念集合之间的匹配度, 符合故障诊断推理的要求.

根据以上分析, 本文提出基于二分图的故障规则匹配优化算法, 具体描述如下:

**算法 1** 基于二分图的故障规则匹配优化算法.

输入: 单个概念单元之间的相似度  $\text{Sim}(c_i, v_j)$ ;

输出: 规则匹配度 RuleMatch.

构造二分图  $G = (C, V, E)$ ,

If  $(|C| == |V|)$ ,

{

    调用 KM 算法计算  $G$  的最优匹配  $M$ ;

}

Else if  $(|C| < |V|)$ ,

```

{
  转换二分图  $G(C, V, E)$  到  $G'(C', V, E')$ , 其中
   $C' = C \cup Z, Z = \{z_1, z_2, \dots, z_k\},$ 
   $k = |V| - |C|;$ 
   $E' = E \cup \{(z_i, v_j) | 1 \leq i \leq |Z|,$ 
   $1 \leq j \leq |V|, \omega_{i,j} = 0\};$ 
  调用 KM 算法计算  $G'$  的最优匹配  $M'$ ;
  将  $M'$  转换为  $M$ : 去除匹配  $M'$  中权重为 0 的
  边;
}
Else if ( $|C| > |V|$ ),
{
  转换二分图  $G(C, V, E)$  到  $G'(C', V, E')$ , 其中
   $V' = V \cup Z, Z = \{z_1, z_2, \dots, z_k\},$ 
   $k = |C| - |V|;$ 
   $E' = E \cup \{(c_i, z_j) | 1 \leq i \leq |C|,$ 
   $1 \leq j \leq |Z|, \omega_{i,j} = 0\};$ 
  调用 KM 算法计算  $G'$  的最优匹配  $M'$ ;
  将  $M'$  转换为  $M$ : 去除匹配  $M'$  中权重为 0 的
  边 (即覆盖了虚拟顶点的边);
}

```

$$\text{RuleMatch} = \frac{M \text{ 中边的权重之和}}{M \text{ 中边的数目}};$$

返回 RuleMatch.

KM 算法经过松弛优化的时间复杂度为  $O(n^3)^{[8]}$ , 其中  $n = |C| = |V|$ , 因而算法的时间开销为常数, 运行效率较高.

### 4 实验测试与分析

对传统的匹配度故障规则推理匹配算法和基于二分图的故障诊断规则匹配优化算法进行比较, 主要通过相同的实验数据和实验环境下对相同的规则进行匹配, 对规则匹配的准确率和时间运行效率进行比较.

本实验在 PC-CPU 3.0, 1 G 内存和 Windows XP 操作系统上运行. 首先, 实验数据使用异步电动机的定子和转子及其故障征兆和故障原因之间的相关数据. 在异步电动机的定子和转子发生故障时, 同一种故障现象可能对应多种故障原因; 同一种故障原因又可能对应多种故障现象. 因此, 为了能从故障现象中诊断出电动机发生故障的原因, 需要预先确定出故障现象与故障原因之间的关系<sup>[9]</sup>. 所包括的故障征兆有 {电机振动  $x_1$ , 噪声  $x_2$ , 温度高  $x_3$ , 局部放电  $x_4$ , 有噪声  $x_5$ , 电流摆动  $x_6$ , 转速降低  $x_7$ , 转矩减速  $x_8$ }; 故障原因有 {定子转子相摩擦  $v_1$ , 转子开焊  $v_2$ , 接缝松动

$v_3$ , 定子绕组接地  $v_4$ , 定子绕组断线  $v_5$ , 定子绕组短路  $v_6$ , 定子绕组接错  $v_7$ , 绕组破损  $v_8$ }. 然后, 采用基于多重多维 Fuzzy 推理模型的 Mamdani 算法提取出来的规则, 使用不同的匹配度算法和本文提出的方法进行故障规则匹配, 从而实现规则匹配准确率和时间运行效率的比较.

例如检测到的故障现象分别为: 电机振动  $c_1$ , 温度高  $c_3$ , 局部放电  $c_4$ , 有噪音  $c_5$ , 即故障征兆集合概念为  $C = \{c_1, c_3, c_4, c_5\}$ . 根据故障知识库或专家经验给出故障现象与产生该故障原因之间的对应模糊隶属度, 如表 1 所示.

表 1 故障征兆-原因隶属度

	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$	$v_7$	$v_8$
$c_1$	0.32	0.21	0.49	0	0.53	0.26	0.91	0.74
$c_3$	0.72	0	0.38	0.56	0.21	0	0.65	0.11
$c_4$	0.24	0.39	0.89	0.46	0.82	0	0.76	0.59
$c_5$	0.83	0.51	1.00	0.22	0.47	0.66	0.94	0.33

在计算本文提出的二分图匹配度时, 如果故障规则前件概念集合为  $V = \{v_1, v_3, v_5, v_7, v_8\}$ ,  $C$  与  $V$  单个概念单元之间的匹配度  $\text{Sim}(c_i, v_j)$  如表 2 所示.

表 2  $C$  与  $V$  概念单元之间的相似度

	$v_1$	$v_3$	$v_5$	$v_7$	$v_8$
$c_1$	0.32	0.49	0.53	0.91	0.74
$c_3$	0.72	0.38	0.21	0.65	0.11
$c_4$	0.24	0.89	0.82	0.76	0.59
$c_5$	0.83	1.00	0.47	0.94	0.33

下面采用基于二分图的故障规则匹配优化算法来计算匹配度, 具体计算过程如下.

1) 构造带权二分图  $G$ , 如图 2 所示.

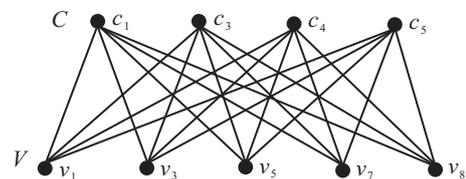


图 2 构造二分图模型

2) 由于  $|C| = 4, |V| = 5, |C| < |V|$ , 对二分图进行转换, 添加虚拟顶点和权重为 0 的边. 转换后的二分图为  $G'$ , 如图 3 所示.

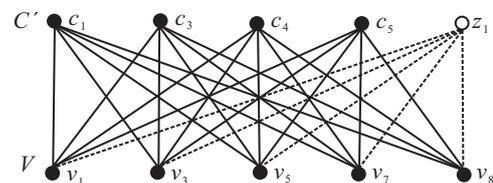


图 3 二分图转换

3) 应用 KM 算法求  $G'$  的最优匹配为

$$M' = \{c_1 v_7, c_3 v_1, c_4 v_5, c_5 v_3, z_1 v_8\}.$$

去除权重为0的边  $z_1v_8$  后,得

$$M = \{c_1v_7, c_3v_1, c_4v_5, c_5v_3\}.$$

#### 4) 计算语义匹配度

$$\text{SemanticMatch} =$$

$$(0.91 + 0.72 + 0.82 + 1.00)/4 = 0.86.$$

通过以上方法即可计算故障征兆与故障知识库中各项规则前件之间的匹配程度.一旦匹配度大于规则阈值,便触发规则,得出故障原因.

实验中,将本文匹配推理方法与基于海明距离的匹配度, AARS 匹配度<sup>[3]</sup>和 MF 匹配度<sup>[4]</sup>的相似度推理方法进行了比较,规则匹配效率结果如图4所示.其中:横坐标为输入故障现象的数量,纵坐标为故障规则匹配准确率(单位: %).从图4可以看出,随着输入故障现象数量的增加,本文方法的规则匹配准确率在总体上优于传统的相似度规则推理算法.

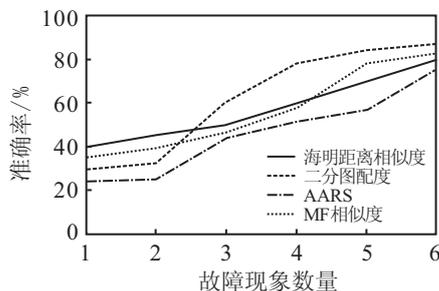


图4 几种相似度推理准确率比较

图5给出了规则匹配时间效率结果.其中:横坐标为输入故障现象的数量,纵坐标为故障规则匹配所耗时间.可以看出,与传统的相似度推理方法相比,本文方法规则匹配所消耗的时间有所降低,而且随着输入故障现象数量的增加,本文方法的规则匹配所消耗的时间趋于平稳,不会像传统的相似度规则推理算法出现较大的波动.

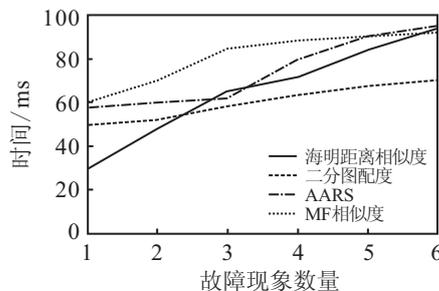


图5 几种相似度推理消耗时间的比较

## 5 结 论

本文提出了基于二分图的故障规则匹配优化算

法,应用二分图的最优匹配算法对故障规则匹配过程进行了优化,并通过实验对几种相似度推理方法进行了比较分析.实验结果表明,基于二分图的故障规则匹配优化算法可以有效提高规则匹配的准确率,并降低规则匹配消耗的时间.

## 参考文献(References)

- [1] Zadeh L A. Outline of a new approach to the analysis of complex systems and decision processes[J]. IEEE Trans on Systems, Man, and Cybernetics, 1973, 3(1): 28-44.
- [2] 罗承忠. 模糊集引论[M]. 北京: 北京师范大学出版社, 1989.  
(Luo C Z. Introduction to fuzzy set[M]. Beijing: Beijing Normal University Press, 1989.)
- [3] Turksen I B, Zhao Z. An approximate analogical reasoning approach based on similarity measures[J]. IEEE Trans on Systems, Man, and Cybernetics, 1988, 18(6): 1049-1056.
- [4] Chen S M. A new approach to handling fuzzy decision-making problems[J]. IEEE Trans on Systems, Man, and Cybernetics, 1988, 18(6): 1012-1034.
- [5] 肖位枢. 图论及其算法[M]. 北京: 航空工业出版社, 1993.  
(Xiao W S. Graph theory and its algorithm[M]. Beijing: Aviation Industry Press, 1993.)
- [6] Wang Y, Makedon F, Ford J. A bipartite graph matching framework for finding correspondences between structural elements in two pteins[C]. Proc of 26th Annual Int Conf on IEEE Engineering in Medicine and Biology Society. San Francisco: IEEE Press, 2004: 2972-2975.
- [7] 邓水光, 尹建伟, 李莹, 等. 基于二分图匹配的语义 web 服务发现方法[J]. 计算机学报, 2008, 31(8): 1364-1375.  
(Deng S G, Yin J W, Li Y, et al. A method of semantic web service discovery based on bipartite graph matching[J]. Chinese J of Computers, 2008, 31(8): 1364-1375.)
- [8] 刘晓慧, 佟伟光, 林树宽. 基于模糊推理的专家系统的研究与实现[J]. 沈阳工程学院学报: 自然科学版, 2007, 3(2): 162-165.  
(Liu X H, Tong W G, Lin S K. Study and realization of expert system based on fuzzy reasoning[J]. J of Shenyang Institute of Engineering: Natural Science, 2007, 3(2): 162-165.)
- [9] Lovasz L, Plummer M. Matching theory[M]. Amsterdam: North-Holland, 1986.