

# 评估 STAR 网络可靠性的新方法

梁家荣, 花仁杰

(广西大学计算机与电子信息学院, 广西 南宁 530004)

**摘要:** 针对具有故障部件的 STAR 互连网络的可靠性问题, 在有条件的容错模型和概率模型的基础上, 提出了一种新的评估 STAR 网络可靠性的方法。该方法通过网络的二连通率来进行评估。首先提出了一个自适应的容错并行路由算法, 通过该算法的多次执行, 得到用于计算网络二连通率的相关条件参数, 在此基础上建立了用于推断网络二连通率的贝叶斯网模型。然后通过基于贝叶斯网的概率推理, 可得到反映网络可靠性的网络二连通率。研究结果表明, 该方法拓展了人们在网络可靠性评价方面的视野, 具有一定的理论和实际意义。

**关键词:** 可靠性; 二连通率; 概率推理; 自适应性; 路由算法; 贝叶斯网

**中图分类号:** TP 393

**文献标志码:** A

## New method used for evaluating reliability of STAR networks

LIANG Jia-rong, HUA Ren-jie

(Coll. of Computer and Electronic Information, Guangxi Univ., Nanning 530004, China)

**Abstract:** To solve reliability problems on STAR interconnection networks with fault parts, under the conditional fault-tolerant model and the probabilistic model, a new method used for assessing reliability of STAR networks is proposed. The reliability of networks is evaluated by two-connection probability of networks in the method. A adaptive fault-tolerant parallel routing algorithm is proposed firstly, and relational condition parameters used for computing two-connection probability of networks are derived by its performance at many times, furthermore a Bayesian networks model for deducing two-connection probability of networks is also established. The two-connection probability of networks used for reflecting reliability of networks is derived by probabilistic inference based on this Bayesian networks. The results show that this method broadens people's horizons toward the evaluation for reliability of networks, so it has certainly theoretical and practical significance.

**Keywords:** reliability; two-connection probability; probabilistic inference; auto-adapted; routing algorithm; Bayesian networks

## 0 引言

近年来, 大规模并行计算机已越来越多地应用到各个领域, 对它的依赖性也越来越强, 因此一旦它出现问题, 将会带来无可估量的损失。互连网络作为大规模并行计算机系统的重要组成部分, 随着大规模并行计算机系统的不断发展, 互连网络的规模也不断扩大, 这无疑增加了互连网络可靠性的评价难度。而互连网络可靠性的高低, 在一定程度上直接反映了大规模并行计算机系统可靠性的质量, 因此其作为研究对象被很多学者所关注。文献[1]在基于遗传算法的神经网络的基础上, 对结构系统的可靠性进行了分析, 提出了一种用于可靠性评估的新模型, 而文献[2-3]分别对概率网络和通信网络的可靠性的评价方法进行了研究。文献[4]对多阶段系统的可靠性进行了研究, 并建立了基于贝叶斯网

的分析模型。这些方法各有各的优点, 具有一定的借鉴意义, 比如文献[2-4]中的分析模型都是建立在数学语言的基础上, 具有一定的严密性, 而文献[1]中的方法则是具有一定的智能性。

在众多的互连网络模型中,  $n$ -STAR 网络具有良好的点边对称性、很强的分层结构以及良好的容错性能等优点<sup>[5-8]</sup>, 在多个领域被作为大规模并行计算机系统的网络模型, 其可靠性的计算也被广泛地开展。文献[6]在概率模型的基础上对子网概率进行了研究, 而文献[7-8]在节点和边失效模型的基础上, 对使得所有子网均失效所需要的节点或边的失效数进行了研究。文献[6-8]的一个共同特点就是都充分考虑了网络的失效情况, 具有很强的容错性。我们知道, 互连网络可靠性的高低, 在很大程度上取决于网络的通信能力, 而路由算法在网络的通信中具有至关重要

收稿日期: 2009-03-12; 修回日期: 2009-06-19。

基金项目: 国家自然科学基金(60564001); 教育部“新世纪优秀人才支持计划”专项基金(NCET-06-0756)资助课题

作者简介: 梁家荣(1966-), 男, 教授, 博士后, 主要研究方向为网络可靠性分析和数据挖掘。E-mail: liangjr@gxu.edu.cn

的作用。路由算法可进一步分为确定性路由算法和自适应路由算法。确定性路由算法是消息在源节点路由到目的节点的过程中,只采用一条固定的路由路径,而自适应路由算法则可根据网络中的节点和链路的状态信息在源节点和目的节点之间选择一条可行的无错路径。很容易看出,确定性路由算法虽然简单、有效、易实施,但却极易受到网络出错节点或链路的影响。因为能减少网络延时,增加网络吞吐量及对节点和链路出错情况的不敏感,自适应路由比确定性路由更受欢迎,也得到更为广泛的研究<sup>[9-12]</sup>,但是基于 STAR 网络的自适应路由算法就很少,尤其是在并行路由方面就更少,虽然文献[13-14]对并行算法进行了研究,但都没有考虑网络的失效情况,在容错性方面较差些。因此加强基于 STAR 网络的自适应容错并行路由算法的研究,有助于网络可靠性评价方法的提出。

在本文的研究中,我们将贝叶斯网的概率推理用于网络可靠性的评价中,在有条件的容错模型和概率模型的基础上,以 STAR 互联网络为网络模型,提出了一个新的网络可靠性的评价方法以及新的评价指标,即二连通率。该方法的具体做法是首先提出一个自适应的容错并行路由算法,通过该算法的多次运行,得到用于计算网络二连通率的相关条件参数。有了这些相关条件参数以后,可以建立用于推断网络二连通率的贝叶斯网分析模型。最后,通过基于贝叶斯网的概率推理,计算出网络的二连通率。这样,在同样规模以及同样的容错模型的基础上,通过二连通率的比较,我们可以知道网络可靠性的高低。

## 1 准备工作

一个  $n$  维 STAR 网络  $G_n$  是由  $n!$  个节点构成,每个节点可由 1 到  $n$  的  $n$  个不同数的排序表示,两个节点  $s, t$  是相邻节点且连接边记为标签  $i$  当且仅当  $s$  排序表示的第一个位置上的数值与第  $i$  个位置上的数值交换后可以得到  $t$  的排序表示,其中  $i \in [2, n]$ ,据此,网络中任意一个节点都有  $n-1$  个邻接节点。在本文中我们定义了一些跟 STAR 网络相关的符号,表示如下。

$N$  表示失效节点数; $INF(t, f)$  表示网络中节点状态的信息向量,其中  $t$  表示该节点是否被占用,  $f$  表示该节点是否失效,在这里  $t, f \in \{0, 1\}$ ;  $G'$  表示从  $G_n$  中去掉失效节点及相关边以后的新网络;  $P_i$  表示当前执行算法的一对正确节点间从源节点的每个正确邻节点出发的路径;  $d$  表示图  $G'$  中节点与节点之间的最长路径的长度值  $n! - N - 1$ 。

一个贝叶斯网是一个有向无圈图,其中节点代表随机变量,节点间的边代表变量之间的直接依赖关系,每个节点都附有一个概率分布,根节点  $X$  所附的是它的边缘分布  $P(X)$ ,而非根节点  $X$  所附的是条件概率分布  $P(X|\pi(X))$ 。

## 2 自适应的容错并行路由算法

### 2.1 算法提出

我们这个算法的前提是网络中最多有  $n-2$  个失效节点,基于这样一个前提的原因是当网络中有  $n-1$  个(或者更多个)节点失效时,网络中会出现孤立的正确节点(或者

孤立的由正确节点构成的局部结构)这种特殊的情况,这样就造成了资源的浪费。我们将提出的新算法命名为自适应容错并行路由算法(adaptive fault-tolerant parallel routing algorithm, AFTPRA)。

输入:任意的两个正确节点,  $u$  和  $v$  (不妨设  $u$  是源点,  $v$  是终点)

**算法 1** AFTPRA 算法

**步骤 1** 扫描整个网络,确定网络的容错模型以及网络中的失效节点个数  $N$ ;

**步骤 2** 根据扫描结果,建立各个节点相对应的信息向量  $INF(t, f)$ ;

**步骤 3** 将分量  $f$  的值为 1 的各个节点以及相关边从网络  $G_n$  中分离后,新网络记为  $G'$ ;

**步骤 4** 在  $G'$  中从  $u$  的每个正确邻节点  $u_i$  出发,并行做下面的事情

$P_i$  初始为包含  $u$  的单集合;

用  $w$  表示当前选定的邻节点,则  $w$  的初始值是  $u_i$ ;

FOR( $j=1; j \leq d; j++$ )

{

如果  $w=v$

把  $w$  加入到  $P_i$  中,路径  $P_i$  建立成功,跳出循环;

如果  $w$  的所有邻节点中包含了  $v$

把  $w, v$  先后加入到  $P_i$  中,并把  $w$  的  $INF$  更新为

$(1, 0)$ , 路径  $P_i$  建立成功,跳出循环;

如果  $w$  的所有邻节点中无  $t$  分量值为 0 的邻节点

{如果  $w=u_i$

跳出循环;

否则

{把  $w$  换成  $P_i$  中最后的一个节点;

将该节点从  $P_i$  中删除,并且其  $INF$  还原成

$(0, 0)$ ;

}

}

如果  $w$  的所有邻节点中有  $t$  分量值为 0 的邻节点

把  $w$  加入到  $P_i$  中,并把  $w$  的  $INF$  更新为  $(1, 0)$ ,

接着在  $w$  的所有  $t$  分量值为 0 的邻节点中选择到达  $w$  所需通信时间最短的邻节点(若有相同最短时间的邻节点,则任选一个),然后把  $w$  换成该邻节点;

}

**步骤 5** 统计  $\{P_i\}$  中基数大于 1 的集合的个数,不妨用  $m$  记之,则  $m = |\{P_i | |P_i| > 1\}|$ ;

**步骤 6** 输出从  $u$  到  $v$  的并行路径数  $m$ 。

### 2.2 AFTPRA 算法性质的分析

**定义 1** 在 AFTPRA 算法执行的过程中,从源节点出发的所有可能的正确路径在向前路由的过程中,可能会为了争夺同一节点而相互等待对方返回上一节点重新寻找新的节点,以给自己让出该节点,这样就有可能造成算法执行的中断,路由失败,我们将这种现象称为并行路由的死锁问题。

**定理 1** AFTPRA 算法是无死锁的

**证明** 算法中最关键的是步骤 4,显然只要步骤 4 不死锁,算法就不会死锁。在步骤 4 中,路由的选择标准有 3: ① 两节点间的通信时间最短; ② 要选择的邻节点没被占用; ③ 对于同一节点,若多条路径同时到达,则采用先到先占用的原则。显然第 3 条标准保证了路径之间节点的不相交性,以及步骤 4 中并行执行的非等待性。对于每条可能的正确

路径,在向前路由的过程中如果遇到了争夺节点的情况,算法先根据标准①、②进行选择,如若不成,则采取返回上一节点的办法,这样就不用进行无谓的等待,如果一直返回到初始节点,那么该路径路由失败,而同时也说明网络中的正确节点已分属于不同的路径,而根据 STAR 网络的容错特性,网络中最多有  $n-2$  个节点失效,这样必然有一条或多条可能的正确路径路由成功。综上所述,AFTPRA 算法是无死锁的。

**定理 2** AFTPRA 算法是无活锁的

**证明** 很明显,算法中对于每条可能的正确路径采用最短通信时间和选择空闲节点的路由标准,这保证了单条路径路由的可行性,同时多条路径之间采用先到先得的原则保证了多条可能的正确路径在同时路由时必然有至少一条路径路由成功,这样就不会产生永远到达不了目的节点的情况,因此算法是无活锁的。

**定理 3** AFTPRA 算法是自适应的

**证明** 在算法执行的过程为避免死锁和活锁所采用的路由标准和原则就是自适应的体现。每条路径根据网络节点是否被占用以及节点间通信时间的长短来选择下一节点;多个路径之间根据各自路由由状态以及合理的原则来选择各自的下一节点,这就保证了算法能够根据网络的最新状态来执行,以使得每对正确节点间的多条并行路径被找出,因此说算法具有自适应性。

**定理 4** AFTPRA 算法在最多  $O(d)$  的时间复杂度和最多  $O(n)$  的空间复杂度内可以找到若干条指定通信节点间的并行路径。

**证明** 算法执行可以找到若干条并行路径,这不用多说。算法在步骤 4 中需要并行执行,每个单线程需要一个处理单元,而根据 STAR 网络的特性,源节点最多有  $n-1$  个邻节点,所以步骤 4 中最多需要  $n-1$  个处理单元,算法其余各步依次执行,因为是串行执行,所以只需一个处理单元即可,这样整个算法的执行需要  $O(n)$  个空间复杂度。再考虑时间复杂度,算法在步骤 4 中并行执行,对于每个可能的路径,串行执行下去,依次需要  $1, 1, d$  个时间单元,算法其余各步依次串行执行,各需要 1 个时间单元来处理,这样整个算法的执行需要  $O(1+1+1+1+1+d+1+1)$ ,即  $O(d+7)=O(d)$  个时间复杂度。证毕

### 3 贝叶斯网的建立和参数设置

#### 3.1 贝叶斯网的建立过程

通过并行算法的多次执行,我们可以知道目前网络中的一些跟建立贝叶斯网有关的参数:

- (1) 网络中的失效节点数为  $N(N \leq n-2)$ 。
- (2) 网络中正确节点的个数为  $n! - N$ ,则正确节点对的组合数为  $\frac{(n! - N)!}{2! (n! - N - 2)!}$ ,用  $c$  记之。

(3) 网络中任一对正确节点,通过算法的执行都可以找到节点间条数确定的并行路径,不妨用  $m_i$  来表示每对正确节点间并行路径的条数,显然下标  $i \in [1, c]$ 。

- (4) 任意一对正确节点间,各条正确路径的长度。

对于每对正确节点  $(u, v)$ ,通过采用并行算法能够找到  $m$  条并行路径,这样我们以算法执行时所假定的源节点作为所建贝叶斯网的一个根节点(随机变量  $X$ ),将各条路径

上的所有节点统一作为所建贝叶斯网第二层上的一个中间节点(随机变量  $Y$ ),然后插入一个新节点作为所建贝叶斯网第三层上的一个中间节点(随机变量  $Z$ ),他们之间具有依次的依赖关系,用有向边连接表示,这样就构造了用于评估网络可靠性的贝叶斯网中的一个局部结构,如图 1 所示,通过该结构我们可以计算出某对正确节点间的连通概率。

因为整个网络中这样的正确节点对的组合数为  $c$ ,所以所建立的贝叶斯网中就有  $c$  个如图 1 所示的局部结构。在所建立的贝叶斯网中插入一个新节点(随机变量  $Pr$ ),作为贝叶斯网第四层上的一个叶节点(随机变量  $Pr$ ),分别用有向边将贝叶斯网中第三层上的  $c$  个中间节点与该叶节点连接。这样用于评估网络可靠性的贝叶斯网就建立起来了,如图 2 所示。

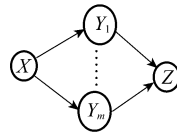


图 1 贝叶斯网的一个局部结构

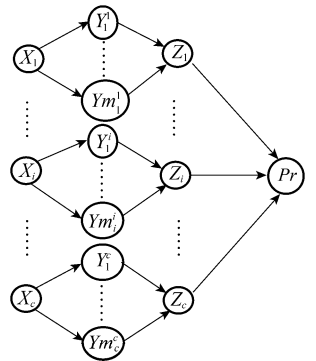


图 2 用于评估网络可靠性的贝叶斯网

#### 3.2 参数的设定

贝叶斯网的参数,即各随机变量的概率分布,一般是通过分析获得,有时也可以从问题的特性直接得到。在这里,我们根据基于概率模型的 STAR 网络中的容错问题的特性采取手动设定的办法,对如图 2 所示的贝叶斯网进行参数设定。设定如下

$$P(X_i) = \begin{cases} p \\ 1-p \end{cases} \quad (1)$$

如果  $X_i$  的状态失效,取  $p$  值;如果  $X_i$  的状态完好,取  $1-p$  值。其中,  $i \in [1, c]$ ,  $p$  为节点的失效概率。

$$P(Y_j^i | X_i) = \begin{cases} 1 - (1-p)^{l_i} \\ (1-p)^{l_i} \end{cases} \quad (2)$$

如果第  $j$  条路径上的所有节点状态完好,且与  $X_i$  的完好状态一致,取  $(1-p)^{l_i}$  值,否则取  $1 - (1-p)^{l_i}$  值。其中,  $j \in [1, m_i]$ ,  $l_i$  表示第  $i$  对节点间的第  $j$  条路径的长度。

$$P(Z_i | Y_1^i, Y_2^i, \dots, Y_{m_i}^i) = \begin{cases} 1 - \sum_{j=1}^{m_i} P(Y_j^i) \\ \sum_{j=1}^{m_i} P(Y_j^i) \end{cases} \quad (3)$$

如果第  $i$  对节点间有正确路径保持连通,取  $\sum_{j=1}^{m_i} P(Y_j^i)$  值,否则取  $1 - \sum_{j=1}^{m_i} P(Y_j^i)$  值。其中  $m_i$  表示第  $i$  对节点间并行

路径的条数。

$$P(Pr | Z_1, Z_2, \dots, Z_c) = \begin{cases} 1 - \prod_{i=1}^c P(Z_i) \\ \prod_{i=1}^c P(Z_i) \end{cases} \quad (4)$$

如果网络具有二连通性,取  $\prod_{i=1}^c P(Z_i)$  值,否则取  $1 - \prod_{i=1}^c P(Z_i)$  值。

式中,  $X_i$  代表每对正确节点对中假设的源节点,它有失效和完好两种状态;  $Y_j^i$  代表第  $i$  对正确节点间第  $j$  条正确路径连通的情况,它有连通与非连通两种状态;  $Z_i$  代表第  $i$  对正确节点保持连通的情况,它有保持连通与不连通的两种状态;  $Pr$  代表网络是否具有二连通性的情况,它具有与不具有两种状态。

对于给定的网络,网络中节点失效的概率我们可以控制给出,网络中已失效节点的情况,即容错模型我们也提前进行了一定的限制,这样再通过 AFTPRA 算法的执行,得到计算网络二连通率所需要的参数值,通过式(1)~式(4)的计算,就可以得到反映网络可靠性的二连通率。

### 4 理论计算与实例演示

本节我们结合给定的一个具体的 STAR 网络,来对本文所介绍的方法进行实践。注意,AFTPRA 算法是可以编程实现的。对于高维的 STAR 网络,通过高性能计算机的计算、程序实现,我们可以得到本方法中所需要的参数数据。限于实验室条件,我们不妨以 3-STAR 网络为例,来进行过程的演示。图 3 为一个 3-STAR 图。

假定排序为 231 的节点失效,如图 3 中的实心黑圆所示。则将节点 231 以及相关的两条边从图 3 中删除后,剩余图  $G'$  如图 4 所示。在图  $G'$  中,总共有 10 对正确节点对,即  $c=10$ ,分别是 (321, 123)、(321, 213)、(321, 312)、(321, 132)、(123, 213)、(123, 312)、(123, 132)、(213, 312)、(213, 132)、(312, 132)。对于每对正确节点对,运行 AFTPRA 算法,得到相关参数:  $\{m_i\} = \{1, 1, 1, 1, 1, 1, 1, 1, 1, 1\}$ ,  $i \in [1, 10]$ 。显然每对节点间都只有 1 条路径。这样我们用  $\{l_i\}$  来表示这 10 路径的长度:  $\{l_i\} = \{1, 2, 3, 4, 1, 2, 3, 1, 2, 1\}$ ,  $i \in [1, 10]$ 。我们假定节点失效的概率为  $p$ ,则利用式(2)可以计算出  $Y_j^i$  的值,显然  $j$  只能等于 1,也就是说每对节点间都仅有 1 条路径,而且通过式(2)可以依次计算出他们处于连通状态的概率,用集合  $\{P(Y_i^1)\}$ ,  $i \in [1, 10]$  表示他们,则通过计算得结果为  $\{P(Y_i^1)\} = \{(1-p)^2, (1-p)^3, (1-p)^4, (1-p)^5, (1-p)^2, (1-p)^3, (1-p)^4, (1-p)^2, (1-p)^3, (1-p)^2\}$ ,  $i \in [1, 10]$ 。然后再通过公式(3)计算出  $P(Z_i)$ ,显然由上面的条件,我们可以知道  $P(Z_i) = P(Y_i^1)$ ,  $i \in [1, 10]$ 。最后可求得  $P(Pr)$  如下

$$P(Pr) = \prod_{i=1}^c P(Z_i) = \prod_{i=1}^{10} P(Z_i) = (1-p)^{30}$$

因为网络节点的失效概率在现代大型集成电路中是可以被严格地控制的,所以我们通过设定,取得我们想要的网络二连通率。当然在同样的节点失效概率和规模的情况下,我们通过比较网络的  $P(Pr)$ ,可以做出网络可靠性的判断。

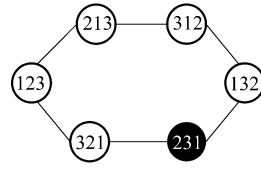


图 3 3-STAR 图

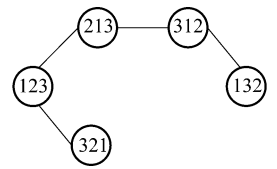


图 4 剩余图  $G'$

### 5 结束语

针对互连网络可靠性的评估,本文以 STAR 网络为网络模型,从自适应容错并行路由算法出发,建立了用于计算网络二连通率的贝叶斯网分析模型,并以此为指标,对网络的可靠性进行了评估。研究表明,本文中所提出的方法在一定程度上解决了互连网络可靠性评估方法的乏力问题,拓展了人们在网络可靠性评估方面的视野,具有一定的理论与实际意义。

### 参考文献:

- [1] Cheng Jin, Li Q S. Reliability analysis of structures using artificial neural network based genetic algorithms [J]. *Computer Methods in Applied Mechanics and Engineering*, 2008, 197 (45): 3742 - 3750.
- [2] 吴海波. 关于概率网络可靠性的一种评估方法[J]. *信息安全与通信保密*, 2008, 4(3): 43 - 45.
- [3] 钟子果, 胡爱群, 陈勇. 具有不完全可靠节点的无向网络终端可靠性评价方法[J]. *电路与系统学报*, 2005, 10(5): 136 - 143.
- [4] 刘东, 张春元, 邢维艳, 等. 基于贝叶斯网络的多阶段系统可靠性分析模型[J]. *计算机学报*, 2008, 31(10): 1814 - 1825.
- [5] Navid I, Hamid S A, Selim G A. Some topological properties of star graphs: The surface area and volume[J]. *Discrete Mathematics*, 2009, 309(3): 560 - 569.
- [6] Wu Xiaolong, Latifi Shahram. Substar reliability analysis in star networks[J]. *Information Sciences*, 2008, 178(10): 2337 - 2348.
- [7] Shahram L, Saberinia E, Wu Xiaolong. Robustness of star graph network under link failure[J]. *Information Sciences*, 2008, 178 (3): 802 - 806.
- [8] Shahram L. A study of fault tolerance in star graph[J]. *Information Processing Letters*, 2007, 102(5): 196 - 200.
- [9] Xu Jia, Li Zhi, Li Qianmu, et al. An adaptive clustering routing transition protocol in ad hoc networks[J]. *Computer Communications*, 2008, 31(10): 1952 - 1960.
- [10] Ahmad P, Hamid S A. An accurate mathematical performance model of partially adaptive routing in binary n-cube multiprocessors[J]. *Mathematical and Computer Modelling*, 2008, 48 (1): 34 - 35.
- [11] 王雷, 林亚平, 陈治平, 等. 超立方体系统中基于安全通路向量的容错路由[J]. *软件学报*, 2004, 15(5): 783 - 790.
- [12] 王高才, 陈建二, 陈松乔, 等. Mesh 网络路由算法容错性的概率分析[J]. *计算机学报*, 2004, 27(3): 319 - 327.
- [13] Youngioo C, Llyong C. A parallel routing algorithm on circulant networks employing the Hamiltonian circuit latin square [J]. *Information Sciences*, 2006, 176(21): 3132 - 3142.
- [14] Dongmin C, Okbin L, Llyong C. A parallel routing algorithm on recursive cube of rings networks employing Hamiltonian circuit Latin square[J]. *Information Sciences*, 2008, 178 (6): 1533 - 1541.