

一种基于模式的故障诊断方法

崔海春^{1,2} 欧阳丹彤^{1,2} 王晓宇^{1,2} 纪树平³

(吉林大学计算机科学与技术学院 长春 130012)¹

(吉林大学符号计算与知识工程教育部重点实验室 长春 130012)²

(中国科学院软件研究所软件工程实验室 北京 100086)³

摘要 提出一种改进的基于模式的故障诊断方法。将故障定义为可以用正则表达式描述的模式,避免了传统的用“特殊”事件表示故障的局限性。将正则表达式转化成等价的确定性有限自动机,便于构造故障模式空间。故障模式空间包含系统所有的故障信息,可以描述单故障、故障序列、多故障、间歇性故障等多种故障形式。最后给出基于故障模式空间的诊断性定义和诊断系统。

关键词 系统模型,故障模式,可诊断性,诊断问题

中图分类号 TP31 文献标识码 A

Approach of Fault Diagnosis Based on Patterns

CUI Hai-chun^{1,2} OUYANG Dan-tong^{1,2} WANG Xiao-yu^{1,2} JI Shu-ping³

(College of Computer Science and Technology, Jilin University, Changchun 130012, China)¹

(Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China)²

(Institute for Software Technology, Chinese Academy of Sciences, Beijing 100086, China)³

Abstract An improved method of fault diagnosis based on patterns was proposed. Faults are defined as patterns that are described by regular expressions to avoid the limitation of traditionally defining faults which uses ‘special’ events. Regular expressions can be equivalently transformed into deterministic finite automaton, so that fault pattern space can be easily constructed. All the information of faults can be contained in the fault pattern space, which include fault sequences, multiple faults, and intermittent faults. Finally, the definition of diagnosability based on fault pattern space was given and a diagnostic system was implemented.

Keywords System model, Fault patterns, Diagnosability, Diagnosis problem

1 引言

离散事件系统在动态设备故障诊断中扮演着重要的角色,这是因为离散事件系统在各个状态接近静态系统,许多混合或连续的动态系统又可以使用离散事件系统近似表示。采用基于模型的推理方法对离散事件系统进行故障诊断也取得了较多研究成果,有基于语法的^[1-4],如自动机、Petri网;也有基于语义的^[5-7],如时序逻辑、语义规则。不同文献中,系统故障有不同的定义方法,经典的故障诊断方法中将故障定义为一些特殊的事件,然后利用诊断算法探测这些特殊事件来判断故障是否发生;此外还有将故障定义为与正常行为有偏差的系统行为,用线性时序逻辑描述系统的正常行为模型,称为规范语言。所有不符合规范语言的系统行为被定义为故障行为^[7];文献[4]给出了故障的另一种新定义,即将每一类系统

故障定义为一种模式,故障描述是基于自动机实现的。

在传统的诊断方法中,将故障定义为发生在部件层的某些特殊事件,导致出现了各种各样的可诊断性定义,相关的定义及算法并不能在类似的诊断问题中重新利用,没有实现诊断问题与系统描述的进一步分离^[4],而且故障种类比较单一,在实际大型的系统中,故障定义不应受系统事件集的约束。文献[4]中将故障定义为完备有限自动机使系统故障行为为多样化^[11],所谓完备自动机是指,自动机的每一个状态都可以被系统任意事件触发到另一个状态^[4]。另外将故障模式用自动机描述,可以同步得到多故障、间歇性故障等不同的故障形式。这种用完备自动机描述故障模式的方法存在以下不足。首先将故障模式描述为完备的自动机只是基于故障的语法形式,而并没有从语义上进一步提取系统故障信息,将系统故障与系统模型完全分离,与实际不符;其次不同故障情形需要单

到稿日期:2012-01-20 返修日期:2012-06-10 本文受国家自然科学基金(60973089,60873148,60773097,61003101),吉林省科技发展计划项目基金(20101501,20100185,20090108,20080107,201101039),国家教育部博士点专项基金(20100061110031),浙江师范大学计算机软件与理论省级重中之重学科开放基金,浙江省自然科学基金(Y1100191),欧盟合作项目(155776-EM-1-2009-1-IT-ERAMUNDUS-ECW-L12),吉林大学符号计算与知识工程教育部重点实验室开放项目(93K-17-2009-K05)资助。

崔海春(1988-),女,硕士生,主要研究方向为基于模型诊断,E-mail:cuihaichun8@163.com;欧阳丹彤(1968-),女,博士,教授,博士生导师,主要研究方向为基于模型的诊断、自动推理和模型检测。

独处理,通过穷举的方式才能列出所有的故障情形,例如当求解含有某两种故障的故障模式时,需要将这两种故障模式的自动机同步得到,显然这种方法太复杂。尤其当系统规模很大、故障种类很多时,很难得到所有的故障情形。

结合文献[5]中基于语义故障模式的描述方法,本文提出一种改进的基于故障模式的诊断方法,每一种故障模式用定义在系统事件集上的正则表达式进行描述。这种描述方法使得故障的发生不再由一个转移引起,可能由多个转移共同导致,体现了系统故障的发生是渐变的。由于正则表达式以其强大灵活的表达能力而迅速成为描述新一代规则的主要工具,开源社区和商业界都开始广泛使用正则表达式来增强协议特征和安全特征的描述。所以用正则表达式表示故障模式,使得故障描述更加简单,且更容易处理。此外正则表达式可以转化成等价的有限自动机^[9],对应故障模式的识别过程,而且现有的正则表达式匹配技术比较成熟,使得快速精确的诊断故障成为可能。文中改进了构造故障模式空间的算法,提供了一个可以包含系统所有故障信息、所有故障情形的模式空间,避免文献[4]中用穷举的方法求解各种故障情形的高复杂性;此外给出基于故障模式空间的可诊断性定义,并分析了这种定义方法的正确性;最后给出计算候选诊断的算法。

本文第2节给出一些相关概念;第3节详细叙述故障模式空间的构建过程以及作用;第4节给出基于故障模式的可诊断性定义、诊断问题描述以及候选诊断求解;第5节分析算法优缺点及比较实验运行结果;最后进行总结和展望。

2 背景知识

定义 1(系统模型^[3]) 系统模型 M 定义为一个有限状态机 $G = \langle Q, \Sigma, T, q_0 \rangle$, 其中 Q 是状态集; Σ 是所有发生在系统中的事件集; $T \subseteq Q \times \Sigma \times Q$ 表示状态转移函数集; q_0 是初始状态。 Σ 分为可观测事件集 Σ_o 和不可观测事件集 Σ_{no} , 且 $\Sigma_o \cap \Sigma_{no} = \emptyset$ 。所有从初始状态发出的事件序列集为 M 生成的语言, 记为 $L(M)$ 。任一个事件序列 $s \in L(M)$, 路径 $Path(s)$ 表示事件序列 s 所经过的有序状态集。例如系统模型为图 1(a), 当事件序列 s 为 $T1T3T3T2$ 时, $Path(s) = \langle q_0, q_1, q_1, q_2 \rangle$ 。

定义 2(可观测事件) 若 $\exists o \in \Sigma$, 并且 o 可被传感器探测且传输到系统外部被观测到, 则称 o 是可观测事件。

投影函数 $map: \Sigma^* \rightarrow \Sigma_o^*$ 定义为 $map(s\sigma) = map(s)\sigma$, 如果 $\sigma \in \Sigma_o$; 否则 $map(s\sigma) = map(s)$ 。不失一般性, 对语言 $L \subseteq \Sigma^*$, $map(L) = \{map(s) | s \in L\}$ 。

值得区分的是, 观测和可观测事件不同。观测与系统的实际运行有关, 而可观测事件是系统的属性, 在系统的一次运行中, 只有被传感器探测到的可观测事件才称为观测。由于观测实际由系统部件发出, 从每个部件获得的观测序列之间并没有包含任何时序信息, 通过构建观测空间可以表示所有观测之间的时序关系^[2]。

定义 3(延迟可观测等价性, 记为 Ξ_M ^[4]) 设 $L(M)$ 是系统 M 生成的语言, $\Xi_M \subseteq L(M) \times L(M)$ 是一个二元关系, $\forall s, s' \in L(M)$, $s \Xi_M s'$ iff $map(s) = map(s')$ 。

很显然, Ξ_M 是一种等价关系, $[s] = \{l | map(l) = map(s) \text{ and } l \in L(M)\}$ 表示包含 s 的等价类。

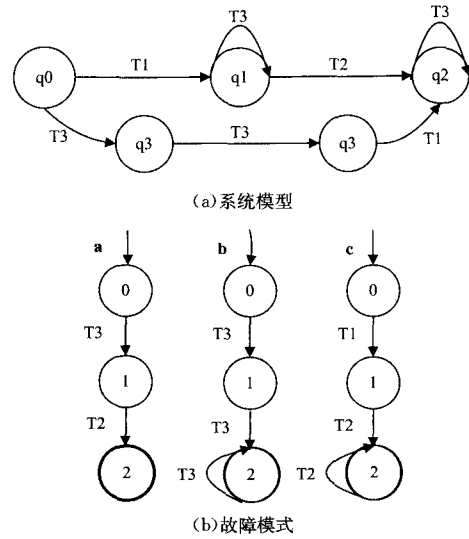


图 1

在系统模型 M 中, 每一种故障模式 P 由 M 中事件集合定义的正则表达式描述, P 所包含的事件个数表示故障模式的规模。例如图 1(a) 为系统模型, 可以定义 $T1T2^+$ 为一个规模为 2 的故障模式。正则表达式描述了一种字符串匹配的模式, 其有自己的语法规则, 功能非常灵活和强大, 定义的故障模式便于理解, 而且这种故障定义方式不受系统模型中事件数量的限制。与传统的定义方法相比, 其可以表示大量的故障信息。此外故障模式可以表示每一种故障发生是渐变的, 可能由系统的一系列行为共同导致。由于正则表达式可以转化成等价的有限自动机, 对应的模式匹配是基于自动机的, 因此可以通过匹配故障模式探测系统发生的故障。例如, 图 1(b) 为图 1(a) 对应的 4 类故障模式的识别过程, 其中 b 类故障的故障模式为 $T3T3T3^*$; a 类故障的故障模式为 $T3T2$; 等。黑色粗体标记的状态对应每类故障模式的接受状态。每一类故障模式 P 对应一个确定性有限状态机, 记为 $A_P = \langle Q_P, \Sigma_P, T_P, q_{P0}, Q_{PF} \rangle$, 其中 Q_{PF} 表示对应故障模式的接受状态。

3 故障模式空间

在诊断系统中, 根据当前的观测序列, 搜索系统可能的运行轨迹, 通过匹配对应的故障模式来判断系统发生哪类故障。在大型系统中, 还要考虑多故障、间歇性故障等多种故障情形, 所以有必要构建与系统模型对应的故障模式空间。

定义 4(合成算子, 记为 Θ) 设故障模式 P_1, P_2 对应的确定性有限状态自动机分别为 $A_1 = \langle Q_1, \Sigma_1, T_1, q_{10}, Q_{1F} \rangle$, $A_2 = \langle Q_2, \Sigma_2, T_2, q_{20}, Q_{2F} \rangle$, 故障标签集合 $LF = \{\epsilon, f1, f2\}$ 。 A_1, A_2 的合成记为 $A_1 \Theta A_2 = \langle Q, \Sigma, T, q_0 \rangle$, 其中 $\Sigma = \Sigma_1 \cup \Sigma_2$; $Q \subseteq Q_1 \times Q_2 \times LF$; $q_0 = (q_{10}, q_{20}, \epsilon)$; $T = \{(q_i, q_j, lf) \xrightarrow{t} (q_i', q_j', lf') | (\text{if } \exists t(q_i \xrightarrow{t} q_i' \text{ or } q_j \xrightarrow{t} q_j')) \text{ and } (\text{if } \exists t(q_i \xrightarrow{t} q_i' \text{ or } q_j \xrightarrow{t} q_j')) \text{ and } (\text{if } \neg \exists t(q_k \xrightarrow{t} q_k') \text{ then } q_k' = q_{k0}, \text{ 其中 } k=i, j) \text{ and } (\text{if } q_i' \in Q_{1F}, lf' = lf \cup \{f1\}) \text{ and } (\text{if } q_j' \in Q_{2F}, lf' = lf \cup \{f2\})\}$ 。

设 P_1, P_2, \dots, P_N 为与系统模型 M 有关的故障模式, A_1, A_2, \dots, A_n 分别对应各个故障模式的自动机表示, 与系统模型 M 对应的故障模式空间通过算法 1 实现。

算法 1 PATTERN_SPACE

```

1. INPUT:  $A_1, A_2, \dots, A_n$ 
2. OUTPUT: Pst
3. Begin
4.   Pst=NULL;
5.   FOR  $i=0$  TO  $K$ 
6.     Pst=Pst  $\Theta$   $A_i^*$ 
7.   Endfor
8. End

```

算法1是文献[5]中构建故障模式空间算法的改进,主要是去掉了原来算法的第二步,即一旦与当前模式不匹配时,就立即返回到初始状态,不管被触发事件的类型。而在原算法中,与当前模式不匹配时,需要分析被触发的事件是否属于通讯事件集,如果不属于,则等待;否则返回初始状态。由于原算法是针对分布式系统提出的,故障模式是基于部件的,构建故障模式空间时需要考虑通讯事件,而在集中式诊断系统中,不必考虑部件之间的链接关系。图1的故障模式通过算法1构建的模式空间如图2所示。从图2可以看到,生成的故障模式空间Pst包含8个状态、17个转移,状态000为故障模式空间的初始状态。

其中带有故障标签的状态称为故障模式空间的接受状态,图2中的接受状态为002,020,120,200。设故障模式P对应的故障标签为F,Pst中所有接受P的状态集记为 Q_F ,所有到达有故障标签F的状态轨迹集合记为 $L_F(Pts)$ 。显然任意轨迹 $s \in L_F(Pts(M)) \cap L(M)$ 被F类的故障模式识别,例如当系统的运行轨迹为“T3;T1:T2”时,发生c类故障;当 $Path(s) \cap Q_F \neq \emptyset$ 时,可判定系统发生F类故障,例如当系统运行轨迹 $s = \langle 'T1T2T2T3' \rangle$ 时, $Path = \langle \langle '000', '001', '002', '002', '110' \rangle \rangle$ 。后面介绍的可诊断性定义和诊断系统的构建与故障模式空间密切相关。

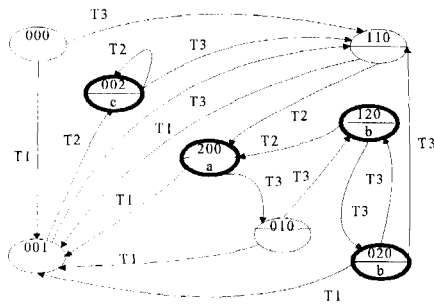


图2 故障模式空间

故障模式空间可以表示多故障、间歇性故障、单故障序列等所有的故障情形,例如当系统轨迹为T3T3T2时,根据状态中的故障标签,就可以判定系统发生a,b两类故障。而文献[4]中的处理方法是将表示a,b两类自动机的故障模式同步得到。当系统规模很大、故障模式的种类很多时,若再考虑故障之间的时序关系,则很难列出所有的故障情形,此外还需要大量构建新的自动机。

4 诊断问题

基于模型故障诊断的基本思想是给定系统的自动机模型和观测序列,由自动机模型的初始状态出发,重建与观测序列一致的所有可能的轨迹,然后从这些轨迹中找到所有可能的故障。同理,在基于故障模式的诊断系统中,诊断问题可以描述为当给定具体的系统模型、故障模式集合和观测时,求解候选诊断集合。

4.1 模式可诊断性分析

可诊断性作为系统模型的一种属性,一直被人们所关注。Sampath最早提出了离散事件系统的可诊断性定义,即f-diagnosability^[1],并给出了通过构建诊断器来判断系统是否可诊断的算法,这种判定算法的复杂性是双指数级的,在大型系统中很难得到应用。随后经典的Twin模型^[8]使系统的可诊断性判定算法降低到多项式级,但是随着系统规模的扩大,这种方法的可诊断性判定仍然很复杂;之后又出现分布式系统的可诊断性^[3]、分散式诊断方法^[10],这些方法将系统分为不同的子系统,通过分析子系统的可诊断性和子系统之间的关系来判定整个系统是否可诊断,但是子系统之间的关系需要协调器控制。随着系统故障定义方式的多样化,又出现了基于故障模式的可诊断性定义^[4],记为 Ω -diagnosability,以及关于可诊断性优化^[12]等。这里给出另一种基于故障模式的可诊断性定义,每一种故障模式用定义在系统事件集上的正则表达式描述。设系统模型M的事件集 $\Sigma = \{T1, T2, T3\}$,Pst_D是经过故障模式空间Pst扩充得到的完备自动机,称为完备的故障模式空间,如图3所示,显然 $L(Pst_D) = \Sigma^*$ 。

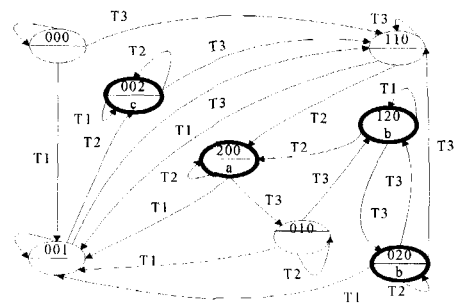


图3 扩展后的故障模式空间 Pst_D

定义5(模式可诊断性,记为P-diagnosability) 设P是完备故障模式空间,M为系统模型,如果系统是模式可诊断的,则 $\exists n \in \mathbb{Z}, \forall s \in L_F(P) \cap L(M), \forall t \in (L(M))/s \cap \Sigma_0, \text{if } \|P(t)\| \geq n, \text{then } \forall u \in [st], Path(u) \cap Q_F(P) \neq \emptyset$ 。

Ω -diagnosability的定义中每一个故障模式都是完备自动机, Ω -diagnosability实际上是f-diagnosability的自动机表示,文献[4]证明了两者是等价的。而P-diagnosability实际上是一种特殊的 Ω -diagnosability。我们给出的完备的故障模式空间也是一个完备自动机,整个完备的故障模式空间可以作为一个故障模式,与 Ω -diagnosability的定义相符。所以文献[4]中可诊断性判定算法在这里同样适用。

4.2 诊断问题求解

在故障模式空间构建中,系统可诊断性判定都是离线实现的,在系统运行中,首先根据实际的观测空间判定系统可能的运行轨迹,然后再同步故障模式空间得到候选诊断集合。假设系统模型和观测空间分别为: $M = \langle Q_M, \Sigma_M, T_M, q_{M0} \rangle$ 和 $O = \langle Q_O, \Sigma_O, T_O, q_{O0}, Q_{OF} \rangle$,得到的 Δ 为系统当前所有可能的运行轨迹,记为 $\Delta = \langle Q_\Delta, \Sigma_\Delta, T_\Delta, q_{\Delta0}, Q_{\Delta F} \rangle$,其中 $Q_{\Delta F}$ 表示与观测路径对应的接受状态集合。算法2实现搜索系统可能的运行轨迹,有限自动机 $Diagor = \langle Q_D, \Sigma_D, T_D, q_{D0}, Q_{DF} \rangle$ 表示候选诊断的计算过程,可以由算法3得到。

算法2 Generate_System_trajetory

```

1. INPUT:  $M = \langle Q_M, \Sigma_M, T_M, q_{M0} \rangle$ 
2.    $O = \langle Q_O, \Sigma_O, T_O, q_{O0}, Q_{OF} \rangle$ 
3. OUPUT:  $\Delta = \langle Q_\Delta, \Sigma_\Delta, T_\Delta, q_{\Delta0}, Q_{\Delta F} \rangle$ 

```

```

4. Initial;
5.  $q_{\Delta 0} = (q_{M0}, q_{O0}); Q_{\Delta} = \{q_{\Delta 0}\}; \Sigma_{\Delta} = \Sigma_M$ 
6. BEGIN
7. For each  $q_{\Delta} = (q_M, q_O) \in Q_{\Delta}$ 
8.   For each  $q_M \xrightarrow{t} q'_M$ 
9.     If  $t$  is visible
10.      and  $q_O \xrightarrow{t} q'_O \in T_O$  then
11.        $T_{\Delta} = T_{\Delta} \cup \{(q_M, q_O) \xrightarrow{t} (q'_M, q'_O)\}$ 
12.       Else  $q'_O = q_O$ 
13.        $T_{\Delta} = T_{\Delta} \cup \{(q_M, q_O) \xrightarrow{t} (q'_M, q'_O)\}$ 
14.       Endif
15.        $Q_{\Delta} = Q_{\Delta} \cup \{(q'_M, q'_O)\}$ 
16.       If  $q'_O \in Q_{OF}$  then
17.          $Q_{\Delta F} = Q_{\Delta F} \cup \{q_{\Delta}\}$ 
18.       Endif
19.     Endfor
20.   Endfor
21. delete( $Q_{\Delta}, T_{\Delta}$ )
22. END

```

算法 2 生成系统的运行空间, 函数 delete() 的主要功能是删除 Δ 中无用的状态和转移, 使得在计算候选诊断集时减少内存空间。

定义 6(同步算子, 记 \parallel) 设 $M = \langle Q_M, \Sigma_M, T_M, q_{M0}, Q_{MF} \rangle, P = \langle Q_P, \Sigma_P, T_P, q_{P0} \rangle$ 分别表示系统模型和模式空间 $M \parallel P = \langle Q, \Sigma, T, q_0, Q_F \rangle$, 其中 $Q \subseteq Q_M \times Q_P; \Sigma = \Sigma_M \cup \Sigma_P; T = \{(q_M, q_P) \xrightarrow{t} (q'_M, q'_P) \mid (\text{if } \exists t (q_M \xrightarrow{t} q'_M)) \text{ and } (\text{if } \exists t (q_P \xrightarrow{t} q'_P)) \text{ then } q'_P = q_{P0}) \text{ and } (\forall lf \in Lf(q_M, q_P), Lf(q'_M, q'_P) = Lf(q'_M, q'_P) \cup (lf, Lf(q'_P)))\}; Q_F = Q_{MF}$ 。

算法 3 Diagor_Construction

```

1. INPUT:  $M = \langle Q_M, \Sigma_M, T_M, q_{M0}, Q_{MF} \rangle$ 
2.  $P = \langle Q_P, \Sigma_P, T_P, q_{P0} \rangle$ 
3. OUTPUT:  $\text{Diagor} = \langle Q_D, \Sigma_D, T_D, q_{D0}, Q_{DF} \rangle$ 
4. BEGIN
5.  $\text{Diagor} = M \parallel P$ 
6. END

```

在线计算候选诊断的过程, 通过定义 6 实现, 所有状态集 Q_{DF} 中的故障标签为候选诊断集合。算法 2 和算法 3 改进了文献[5]中计算候选诊断的算法, 主要体现在先通过算法 2 获得当前的系统运行空间, 再删除一些对计算候选诊断无用的转移和状态, 以降低原算法中状态空间的复杂性。下面通过一个实例说明计算候选诊断集的过程。

4.3 运行实例

假设图 1(a) 是系统模型, 图 3(a) 是对应的观测空间, 按照算法 2 得到对应的系统运行空间, 如图 3(b) 所示, 虚线标记的状态和转移通过 delete() 函数删除。

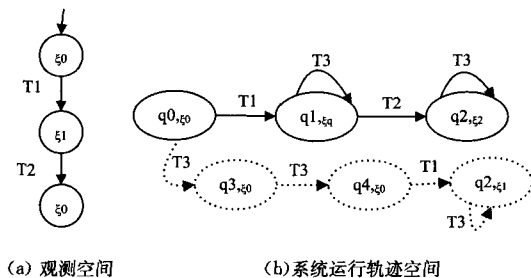


图 3

$T1, T2$ 是可观测事件, $T3$ 是不可观测事件, 系统的观测空间如图 4(a) 所示。利用算法 3, 得到的诊断过程如图 4(b) 中实线标记的状态和转移, 黑体标记的状态中故障标签为候选诊断, 显然有 6 个候选诊断, 即为 $\{\{a\}, \{c\}, \{a, b\}, \{b, c\}\}$ 。值得注意的是, 原算法中的诊断过程更复杂, 还包括图 4 中虚线标记的状态和转移, 显然虚线状态没有任何候选诊断, 因此在算法 2 中提前删除, 以降低复杂性。

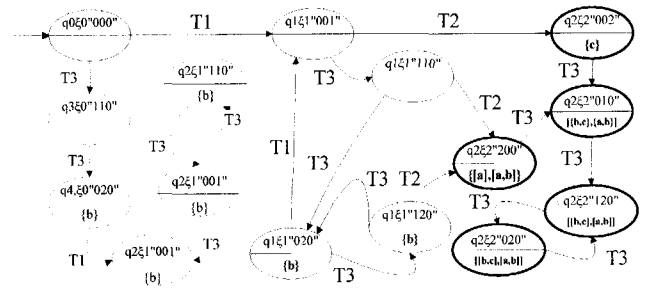


图 4 诊断过程

5 算法分析及实验结果

基于故障模式的可诊断性定义和诊断系统的构建都是基于故障模式空间的。与传统的基于模型诊断方法相比, 通过构建故障模式空间来表示系统故障信息有以下两点好处: 1) 可以根据系统的实际运行情况动态地增加或者删除故障类型, 进一步完善故障模式空间, 避免在传统的故障诊断算法中用特殊事件表示故障的单一性、有限性。此外构建的诊断系统更能准确地判断系统的运行情况, 可以及时更新系统的设备, 降低不必要的损失; 2) 故障模式空间包含了系统所有的故障信息, 可以将不同故障情形统一处理。在传统的诊断方法中, 将故障定义为发生在部件层的某些特殊事件, 导致出现了各种各样的可诊断性定义, 甚至依赖于系统的事件集, 使得可诊断性的判定很难实现, 相关的定义及算法并不能在类似的诊断问题中重新利用, 使得基于模型的诊断问题与系统描述并不能进一步分离, 而且对于多故障等情形需要单独处理。显然基于模式的故障诊断方法可以更好地应用在实际中。

而文献[4]中提出基于完备自动机的故障模式只有通过穷举的方法才能列举所有的故障情形, 在诊断系统故障时, 需要构建不同故障模式的诊断系统, 不能一次性判断出系统所发生的故障。这种用完备自动机描述故障模式的方法还存在以下不足: 首先将故障模式描述为完备的自动机, 这并没有从语义上提取系统故障信息, 也没有解决传统故障定义基于语法的局限性; 其次不同故障情形需要单独处理, 通过穷举的方式才能列出所有的故障情形, 尤其在系统规模很大、故障种类很多时, 很难得到所有的故障情形, 使得多故障诊断很难实现。表 1 是两种构建故障模式空间所需要时间的对比, 其中实验运行平台: Visual studio 2010, 编程语言: C++。CTA 中的数据是根据文献[4]提出的多故障模式空间构建算法所得, 而后面的数据是根据本文提出的计算方法所得, 其中 RE(80) 表示故障模式的平均规模是 80, 其他的类似。显然当故障个数逐渐增多时, 用文献[4]中提出的算法计算故障模式空间需要的时间远远大于本文提出的方法。此外表 1 中的数据还表明用本文提出的方法计算故障模式空间的时间复杂性与表示故障模式的正则表达式的规模有关, 所以有必要进一步根据系统的实际情况缩小故障模式的规模以降低计算复杂性。

表1 (单位:s)

故障表示	故障个数					
	5	10	20	50	80	100
CTA	0	0	15	93	640	2174
RE(80)	30	80	161	280	440	625
RE(50)	15	40	72	115	195	295
RE(20)	15	43	70	83	98	150
RE(10)	15	31	48	68	100	135

在计算候选诊断集时,首先生成与当前观测一致的系统运行空间,删除不包含候选诊断的系统行为,然后根据系统的实际运行轨迹和故障模式空间计算可能发生的故障,得到候选诊断集。在系统的每次运行中与观测一致的行为模型远远小于系统的全局模型,显然这种分开处理的方法与文献[5]相比,可以降低原算法带来的高复杂性,得到的诊断系统具有更强的实用性。此外还可以通过添加时序信息缩小观测空间,或者在匹配过程中增加启发式函数以降低计算候选诊断集的时间复杂性。

结束语 本文提出了在基于正则表达式定义的故障下系统的可诊断性定义以及故障诊断过程,其克服了传统诊断方法的一些不足。在传统的故障诊断方法中将故障定义为一类特殊事件,规定当这些特殊事件发生时,系统发生故障。这种定义故障的方式存在如下局限,首先系统的可诊断性定义和诊断系统构建依赖于具体的系统模型,不能重复利用;其次,故障描述比较单一,没有将故障发生的因果关系和时序关系表示出来;最后,传统的故障诊断方法不能同时解决单故障、多故障、间歇性故障等不同的故障情形。

本文主要在以下几方面做了改进。首先与文献[4]比较,主要有(1)故障模式用正则表达式描述,代替了完备的自动机。这种故障定义方式使其可以在语义层上描述故障,也可以表示故障与系统之间的因果关系和时序关系,避免了文献[4]中故障定义完全与系统模型分离的不足。还可以描述故障的发生过程,认为故障是由一些类转移共同导致的;(2)通过构建故障模式空间描述多故障、间歇性故障等所有不同的故障情形,而不是通过穷举考虑每一种情况。其次在文献[5]的基础上,改进了构建故障模式空间的算法和计算候选诊断的算法,提出基于故障模式空间的可诊断性定义以及诊断系统。

正则表达式具有很强的表达能力,但毕竟不是专门为故障诊断设计的,在描述系统故障特征时可能会扩大问题的范

围,从而造成匹配困难。作者进一步的研究是设计出更适合表示离散事件系统中系统故障的规则,提出高效的故障诊断方法。

参考文献

- [1] Sampath M, Sengupta R, Lafortune S, et al. Diagnosability of discrete-event systems[J]. IEEE Transactions on Automatic Control, 1995, 40(9): 1555-1575
- [2] Baroni P, Lamperti G, Zanella M. Diagnosis of large active systems[J]. Artificial Intelligence, 1999, 110(1): 135-183
- [3] Pencolè Y. Diagnosability analysis of distributed discrete event systems[A]//the 16th European Conference on Artificial Intelligence, 2004[C]. Valencia, Spain, 2004: 43-47
- [4] Jéron T, Marchand H. Supervision patterns indiscrete event systems diagnosis[A]//the 17th International Workshop on Principles of Diagnosis, 2006[C]. Peñaranda de Duero, Spain, 2006: 117-124
- [5] Lamperti G, Zanella M. Injecting semantics in-to diagnosis of discrete-event systems[A]//the 21st International Workshop on Principles of Diagnosis, 2010[C]. Portland, America, 2010: 233-240
- [6] Lamperti G, Zanella M. Context-sensitive diagnosis of discrete-event systems[A]//Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, 2011[C]. Catalonia, Spain, 2011: 969-975
- [7] Jiang S, Kumar R. Failure diagnosis of discrete event systems with linear-time temporal logic specifications[J]. IEEE Transactions on Automatic Control, 2004, 49(6): 128-133
- [8] Jiang S, Huang Z. A polynomial algorithm for testing diagnosability of discrete event system[J]. IEEE Transactions on Automatic Control, 2001, 46(8): 1318-1321
- [9] 刘磊. 编译原理及实现技术(第二版)[M]. 北京:机械工业出版社, 2010
- [10] Pencolè Y, Cordier M. A formal framework for the decentralized diagnosis of large scale discrete event systems and its application to telecommunication networks[J]. Artificial Intelligence, 2005, 164(1/2): 121-170
- [11] Briones L. Optimal observability for diagnosability [A]// the 19th International Workshop on Principles of Diagnosis, 2008 [C]. Blue Mountains, NSW, Australia, 2008: 31-38

(上接第152页)

稳定、聚类精度高。另一方面,粗糙 K-modes 聚类算法计算复杂,同时随着数据量的增大,单台服务器容易出现内存不足等性能瓶颈。为了解决该问题,本文对粗糙 K-modes 算法进行了 MapReduce 并行设计,实验表明,算法在云计算平台上的性能得到提升,能够对海量数据进行有效聚类分析。

参考文献

- [1] Han Jia-wei, Kam B M. Data mining concepts and techniques [M]. San Francisco, USA: Morgan Kaufmann, 2001
- [2] Huang Zhe-xue. Extensions to the k-means algorithm for clustering large data sets with categorical values[C]//Data Mining and Knowledge Discovery. Netherlands Kluwer Academic Publishers, 1998: 283-304
- [3] Ng K N, Lim J, Huang J Z, et al. On the impact of dissimilarity measure in k-modes clustering algorithm [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(3): 503-507
- [4] 李仁侃, 叶东毅. 粗糙 K-Modes 聚类算法[J]. 计算机应用, 2011, 31(1): 97-100
- [5] 梁吉业, 白亮, 曹付元. 基于新的距离度量的 K-Modes 聚类算法[J]. 计算机研究与发展, 2010, 47(10): 1749-1755
- [6] 赵卫中, 马慧芳, 等. 基于云计算平台 Hadoop 的并行 k-means 聚类算法设计研究 [J]. 计算机科学, 2011, 38(10): 166-176
- [7] Dean J, Ghemawat S. MapReduce: simplified data processing on large clusters[C]//Proceedings of Operating Systems Design and Implementation. San Francisco, CA, 2004: 137-150
- [8] Huang Zhe-xue. Clustering large data sets with mixed numeric and categorical values[C]//Proc of PAKDD'97. Singapore: World Scientific, 1997: 21-35
- [9] 赵恒, 杨万海. 模糊 K-Modes 聚类精度分析[J]. 计算机工程, 2003, 29(12): 27-28