②

10-16

一个系统级故障诊断问题*

A System-Level Fault Diagnosis Problem

TP306.3
TP368.1

杨晓帆 陈廷槐 何中市
Yang Xiaofan Chen Tinghuai He Zhongshi
(重庆大学计算机研究所,重庆 630044;第一作者 32 岁,男,副教授,博士)

摘 要 寻找症候的最小相容集,这是系统级故障诊断的一个重要问题,在 Chwa-Hakimi 模型下,我们证明了两个结果:1)对于二分图系统,该问题是 NP 难的;2)对于森林系统,该问题是多项式时间可解的。

关键词 故障诊断;多处理机系统;图论 二分图系统
中国图书馆资料分类法分类号 TP306.3

Abstract In system-level fault diagnosis, it is an important problem to find a minimum compatible set of the syndrome. Under the model of Chwa & Hakimi, we prove two results: 1) this problem is NP-hard for bipartite graph systems; 2) this problem is polynomial-time solvable for forest systems.

Keywords fault diagnosis; mutiprocessor system; graph theory

0 引 言

随着多处理机系统规模的扩大和系统可用性要求的提高,迫切需要快速识别系统中的故障处理机。1967 年,Preparata, Metze 和 Chien[1]建立了系统级故障诊断理论,其基本思想是:让系统中的处理机相互测试,通过对测试结果进行分析来找出故障处理机。

在系统级故障诊断中,测试是诊断的基础:1980 年,Malek[2](1981 年,Chwa 和 Hakimi[3]独立地)提出了一种比较实用的测试方法,其基本思想是:让两个相邻处理机执行同一个任务,然后检验所得结果是否相同;如果相同,这两个处理机就都认为对方无故障,否则这两个处理机就都认为对方有故障,这种测试被称为比较测试,全体测试结果被称为症候。

如果一组处理机发生故障有可能导致某个症候出现,就称这个处理机集是这个症候的相容集,寻找症候的最小相容集,这是系统级故障诊断的一个重要问题[1]。以前,主要是在有唯一最小相容集的系统中解这个问题[1~5],并找到了多项式时间算法[4]。令人遗憾的是,大多数分布式系统不满足最小相容集的唯一性假定。

近年来,开始在真实系统中寻找症候的最小相容集[6~8]。例如,Blough 和 Pelc[7]在 Malek

* 收文日期 1996-01-25
国家自然科学基金资助项目

模型下证明：对于二分图系统，可以在 $O(n^{2.5})$时间内解这个问题(n 是处理机总数)；对于森林系统，可以在 $O(n)$时间内解这个问题。

由于 Malek 模型是 Chwa-Hakimi 模型的特例，在 Chwa-Hakimi 模型下寻找最小相容集变得更为困难。我们在 Chwa-Hakimi 模型下证明：对于二分图系统，这个问题是 NP 难的；对于森林系统，可以在 $O(n^3)$时间内解这个问题。

1 术语和符号

本文假定各个处理机相互独立地等概率 p 地发生永久故障。

在多处理机系统中，处理机之间的比较测试[2,3]可以用无向图 G 表示，G 中的顶点表示处理机，G 中的边$\{u,v\}$表示在处理机 u 和 v 之间作了比较测试，G 被称为该系统的比较图。

用 $\sigma(\{u,v\})=0$ 表示处理机 u 和 v 都认为对方无故障，用 $\sigma(\{u,v\})=1$ 表示 u 和 v 都认为对方有故障，把全体比较测试结果称为症候，记为 σ.

关于比较值与处理机状态之间关系的假定被称为测试模型。本文仅研究 Chwa-Hakimi 模型(表 1)。为了对照，我们还列出了 Malek 模型(表 2)，显然，Malek 模型是 Chwa-Hakimi 模型的特例。

表 1 Mhwa-Hakimi 模型

处理机 u	处理机 v	$\sigma(\{u,v\})$
无故障	无故障	0
无故意	有故障	1
有故障	无故障	1
有故障	有故障	0 或 1

表 2 Malek 模型

处理机 u	处理机 v	$\sigma(\{u,v\})$
无故障	无故障	0
无故障	有故障	1
有故障	无故障	1
有故障	有故障	1

Chwa-Hakimi 模型告诉我们：如果 F 是全体故障处理机的集合(简称为故障集)，则

1) 如果 $\sigma(\{u,v\})=0$，那么或者 u 和 v 都$\in F$，或者 u 和 v 都$\notin F$.

2) 如果 $\sigma(\{u,v\})=1$，那么不可能 u 和 v 都$\notin F$.

反之，如果 F 满足 1)和 2)，则称 F 是 σ 的相容集。显然，必须从 σ 的相容集中去寻找故障集。

本文需要图论中的下列概念和结论[9]：

1) 二分图是一个图，其顶点集可以被划分成两部分，使得图中每条边的两个端点分别属于两个部分。

2)G 是二分图$\Leftrightarrow G$ 没有长为奇数的圈(奇圈)。

3)森林是没有圈的图。

4)图的顶点覆盖集：该图的一个顶点子集，该图的每条边至少有一个端点属于这个子集。

图论的其余知识见[9]。

现列出一个著名的 NP 难问题[10]。

VC：求连通图 G 的最小顶点覆盖集。

为了方便，对比较图 G 上的症候 σ，引入下列记号：

0-边：比较值为 0 的边　　1-边：比较值为 1 的边

E_0：全体 0-边的集合　　E_1：全体 1-边的集合

笔者主要研究下面三个问题：

- MINDIAG：在 Chwa-Hakimi 模型下，求比较图上症候的最小相容集。
- B-MINDIAG：在 Chwa-Hakimi 模型下，求二分图上症候的最小相容集。
- F-MINDIAG：在 Chwa-Hakimi 模型下，求森林上症候的最小相容集。

2 二分图系统上的诊断

证明 B-MINDIAG 是 NP 难的。

引理 1 设 F 是比较图 G 上症候 σ 的相容集，则 F 是图 G-E_0 的顶点覆盖集。

证明 设 F 是比较图 G 上症候 σ 的相容集。根据相容集的定义，G 中每条 1-边至少有一个端点属于 F，因此 F 是图 G-E_0 的顶点覆盖集。

引理 2 设 F 是比较图 G 上症候 σ 的相容集，$\{G_i; i\in I\}$ 是 $G-E_1$ 的全体连通支，则

1） 对每个 $i\in I, F\cap V(G_i)=\varnothing$ 或者 $F\cap V(G_i)=V(G_i)$.

2） 如果有一条 1-边的两个端点同时属于 $V(G_i)$，那么 $F\cap V(G_i)=V(G_i)$.

证明 根据相容集的定义，每条 0-边的两个端点都 $\in F$ 或者都 $\notin F$。因此，1)成立。根据相容集的定义，每条 1-边至少有一个端点 $\in F$。因此，1)蕴涵 2)。　证毕

定理 3 设 σ 是比较图 G 上的症候，$\{G_i; i\in I\}$ 是 $G-E_1$ 的全体连通片，$\{G_i; i\in I_1\}$ 是 $G-E_1$ 的含有 1-边的两个端点的连通片，则 F 是 σ 的相容集 $\Leftrightarrow F=(\bigcup_{i\in I_1}V(G_i))\bigcup U$，其中

1） U 是 $G-\bigcup_{i\in I_1}V(G_i)-E_0$ 的顶点覆盖集，

2）对每个 $i\in I, F\cap V(G_i)=\varnothing$ 或者 $F\cap V(G_i)=V(G_i)$.

证明 这里引理 1 和引理 2 的直接结果。　证毕

现在给出一个算法，它以一个连通图为输入，输出一个比较图和它上面的一个症候。

算法 GS

输入：图 H，其顶点集为 $\{1,2\cdots\cdots,n\}$，有 m 条边；

输出：比较图 G 和它上面的症候 σ；

①求 H 的生成树 T；（注：T 中有 $n-1$ 条边）

　对于 T 中的每条边 e，$\sigma(e)\leftarrow 0$；

②把 H 的不在 T 中的边从 e_0 到 e_{m-n} 编号；

　$for\ 1\leq i\leq n, i^0\leftarrow i; r_i\leftarrow 0$；

　$G^0\leftarrow T; k\leftarrow 0$；

③如果 G^k+e_k 没有奇圈，则转到④，否则转到⑤；

④构造图 G^{k+1} 如下：

　$V(G^{k+1})\leftarrow V(G^k); E(G_{k+1})\leftarrow E(G^k)\bigcup\{e_k\}$；

　$\sigma(e_k)\leftarrow 1$；

　$k\leftarrow k+1$；转到⑥；

⑤设 $e_k=(i,j), i<j$；

　$r_i\leftarrow r_i+1$；

构造图 G^{k+1} 如下；

$V(G^{k+1})\leftarrow V(G^k)\cup(i'_1)$；$E(G^{k+1})\leftarrow E(G^k)\cup\{\{i^0,i'_1\},\{j^0,i'_1\}\}$；

$\sigma(\{i^0,i'_1\})\leftarrow 0$；$\sigma(\{j^0,i'_1\})\leftarrow 1$；

$k\leftarrow k+1$；转到⑥；

⑥如果 $k<m-n$，则转到③，否则转到⑦；

⑦$r\leftarrow \max(r_i;1\leqslant i\leqslant n)$；

⑧构造图 G 如下：

$V(G)\leftarrow V(G^{m-n})\cup(i^j;1\leqslant i\leqslant n,r_i+1\leqslant j\leqslant r)$；

$E(G)\leftarrow E(G^{m-n})\cup(\{i^0,i_j\};1\leqslant i\leqslant n,r_i+1\leqslant j\leqslant r)$；

对于每条新加的边 e，$\sigma(e)\leftarrow 0$；

GS 算法有四条性质（引理 4～引理 7）。

引理 4　G 是二分图。

证明　由 GS 算法的定义，下面两条性质成立：

1) G^0 没有奇圈；2) 如果 G^{m-n} 没有奇圈，那么 G 也没有奇圈。

于是如果 G 有奇圈，则可以假定对某个 $k(0\leqslant k\leqslant m-n-1)$，下面两条性质成立：

(a) G_k 没有奇圈；(b) G_{k+1} 有奇圈。

设 $e_k=\{i,j\}$，根据 GS 算法的定义，可知

(A) $G_k+\{i,j\}$ 有奇圈；

(B) 在 G_k 上添加一个顶点 i^r 和两条边 $\{i,i^r\}$ 及 $\{j,i^r\}$，得到 G_{k+1}。

一方面，由 (a) 和 (A) 可知：$G_k+\{i,j\}$ 有含边 $\{i,j\}$ 的奇圈，如果 $G_k+\{i,j\}$ 有含边 $\{i,j\}$ 的偶圈，则 G_k 含有奇圈，这与 (a) 矛盾，因此，$G_k+\{i,j\}$ 没有含边 $\{i,j\}$ 的偶圈。

另一方面，由 (a)，(b) 和 (B) 可知：G_{k+1} 有奇圈有含边 $\{i,i^r\}$ 和 $\{j,i^r\}$。于是，$G_k+\{i,j\}$ 有含边 $\{i,j\}$ 的偶圈，矛盾。因此，G 没有奇圈，即 G 是二分图。　证毕

引理 5　F 是 σ 的相容集⇔下面两条性质成立：(a) F 是 $G-E_0$ 的顶点覆盖集；

(b) 有 H 的顶点覆盖集 S，使得 $F=\bigcup_{i\in S}V(G_i)$.

证明　由 GS 算法的定义可得：

1) $G-E_1$ 由 n 个连通支 $G_1,G_2,\cdots\cdots,G_n$ 组成。

2) 每个 G_i 都是星形图，其顶点集是 $\{i^j;0\leqslant i\leqslant r\}$，中心是 i^0，不含 1-边的两个端点。

根据定理 3 和 GS 算法的定义，该引理成立。　证毕

引理 6　如果 F 是 σ 的最小相容集，则 $\{i;i^0\in F\}$ 是 H 的最小顶点覆盖集。

证明　根据引理 5，可得：

1) $\{i;i^0\in F\}$ 是 H 的顶点覆盖集；　2) $|F|=(r+1)\times|\{i;i^0\in F\}|$.

现任取 H 的顶点覆盖集 S^*，根据引理 5，可得：

(a) $\bigcup_{i\in S^*}V(G_i)$ 是 σ 的相容集，　(b) $|\bigcup_{i\in S^*}V(G_i)|=(r+1)\times|S^*|$.

因为 F 是 σ 的最小相容集，所以 $|F|\leqslant|\bigcup_{i\in S^*}V(G_i)|$，即 $|\{i;i^0\in F\}|\leqslant|S^*|$。从而 $\{i;i^0\in F\}$ 是 H 的最小顶点覆盖集。　证毕

引理 7　GS 算法的最坏时间复杂度是 $O(n^2m^2)$.

证明　GS 算法的第一步可以用深度优先搜索法在 $O(n+m)$ 时间内完成，第二步也可以在 $O(n+m)$ 时间内完成，第七步可以用分治法在 $O(\log_2 n)$ 时间内完成，第八步可以在 $O(n^2)$ 时

间内完成，第三步至第六步组成的循环者次数是 $O(m)$，执行一次循环体语句（第四步或第五步）需要 $O(1)$ 时间，执行一次第六步中的判断语句需要 $O(1)$ 时间。

最后分析执行一次第三步中的判断语句所需要的时间。首先设计下列算法来判断连通图 G 中有无奇圈。

算法：ODD-CYCLE

输入：有 n 个顶点和 m 条边的边通图 G；

输出：G 有无奇圈；

1)选定 G 中的顶点 i 作为出发点；

2)用深度优先搜索法把 G 中的顶点分成两个部分 V_1 和 V_2，其中：

$V_0=\{j$：在深度优先搜索法中，从 i 经过偶数长的路径到达 $j\}$，

$V_1=\{j$：在深度优先搜索法中，从 i 经过奇数长的路径到达 $j\}$.

3)如果 G 有一条边的两个端点同时属于 V_0 或者同时属于 V_1，则 G 有奇圈；否则 G 没有奇圈。ODD-CYCLE 算法的第一步和第二步需要 $O(n+m)$ 时间，第三步需要 $O(m\ n^2)$ 时间，因此，每执行一次 GS 算法的第三步中的判断语句需要 $O(m\ n^2)$ 时间。

综上所述，GS 算法的最坏时间复杂度是 $O(m\ n^2)$.

定理 8 B-MINDIAG 是 NP-难的。

证明 根据引理 7，可以在多项式时间内把连通图 H 映射到二分比较图 G 和它上面的症候 σ。如果 B-MINDIAG 有多项式时间算法 A，那么 VC 有下列多项式时间算法。

算法：COVER

输入：图 H；

输出：H 的一个最小顶点覆盖集 S；

1) $(G,\sigma)\leftarrow GS(H)$；

2) $F\leftarrow A(G,\sigma)$；

3) $S\leftarrow\{i:i^0\in F\}$；

根据引理 5 和引理 7，COVER 算法是正确的，并且是多项式时间算法，因为 VC 是 NP 难的，所以 B-MINDIAG 也是 NP 难的。

推论 9 MINDIAG 是 NP 难的。

3 森林系统上的诊断

现证明 F-MINDIAG 是多项式时间可解的。首先设计一个算法，它以比较图上的症候为输入，输出一个图。

算法 SG

输入：有 n 个顶点的比较图 G 上的症候 σ；

输出：图 H；

1)求 $G-E_1$ 的全体连通片 $\{G_i:i\in I\}$；

2) $I_1\leftarrow\{i:i\in I$，G 有一条 1-边的两个端点都属于 $V(G_i)\}$；

3)构造图 H 的顶点集如下：

对于 $1\leqslant i\leqslant n$，$V_i\leftarrow V(G_i)$；

$V(H) \leftarrow \bigcup_{i\in I_1} V_i$;

4)构造图 H 的边集：对 $i<j$，如果 G 有一条边把 $V(G_i)$和 $V(G_j)$连接起来，则把 V_i 中每个顶点和 V_j 中所有顶点连接起来；SG 算法有四条性质(引理 10～13)。

引理 10　如果 G 是森林，那么 H 也是森林。

证明　如果 H 不是森林，不妨设 $v_0 v_1 \cdots\cdots v_k$ 是 H 的圈。

对于每条边$\{v_i, v_{i+1}\}$(令 $v_{k+1}=v_0$)，不妨设 $v_i \in V_s, v_{i-1} \in V_t$. 由 SG 算法的定义，可得：

1)G 有一条边$\{u,v\}$, $u\in V(G_s)$, $v\in V(G_t)$.

2)在 G_s 中有从 v_i 到 u 的 0-路径，在 G_t 中有从 v_{i+1}到 v 的 0-路径。

在 G 中有从 v_i 到 v_{i+1}的路径 P_i. 从而 G 有一条闭路 $P_1 \to P_2 \to \cdots \to P_k$. 即 G 不是森林，矛盾。因此，H 是森林。　证毕

引理 11　如果 F 是 H 的最小顶点覆盖集，那么对每个 V_i，或者 $F\cap V_i = V_i$ 或者 $F \cap V_i = \varnothing$.

证明　设 $T=V(H)-F$,则 T 是 H 的最大独立集。

假如有 V_i 满足 $T\cap V_i \neq V_i$ 和 $T\cap V_i \neq \varnothing$，则有一对顶点 u 和 v 满足　$u\in T\cap V_i$ 和 $v \in V_i - T$. 根据 SG 算法的定义：u 不与 T 中任何顶点相邻 ⇔ v 不与 T 中任何顶点相邻。从而 $T\cup\{v\}$ 是 G 的独立集。这与 T 是 G 的最大独立集这个假定矛盾。因此，对每个 V_i，或者 $T\cap V_i = V_i$ 或者 $T\cap V_i=\varnothing$，这蕴涵对每个 V_i，或者 $T\cap V_i = V_i$ 或者 $T\cap V_i = \varnothing$.　证毕

引理 12　如果 F 是 H 的最小顶点覆盖集，那么 $F\bigcup_{i\in I_1} V(G_i)$是 σ 的最小相容故障集。

证明　这是定理 3 和引理 11 的直接结果。

引理 13　SG 算法的最坏时间复杂度是 $O(n^2 m)$.

证明　SG 算法的第一步需要 $O(n+m)$时间，第二步需要($O(n^2 m)$时间，第三步需要 $O(n)$时间，第四步需要 $O(n^2)$时间。因此，SG 算法的最坏时间复杂度是 $O(n^2 m)$.

引理 14[7]　可以在 $O(n)$时间内找到 n-阶森林的最小顶点覆盖集。

定理 15　可以在 $O(n^3)$时间内解 F-MINDIAG.

证明　构造解 F-MINDIAG 的算法如下：

算法：FOREST

输入：有 n 个顶点的森林比较图 G 和 G 上面的症候 σ;

输出：σ 的最小相容集 F;

1)$H \leftarrow SG(G, \sigma)$;

2)求 H 的最小顶点覆盖集；

3)$F \leftarrow S\bigcup_{i\in I_1} V(G_i)$;

该算法的正确性由引理 10 和引理 12 保证。由引理 14，算法的第一步需要 $O(n^3)$时间。由于是有 n 个顶点的森林，根据引理 13，算法的第二步需要 $O(n)$时间，第三步需要 $O(n)$时间。因此，可以在 $O(n^3)$时间内解 F-MINDIAG.

4　结束语

在系统级故障诊断中，寻找症候的最小相容集是一个重要问题。在 Chwa-Hakimi 模型下，我们证明了两个结果：1)对于二分图系统，该问题是 NP 难的；2)对于森林系统，该问题

是多项式时间可解的。这些结果是对文献[7]中有关结果的深化。

由于大多数分布式系统是二分图系统而不是森林系统(例如：超立方体系统，网格系统)寻找最小相容集是十分困难的。近年来提出的概率诊断方法[8,11,12]是克服这个困难的一条重要途径。

参考文献

1 Preparata F P, Metze G, Chien R T. On the connection assignment problem of diagnosable systems. IEEE Trans. Electron. Comput. ,1967,(12):848～854

2 Malek M. Undirected graph models for system-level fault diagnosis. In Proc. 7th Symp. Comput. Architecture, 1980,31～35

3 Chwa K Y, Hakimi L. Scheme for fault tolerant computing-a comparison of modularly redundant and t-diagnosable system. Inform. Contr. ,1981,212～238

4 Dahbura A T, Masson G M. An $O(n^{2.5})$ fault identification algorithm for diagnosable systems. IEEE Trans. Comput. , 1984,(5):486～494

5 Chen Tinghuai. Fault Diagnosis and Fault Tolerance-A Systematic Approach to Speeial Topics. Berlin: Springer-Verlag,1992,65～93

6 Pelc A. Undirected graph models for system-level fault diagnosis. IEEE Trans. Comput. ,1991(11):1271～1276

7 Blough D M, Pelc A. Complexity of fault diagnosis in comparison models. IEEE Trans. Comput. , 1992,(3): 319～323

8 杨晓帆. 容错和诊断-神经网络和多处理机系统中若干问题的研究:[博士学位论文],重庆:重庆大学计机系,1994

9 Bondy J A, Murty U S R. Graph Theory with Applications. New York:Elsevier, North Holland, 1976. 5～20

10 Garey M R, Johnson I S. Computers and Intractability-A Guide to the Theory of NP-Completeness. San Francisco,CA:Freeman,1979. 17～38

11 Scheinerman E. Almost sure fault tolerance in random graphs. SIAM J. Comput. , 1987,(12)1124～1134

12 Blough D M, Sullivan G F, Masson G M. Efficient diagnosis of multiprocessor systems under probabilistic models. IEEE Trans. Comput. , 1992(8):1126～1136