

正态分布随机Petri网的串并行结构化简

詹惠琴, 古 军, 习友宝

(电子科技大学自动化工程学院 成都 610054)

【摘要】基于正态分布随机Petri网的串行和并行两种基本结构的化简开展研究,推导出在正态分布广义随机Petri网中化简串行模型的等价理论公式,即合成执行时间为各个变迁的执行时间之和,仍服从正态分布,并且合成均值等于所有变迁执行时间的均值之和,合成标准偏差等于所有变迁执行时间标准偏差的方和根。还提出了并行模型的模拟仿真计算方法,分析了两个并行变迁独立不同分布时的情况,并给出了近似估算公式,在公式中,均值较大的变迁权重较大。

关键词 正态分布广义随机Petri网; 数值仿真; 系统模型化简; 串行并行结构

中图分类号 TP050.4

文献标识码 A

Reducing the Series-Parallel Structure of the Normal Generalized Stochastic Petri Nets

ZHAN Hui-qin, GU Jun, and XI You-bao

(School of Automation Engineering, University of Electronic Science and Technology of China Chengdu 610054)

Abstract This paper adopts a novel approach to reduction techniques of serial and parallel connection models in the normal generalized stochastic Petri net (NGSPN). The equivalent theorems of reducing the NGSPN's serial connection model is deduced. An availability method for reducing the parallel connection model is proposed and emulated. We also present an asymptotic formula for reducing the two parallel transitions connection model under independent and non-identically distributed cases. The proposed theorems and methods are especially useful in modeling and analyzing complex and large scale systems with Petri net.

Key words normal distribution general stochastic Petri net; numerical emulation; series-parallel structure; system model reduction;

Petri网的状态空间爆炸问题的研究始终是Petri网研究的主要问题之一。所谓状态空间爆炸是指在Petri网模型中, Petri网的状态总数随着模型的增长(包括网的基本元素:位置、变迁、弧、权等的增加)而呈指数性增长。学者针对不同的网模型,提出了多种Petri网模型的简化方法^[1-4]。

在Petri网模型中,变迁的串、并行是Petri网模型的主要基本形式,在测试系统、制造系统、工作流、控制系统、计算机网络等重要的Petri网模型中存在大量变迁的串、并行结构,所以对Petri网模型变迁的串、并行的化简,有助于解决Petri网状态空间的爆炸问题,特别是能够推导出它的性能等价的数学公式是很有必要的^[5-6]。

目前研究随机Petri网(stochastic Petri nets, SPN)一般都假定与随机时间关联的变迁服从负指数分

布,但是有些系统建立的随机Petri网并不符合负指数分布这个假定条件,而服从其它分布,比如测试系统和硬件电路的Petri网模型的变迁时间服从正态分布。

本文推导出在正态分布广义随机Petri网(normal generalized stochastic Petri nets, NGSPN)中化简串行模型的等价理论公式,提出了并行模型的模拟仿真计算方法,给出了两个并行变迁独立不同分布时的近似估算公式。

1 正态分布广义随机Petri网

SPN中的变迁执行时间假定为服从负指数分布,这在许多情况是适用的。但是,有些Petri网的变迁执行时间不服从负指数分布,如在电路系统中,用变迁表示电路的时延,电路时延服从正态分布^[7]。这是因为组成电路的电子元件的时延具有随机

性, 由于制造中的各种因素, 型号相同的器件的时间性能参数有一定的分散程度, 所以器件手册一般给出参数的典型值和变化范围。还有即使是同一个器件, 由于工作条件(电压、电流、阻抗等)不同, 环境条件(温度、湿度、电磁场、大气压)变化, 也会表现出随机变化的时间特性。由于影响时间特性的因素众多, 且每个影响量不大, 根据概率论中的中心极限定理, 可把一般电路的时间特性作为服从正态分布对待。

1.1 NGSPN定义

定义 1^[8] 设有一个NGSPN=(P, T, F, W, K, M_0, D), 其中:

(1) $P = \{P_1, P_2, \dots, P_n\}$ 为有限的位置集。

(2) $T = \{t_1, t_2, \dots, t_m\}$ 为有限的变迁集; 变迁 T 划分为三个子集 $T = T_f \cup T_d \cup T_s$, $T_f \cap T_d \cap T_s = \emptyset$, T_f 为执行时间是固定时间的变迁集; T_d 为执行时间服从正态分布的变迁集; T_s 为瞬时变迁集。

(3) $F \subseteq (P \times T) \cup (T \times P)$ 为流关系(即有向弧集), $P \cap T = \emptyset, P \cup T \neq \emptyset$ 。F中允许有禁止弧 h , 且 $h \in (P \times T)$ 。

(4) $W: F \rightarrow N^+$ 是弧权函数。

(5) $K: P \rightarrow N^+ \cup \{\infty\}$ 是位置容量函数。

(6) $M_0: P \rightarrow N$ 是初始标识 (marking), 满足: $\forall p \in P, \text{有 } M_0(p) \leq K(p)$ 。

(7) $D = \{\mu_i, \sigma_i \mid t_i \in T_d, t_i \sim N(\mu_i, \sigma_i^2)\}$ 是服从正态分布的所有变迁的均值和标准偏差的集合。

即在NGSPN网中, 变迁的执行时间包括服从正态分布的随机时间、固定时间和瞬时时间三种类型。

1.2 NGSPN的数字特征

SPN标识表示的系统状态与马尔可夫链等价, 所以常用马尔可夫方法分析和求解系统性能^[9]。而NGSPN标识表示的系统状态与马尔可夫链不等价, 所以不能采用通常的方法求解NGSPN模型的稳态性能。

求NGSPN网系统的性能指标可转换为求对应随机变量的数字特征。如系统的运行周期, 就是指定的变迁按照一定顺序(串行或并行)的联合执行时间, 求出联合执行时间的数字特征(均值和标准偏

差), 再用均值表示结果, 标准偏差表示结果的分散程度, 均值加减2~3倍标准偏差就是联合执行时间的范围, 即系统的运行周期范围。

2 串并行模型的化简

2.1 固定时延Petri网

定理 1 在固定时延Petri网中, 变迁 T_1, T_2, \dots, T_n 串行, 合成执行时间为 $t = \sum_{i=1}^n t_i$; 变迁 T_1, T_2, \dots, T_n 并行, 合成执行时间为 $t = \max(t_1, t_2, \dots, t_n)$ 。

根据变迁串行和并行的意义, 定理1显然成立, 其意义是串行变迁的合成执行时间为各组成变迁的执行时间之和, 并行变迁的合成执行时间取各组成变迁执行时间中的最大值。

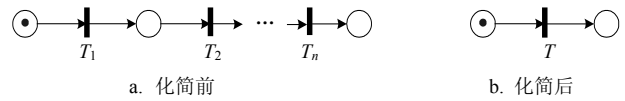


图1 Petri网变迁的串行模型和化简模型

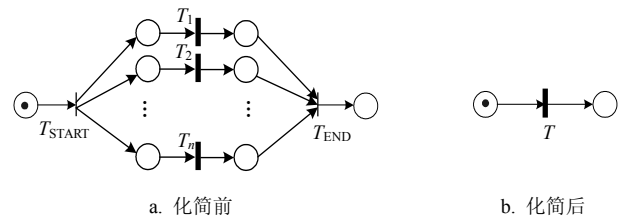


图2 Petri网变迁的并行模型和化简模型

2.2 SPN网的串并行变迁合成执行时间

文献^[10]研究了随机Petri网(SPN)中变迁执行时间服从负指数分布的变迁的串、并行化简, 得到的化简后的性能等价公式如下:

定理 2 假定在SPN中, 变迁 T_1, T_2, \dots, T_n 的执行时间相互独立, 服从参数为 λ_i 的指数分布, 即:

$$P_i(t) = 1 - e^{-\lambda_i t}, \quad i = 1, 2, \dots, n.$$

(1) 变迁 T_1, T_2, \dots, T_n 串行的性能等价执行时间为 t , 有:

$$t = \sum_{i=1}^n \frac{1}{\lambda_i}$$

(2) 变迁 T_1, T_2, \dots, T_n 并行的性能等价执行时间为 t , 有:

$$t = \sum_{i=1}^n \frac{1}{\lambda_i} - \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{1}{\lambda_i + \lambda_j} + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n \frac{1}{\lambda_i + \lambda_j + \lambda_k} + \dots + (-1)^{n-1} \frac{1}{\sum_{i=1}^n \lambda_i}$$

3 NGSPN串行变迁的合成执行时间

定理 3 设NGSPN中, $T_i \in T_d$, 若 T_i 的执行时

间为随机变量 t_i , t_i 满足正态分布, 即 $t_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, \dots, n$, 它们互相独立, $T_j \in T_f, T_j$ 的执行时间为固定时间 $t_j, j = 1, 2, \dots, m$, 所有变迁的执

行顺序为串行关系, 则全部串行变迁的合成执行时间为 $t = \sum_{i=1}^n t_i + \sum_{j=0}^m t_j$, 并且 t 仍服从正态分布 $t \sim$

$N(\sum_{i=1}^n \mu_i + \sum_{j=1}^m t_j, \sum_{i=1}^n \sigma_i^2)$, 且均值的执行时间为随机

变量 $\mu = \sum_{i=1}^n \mu_i + \sum_{j=1}^m t_j$, 标准偏差 $\sigma = \sqrt{\sum_{i=1}^n \sigma_i^2}$ 。

证明 根据随机变量特征函数的定义, 可求得正态分布 $t_i \sim N(\mu_i, \sigma_i^2)$ 的特征函数为

$\varphi_i(u) = e^{i\mu u - \frac{1}{2}\sigma_i^2 u^2}$, 令 $Y = \sum_{i=1}^n t_i$, 则:

$$\varphi_Y(u) = \prod_{i=1}^n \varphi_i(u) = \prod_{i=1}^n e^{i\mu_i u - \frac{1}{2}\sigma_i^2 u^2} = e^{i(\sum_{i=1}^n \mu_i)u - \frac{1}{2}(\sum_{i=1}^n \sigma_i^2)u^2}$$

因为 $\sum_{j=1}^m t_j = k$ 为常数, 故 $t = Y + k$ 的特征函数为:

$$\varphi_Z(u) = e^{iuk} \varphi_Y(u) = e^{iuk} e^{i(\sum_{i=1}^n \mu_i)u - \frac{1}{2}(\sum_{i=1}^n \sigma_i^2)u^2} = e^{i(\sum_{i=1}^n \mu_i + k)u - \frac{1}{2}(\sum_{i=1}^n \sigma_i^2)u^2}$$

根据唯一性定理, 故 t 服从正态分布, 且

$\mu = \sum_{i=1}^n \mu_i + \sum_{j=1}^m t_j$, $\sigma = \sqrt{\sum_{i=1}^n \sigma_i^2}$, 证毕。

定理3表明了在一个NGSPN中, 有限个串行变迁的合成执行时间为各个变迁的执行时间之和, 若这些变迁的执行时间是固定时间或服从正态分布, 那么合成执行时间仍服从正态分布, 并且均值等于各个变迁的执行时间(固定时间时)或执行时间的均值(正态分布时)之和, 方差等于所有执行时间服从正态分布的变迁的各方差的平方和。

4 NGSPN并行变迁的合成执行时间

4.1 理论方法

变迁 T_1, T_2, \dots, T_n 并联, 合成执行时间为 $t = \max(t_1, t_2, \dots, t_n)$, 以下推导变迁执行时间服从正态分布的并行合成时间。

设随机变量 $X_i \sim N(\mu_i, \sigma_i^2)$, $i=1, 2, \dots, n$, 且它们互相独立; $Z = \max(X_1, X_2, \dots, X_n)$; X_i 的概率密度函数 $f_i(x)$ 和分布函数 $F_i(x)$ 分别为:

$$f_{X_i}(x) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-\mu_i)^2}{\sigma_i^2}}$$

$$F_{X_i}(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(t-\mu_i)^2}{\sigma_i^2}} dt$$

因为 X_i 相互独立, Z 的分布函数为:

$$F_Z(z) = \prod_{i=1}^n F_{X_i}(z) = \prod_{i=1}^n \int_{-\infty}^z \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-\mu_i)^2}{\sigma_i^2}} dx$$

Z 的概率密度函数为:

$$f_Z(z) = F_Z'(z) = \left(\prod_{i=1}^n \int_{-\infty}^z \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-\mu_i)^2}{\sigma_i^2}} dx \right)'$$

上面两个公式很复杂, 无解析解。针对要解决的问题, 可转变为求 Z 的均值和标准偏差(方差), 然后使用模拟仿真的方法来求近似解。

4.2 模拟仿真方法

模拟仿真步骤如下:

(1) 用计算机软件提供的随机数发生器产生符合概率分布要求的随机数 x_{ij} , 作为各个 X_i 的 m 组样本, 其中 $i=1, 2, \dots, n$, $j=1, 2, \dots, m$ 。 m 的取值应较大, 才能够反应整个分布的情况, 可取1 000、10 000和100 000等, 本文仿真时取 $m=100 000$ 。

(2) 计算 $z_j = \max(x_{1j}, x_{2j}, \dots, x_{nj})$, 其中 $j=1, 2, \dots, m$ 作为 Z 的 m 个样本。

(3) 使用数据 $z_j, j=1, 2, \dots, m$ 画直方图, 得到 Z 的近似概率密度曲线。

(4) 计算 Z 的均值的估计值为 $\mu_z = \frac{1}{m} \sum_{j=1}^m z_j$, Z 的

标准偏差的估计值为 $\sigma_z = \sqrt{\frac{\sum_{j=1}^m (z_j - E(Z))^2}{n-1}}$

5 并行变迁合成执行时间求解举例

设两个并行变迁($n=2$)独立不同分布, 即随机变量 $X_1 \sim N(\mu_1, \sigma_1^2)$; $X_2 \sim N(\mu_2, \sigma_2^2)$; $Z = \max(X_1, X_2)$ 。在此条件下, Z 的均值和方差都无理论公式, 用以上给出的模拟仿真法计算得到一些例子的均值和标准偏差如表1所示; 随机变量 z 的概率分布直方图, 如图3示。

表1 模拟仿真不同分布时 Z 的均值、标准偏差

X_1 的分布参数	X_2 的分布参数	μ_z	σ_z
$\mu_1 = 2, \sigma_1 = 1$	$\mu_2 = 1, \sigma_2 = 1$	2.196	0.874
$\mu_1 = 5, \sigma_1 = 1$	$\mu_2 = 1, \sigma_2 = 1$	4.9973	0.995
$\mu_1 = 10, \sigma_1 = 1$	$\mu_2 = 1, \sigma_2 = 1$	9.998	1.000
$\mu_1 = 4, \sigma_1 = 4$	$\mu_2 = 1, \sigma_2 = 1$	4.564	3.238
$\mu_1 = 4, \sigma_1 = 1$	$\mu_2 = 1, \sigma_2 = 4$	4.557	1.580
$\mu_1 = 10, \sigma_1 = 5$	$\mu_2 = 2, \sigma_2 = 4$	10.333	4.629

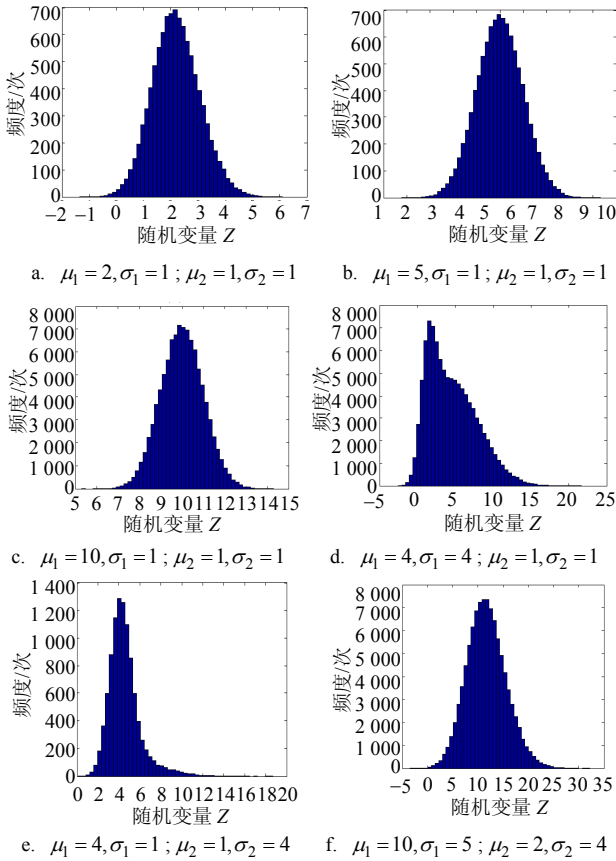


图2 概率分布直方图

然后研究近似估算公式。Z的均值为:

$$\mu_z = f(\mu_1, \mu_2, \sigma_1, \sigma_2),$$

正态分布的随机变量Z靠近均值的概率比远离均值的概率大, 有 $P\{|Z \leq 2\sigma|\} = 95.47$, 而Z每次取最大值, 所以均值大的随机变量对Z影响更多, 所以定义权 C_1 和 C_2 为:

$$\begin{cases} C_1 = \frac{\mu_1}{\mu_1 + \mu_2} \\ C_2 = \frac{\mu_2}{\mu_1 + \mu_2} \end{cases} \quad C_1 + C_2 = 1 \quad (1)$$

假设Z的均值近似估算为:

$$\mu_z = C_1\mu_1 + C_2\mu_2 + (C_1\sigma_1 + C_2\sigma_2)k_\mu \quad (2)$$

Z的标准偏差的近似估算为:

$$\sigma_z = k_\sigma (C_1\sigma_1 + C_2\sigma_2) \quad (3)$$

确定式(2)和(3)中的待定系数为:

$$k_\mu \approx 0.3 \sim 0.7$$

$$k_\sigma \approx 0.8 \sim 1.0$$

若 X_1 和 X_2 的均值相差很多, 则可忽略均值较小的那个随机变量。一般当 $\mu_i > \mu_j$ 时, 若满足 $\mu_i - 2\sigma_i > \mu_j + 2\sigma_j$, 则可近似:

$$Z = \max(X_i, X_j) \approx X_i$$

$$\mu_z \approx \mu_i$$

$$\sigma_z \approx \sigma_i$$

6 结 论

Petri网理论是一种很好的系统建模和性能分析工具^[11]。基于有些系统建立的随机Petri网不符合负指数分布这个条件, 而服从正态分布, 本文提出了正态分布广义随机Petri网(NGSPN)。针对Petri网模型的简化问题, 推导了化简串行模型的等价理论公式; 提出了并行模型的模拟仿真计算方法, 给出了两个并行变迁独立不同分布时的近似估算公式, 具有理论意义和实际价值的。

参 考 文 献

- [1] 詹柔莹. 网络传输层协议的随机Petri网模型及性能分析[J]. 电子科技大学学报, 2001, 30(1): 13-16.
- [2] DWYER M, CLARKE L. Compact petri net representation and its implications for analysis[J]. IEEE Transactions on Software Engineering, 1996, 22(11): 794-811.
- [3] WANG J, DENG Y, ZHOU M. Compositional time Petri nets and reduction rules systems, man and cybernetics[J]. IEEE Transactions on Part B, 2000, 30(4): 562-572.
- [4] ANISIMOV N A, GOLENKOV E A, KHARITONOV D I. Compositional Petri net approach to the development of concurrent and distributed systems[J]. Programming and Computer Software, 2001, 27(6): 309-319.
- [5] 曾成碧, 陈光福. 可达树分析法及其在测试中的应用[J]. 电子科技大学学报, 1999, 28(5): 524-526.
- [6] 龙士工, 罗文俊. 随机Petri网性能等价化简与分析方法研究[J]. 计算机工程, 2006, 32(12): 27-28.
- [7] MURUGAVEL A K, RANGANATHAN N. Petri net modeling of gate and interconnect delays for power estimation[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems 2003, 11(5): 921-927
- [8] 詹惠琴. 测试系统的Petri网建模和性能分析研究[D]. 成都: 电子科技大学, 2005.
- [9] 许京奕, 宋佳兴, 刘卫东, 等. 基于GSPN的工作流模型性能评价方法与应用[J]. 计算机工程与应用, 2005, 41(19): 189-192.
- [10] 田立勤, 林 闯, 周文江. 随机Petri网模型中变迁的串、并联性能等价化简技术[J]. 电子学报, 2002, 30(8): 1134-1137.
- [11] ZHAN Hui-qin. Modeling and analysis of a testing system using hybrid petri net[C]//The Eighth International Conference on Electronic Measurement and Instruments ICEMI'2007. Xi'an: [s.n.], 2007.

编辑 熊思亮