

广义随机 Petri 网下的组合 Web 服务建模与评价

门鹏, 段振华

(西安电子科技大学计算机学院, 710071, 西安)

摘要: 为了保证用语义 Web 服务标记语言(OWL-S)描述的组合 Web 服务的可靠性, 确保组合服务具有较高的系统性能, 提出了一种用广义随机 Petri 网(GSPN)对 OWL-S 过程模型进行建模、分析、验证以及评估的方法. 首先给出从 OWL-S 过程模型到 GSPN 的映射规则和转换方法, 然后针对得到的 GSPN 模型, 采用 S-不变量、T-不变量, 通过可达性等分析方法验证 OWL-S 过程模型的正确性, 最后构造与 GSPN 同构的马尔可夫链, 并根据其稳定状态概率进行系统性能评价. 经数学工具验证结果表明, 所建模型是正确、可靠的, 并发现了组合服务中的瓶颈. 所提方法可支持组合 Web 服务的重建、优化和决策.

关键词: 语义 Web 服务标记语言; 广义随机 Petri 网; 组合 Web 服务; 性能评价; 形式化验证

中图分类号: TP391 **文献标志码:** A **文章编号:** 0253-987X(2008)08-0967-05

Modeling and Evaluation of Composite Web Services Based on Generalized Stochastic Petri Net

MEN Peng, DUAN Zhenhua

(School of Computer, Xidian University, Xi'an 710071, China)

Abstract: To ensure the correctness of composite Web services described by semantic Web services markup language (OWL-S), and to guarantee the high performance of the composite Web services, a generalized stochastic Petri net based method is proposed for modeling, verification and evaluation of the OWL-S process models. The transforming rules from OWL-S process models to generalized stochastic Petri net (GSPN) are given. The analysis of S-invariant, T-invariant and reachability of the Petri net are used to verify the soundness of the process models. Moreover, Markov chains of the models are constructed. Performances of the processes are analyzed in terms of steady state probability distribution of generalized stochastic Petri net. It is verified mathematically that the models are correct and reliable. Also the bottlenecks of composite Web services is found. It is believed that the method will support to the reconstruction, optimization and decision-making of the composite Web services.

Keywords: semantic Web service markup language; generalized stochastic Petri net; composite Web service; performance evaluation; formal verification

Web 服务是面向服务体系架构(SOA)的基石, 而 SOA 实现的关键技术是构造可重用的组合 Web 服务. 随着语义 Web 服务的不断发展, 自动发现、组合、调用的 Web 服务得以实现. 语义 Web 服务标记语言(OWL-S)^[1]是目前典型的语义 Web 服务的描

述方法, 其本体 ServiceModel 可以用来描述 Web 服务的执行过程和逻辑执行顺序, 是实现自动组合 Web 服务的基础.

为了保证用 OWL-S 描述的组合 Web 服务的正确性和可靠性, 可用形式化方法来验证 OWL-S

的过程模型. Narayana^[2]等人采用 Petri 网对 Web 服务描述的本体语言 DAML-S(OWL-S 的前身)进行建模, Ankolekar^[3]等人用模型检测器 SPIN 来验证 OWL-S 过程模型的正确性, 但上述研究大都只是验证了系统的功能属性, 而不关心 OWL-S 组合服务性能等的非功能属性.

本文在保证 OWL-S 过程模型正确性的基础上, 重点讨论了性能评价问题. 根据单个 Web 服务执行的统计时间, 采用广义随机 Petri 网(GSPN)^[4]的分析方法估算了某一组合服务的平均处理时间, 从中发现流程中存在的性能瓶颈问题, 并为过程模型的评估和改进提出了一种决策方法. 同时, 对比了 Narayana^[3]描述服务组合操作语义的 Petri 网方法, 在保留需验证的行为的同时, 将 OWL-S 转换为一个更为简洁的 Petri 网模型, 以提高分析效率.

1 基本概念

1.1 OWL-S

OWL-S^[1]是一种基于语义 Web 的描述 Web 服务的本体语言. 一个 Web 服务可以用 ServiceProfile、ServiceModel 和 ServiceGrounding 3 种本体来描述其 3 方面的内容. ServiceProfile 本体可描述 Web 服务的功能、服务提供者的信息及服务的非功能属性, 包括服务分类、服务质量、完成时间和服务代价等. ServiceGrounding 本体可描述 Web 服务的使用方法. 本文用 ServiceParameter 来描述服务的平均执行时间, 这个值通常是经过测量统计得出的, 并用 ServiceModel 来描述 Web 服务的工作过程, 其子类 Process 描述服务中各个子过程的执行顺序. Process 分为原子过程、组合过程和简单过程. 其中: 原子过程描述服务一次不可分割的执行过程, 没有子过程; 组合过程通过控制结构将原子过程或组合过程组合为一个复杂过程; 简单过程是原子过程和组合过程的抽象, 它不能直接调用.

1.2 广义随机 Petri 网

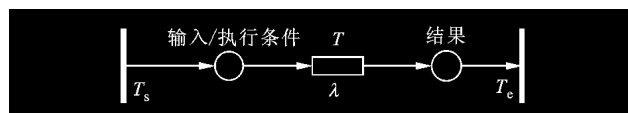
一个 GSPN 可用一个 P/T 系统 $(S, T; F, W, M_0, \lambda)$ ^[4]来描述, 其中的 T 被划分为时间变迁集 T_t 和瞬时变迁集 T_i 2 个子集. 时间变迁的实施需要一定的时间, 瞬时变迁的实施时间可以忽略不计, 即 $T = T_t \cup T_i$, $T_t \cap T_i = \emptyset$, 而在 F 中允许存在禁止弧. 在 GSPN 中, 每个时间变迁的实施时间服从指数分布, 即 $\forall t \in T_t: F_t = 1 - e^{-\lambda t}$. 其中, 实参 $\lambda_i > 0$ 是时间变迁平均实施速率, 表示在可实施情况下单位时间内平均实施的次数, 其倒数 $\tau_i = 1/\lambda_i$ 称为变

迁 t_i 的平均实施时间, λ_i 的集合为 λ . 本文假设发生的单个原子 Web 服务服从指数分布, 因为这种假设与现实中的多数系统是相符的^[4].

2 从 OWL-S 过程模型到 GSPN 的映射规则

2.1 OWL-S 原子过程转换

一个过程的定义包括 3 部分内容: 输入、执行条件和结果, 其中的结果又包含输出和执行效果. 用 GSPN 时间变迁代表原子过程的执行, λ 代表服务平均执行时间(该值是从模拟服务的实际测量中获得的或根据某种要求预测的^[5]), 变迁的输入库所代表原子过程的输入参数和执行条件, 变迁的输出库所代表结果, 那么原子过程的 GSPN 模型如图 1 所示.

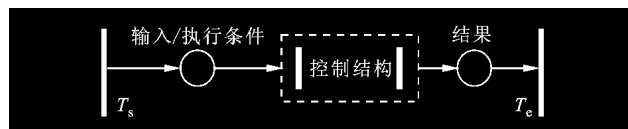


变迁 T_s 、 T_e 分别表示原子过程的开始和结束, 是瞬时变迁, 用竖线表示; 变迁 T 是时间变迁, 用矩形框表示

图 1 原子过程的 GSPN 模型

2.2 OWL-S 组合过程转换

组合过程是依 OWL-S 的控制结构将原子过程、组合过程以及控制结构按特定流程组合而成的, 如图 2 所示.



虚线框代表组合过程的内部流程, 由控制结构组成;
库所输入/执行条件代表输入参数和执行条件;
库所结果代表输出结果

图 2 组合过程的 GSPN 模型

OWL-S 过程模型的 9 种控制结构描述如下.

Sequence: 顺序执行子过程.

Split: 并发执行子过程, 但对同步要求没有进一步描述.

Split-join: 并发执行子过程, 且具有同步控制功能.

If-then-else: 2 个子过程根据条件选择执行.

Choice: 对一组子过程任选其一执行, 选择条件没有明确定义.

Any-order: 一组子过程按任意顺序执行.

Repeat-while: 循环执行子过程,在过程执行之前先判断条件是否成立.

Repeat-until: 循环执行子过程,先执行一次再判断是否继续循环执行.

Iterate: 是一个抽象概念,因此不考虑.

控制结构的 GSPN 描述如图 3 所示,其中一个时间变迁表示一个原子过程、组合过程或者一个控制结构,瞬时变迁表示流程的控制.

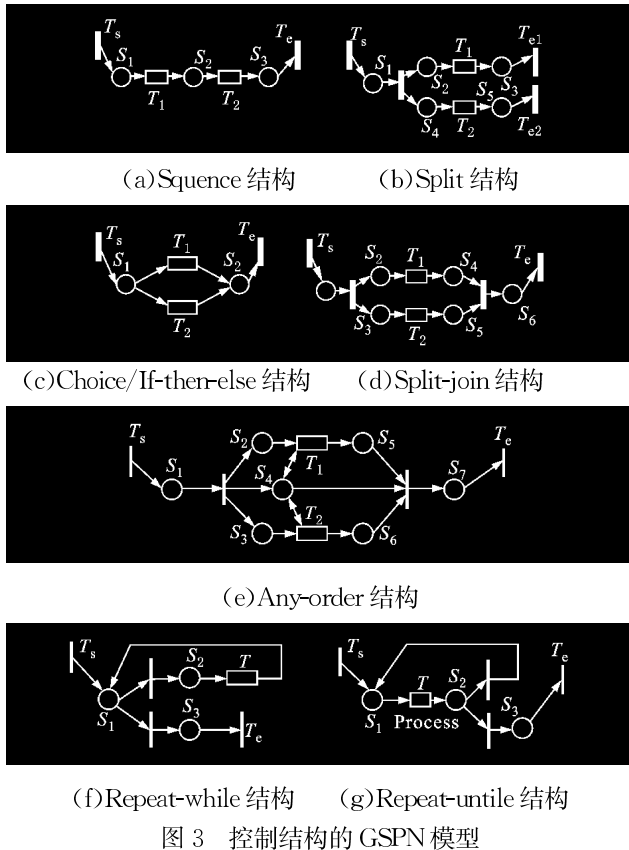


图 3 控制结构的 GSPN 模型

从一个 OWL-S 组合过程到 GSPN 模型的转换,可根据 Process 的描述采取从外到内递归替换的方式进行,当组合服务的控制结构中的时间变迁代表组合过程或者控制结构时,可用它们对应的 GSPN 模型进行替换,直到将时间变迁转换成原子过程为止.为了进行性能评价,可用一个库所将 GSPN 模型的首尾变迁连接起来.

3 OWL-S 过程模型的分析与评价

对 OWL-S 的 GSPN 模型进行正确性分析能够发现组合过程中存在的诸如死锁、非正常终止等问题,而对模型进行性能评价能够估算组合流程的周期时间、资源利用率等性能指标.

分析模型正确性的方法主要有可达图分析和不

变量分析.基于可达图的分析方法,采用枚举的思想来构造所有可达状态的可达集,从而判断系统的性质.基于不变量的分析方法,将 Petri 网结构用矩阵表示,通过引入线性代数法对性质进行分析,并用 S-不变量判断系统的有界性、安全性、守恒性,用 T-不变量判断系统的活性^[6].

GSPN 模型性能评估步骤^[4]如下.

步骤 1: 构造与 GSPN 同构的马尔可夫链, GSPN 的可达图与时间连续的奇次马尔可夫链是同构的.依据 GSPN 中瞬时变迁的触发优于时间变迁的原则,得到相应的可达图,并在每个弧上标出变迁的平均实施速率,由此得到马尔可夫链.

步骤 2: 求约简后的马尔可夫链的转移概率矩阵.

步骤 3: 求与步数相关的实存状态稳定概率分布.

步骤 4: 求稳定状态概率.

步骤 5: 性能指标分析.

4 实例分析

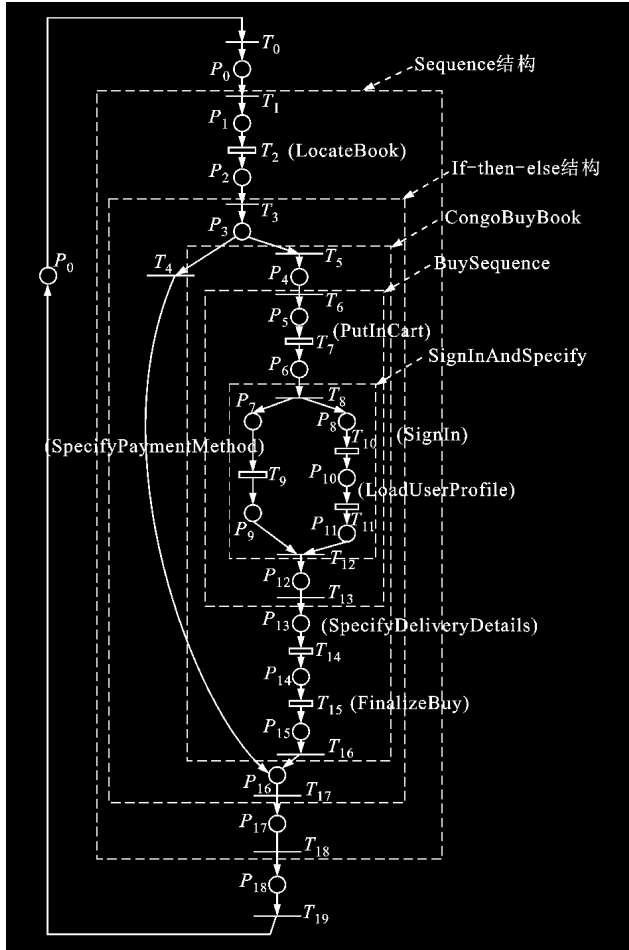
4.1 概述

本文对 OWL-S 规范给出的示例 CongoProcess.owl^[1]的组合过程 FullCongoBuy 进行分析,并对其子过程 BuySequence 作略微修改,同时给出了每个原子过程的平均实施速率. FullCongoBuy 由原子过程 LocateBook 和一个 If-then-else 控制结构顺序组成,当 LocateBook 查询库存为空时, FullCongoBuy 过程结束,否则执行复合过程 CongoBuy-Book 继续完成购买操作. FullCongoBuy 由复合过程 BuySequence 及原子过程 SpecifyDeliveryDetails、FinalizeBuy 顺序组成.复合过程 BuySequence 先使用原子过程 PutInCart 将需要购买的书籍放入购物车,然后用组合过程 SignInAndSpecify 完成登陆,并同时告知支付方式操作. BuySequence 由组合过程 SignInSequence 和原子操作 SpecifyPaymentMethod 使用控制结构 Split-join 并发完成, SignInSequence 先用原子过程 SignIn 完成登陆,再用 LoadUserProfile 加载用户信息.

4.2 建立 GSPN 模型

FullCongoBuy 的 GSPN 模型如图 4 所示.用变迁 T_0 、 T_{19} 作为 FullCongoBuy 开始、结束的标记,库所 P_0 、 P_{18} 代表输入/执行条件和结果.图 4 中,最外层虚线框代表 FullCongoBuy,是由原子过程

LocateBook和控制结构 If-then-else 顺序组成的,并用变迁 T_2 表示 LocateBook. 对分支结构内部的组合过程 CongoBuyBook 连续进行递归替换,直到所包含的所有组合过程均被分解为原子过程,同时用一个时间变迁表示. 最后,用库所 P_{19} 连接变迁 T_{19} (T_0 表示该过程可以反复执行). 此时,实施速率 $\lambda = (\lambda_2, \lambda_7, \lambda_9, \lambda_{10}, \lambda_{11}, \lambda_{14}, \lambda_{15}) = (1/40, 1/25, 1/70, 1/30, 1/20, 1/15, 1/35)$.



每个虚框代表相应组合过程的控制结构; T_2 、 T_7 、 T_9 、 T_{10} 、 T_{11} 、 T_{14} 和 T_{15} 分别代表原子操作 LocateBook、PutInCart、SpecifyPaymentMethod、SignIn、LoadUserProfile、SpecifyDeliveryDetails 及 FinalizeBuy

图4 组合过程 FullCongoBuy 的 GSPN 模型

4.3 验证组合过程的正确性

设 Petri 网的关联矩阵为 C , S -不变量为 α , T -不变量为 β , 则 $C \cdot \alpha = 0$, $C^T \cdot \beta = 0$. 由此可求得 Petri 网(图4)的 S -不变量的支撑集合, 从中看出每个库所都属于某一 S -不变量, 因此 Petri 网是有界的. 另外, 每个变迁都属于某一 T -不变量, 每个变迁均可触发, 因此该模型也是活的. 通过可达性分析可以证明系统的安全性, 这里不再赘述.

4.4 组合过程性能评价

(1)构造与 GSPN 同构的马尔可夫链, 如图5所示.

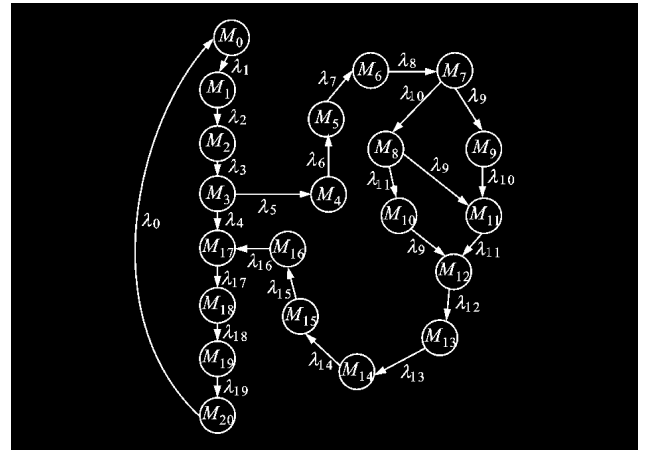


图5 与 GSPN 同构的马尔可夫链

GSPN 的可达集可以分为 2 个不相交的子集: 仅使时间变迁可实施的状态集和可使瞬时变迁可实施的状态集, 前者称为实存状态, 后者称为消失状态. 马尔可夫链的状态集为: $M_0 = \{P_0\}$, $M_1 = \{P_1\}$, $M_2 = \{P_2\}$, $M_3 = \{P_3\}$, $M_4 = \{P_4\}$, $M_5 = \{P_5\}$, $M_6 = \{P_6\}$, $M_7 = \{P_7, P_8\}$, $M_8 = \{P_7, P_{10}\}$, $M_9 = \{P_8, P_9\}$, $M_{10} = \{P_7, P_{11}\}$, $M_{11} = \{P_9, P_{10}\}$, $M_{12} = \{P_9, P_{11}\}$, $M_{13} = \{P_{12}\}$, $M_{14} = \{P_{13}\}$, $M_{15} = \{P_{14}\}$, $M_{16} = \{P_{15}\}$, $M_{17} = \{P_{16}\}$, $M_{18} = \{P_{17}\}$, $M_{19} = \{P_{18}\}$, $M_{20} = \{P_{19}\}$, 其中实存状态集为 $\{M_1, M_5, M_7, M_8, M_9, M_{10}, M_{11}, M_{14}, M_{15}\}$, 其余为消失状态.

(2)求约简后的马尔可夫链的转移概率矩阵. 约简后的马尔可夫链的转移概率矩阵为 $U = F + EG^\infty$, 其中 F 表示 U 中实存状态向实存状态的转移概率, E 表示 U 中实存状态向消失状态的转移概率, G^∞ 中的元素 $g_{ij} = P(r \rightarrow j)$ 表示从给定的消失状态 r 出发, 经过任意步首先到达实存状态 j 的概率. 所以, 有

$$U = \begin{bmatrix} & M_1 & M_5 & M_7 & M_8 & M_9 & M_{10} & M_{11} & M_{14} & M_{15} \\ M_1 & 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ M_5 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ M_7 & 0 & 0 & 0 & 7/10 & 3/10 & 0 & 0 & 0 & 0 \\ M_8 & 0 & 0 & 0 & 0 & 0 & 7/9 & 2/9 & 0 & 0 \\ M_9 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ M_{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ M_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ M_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ M_{15} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(3)求与步数相关的实存状态稳定概率分布. 用行向量 $\mathbf{Y}$ 代表稳定状态概率, 根据 Marsan 求解稳定概率分布的线性方程组, 可求得与步数相关的稳定状态概率, 从而得 $y_1=0.25$, $y_5=0.125$, $y_7=0.125$, $y_8=0.086$, $y_9=0.038$, $y_{10}=0.068$, $y_{11}=0.058$, $y_{14}=0.125$, $y_{15}=0.125$.

(4)求稳定状态概率. 选择约简后的马尔可夫链的一个状态 i 作为参考状态, 则在连续访问状态 i 期间访问状态 j 的次数为 $V_{ij}=\mathbf{Y}_j/\mathbf{Y}_i$. 每个状态的驻留时间为

$$T_{Si} = \begin{cases} 0 & \forall i \in V \\ \left[\sum_{j \in H_i} r_f \right]^{-1} & \forall i \in T \end{cases}$$

式中: H_i 是实存状态 i 下可实施的变迁集合. 返回参考状态 i 的平均时间为 $W_i = \sum_{j \in T} V_{ij} T_{Sj}$. 由此, GSPN 的稳定状态概率最终可表示为

$$P[j] = \begin{cases} 0 & j \in V \\ V_{ij} T_{Sj} / W & j \in T \end{cases}$$

选择 M_1 为参考状态, 分别求得 V_{ij} 、 T_{Sj} 和 W_1 , 从而得到稳定状态概率 $P[M_1]=0.328\ 9$, $P[M_5]=0.102\ 8$, $P[M_7]=0.086$, $P[M_8]=0.044$, $P[M_9]=0.037\ 5$, $P[M_{10}]=0.156\ 6$, $P[M_{11}]=0.038$, $P[M_{14}]=0.061\ 7$, $P[M_{15}]=0.143\ 9$. 时间变迁的利用率 $X(T_i) = \sum_{M \in E} P[M]$, 其中 E 是使 T_i 可触发的所有可达标示集合. 所有时间变迁利用率则为 $X(T_2)=0.328\ 9$, $X(T_7)=0.102\ 8$, $X(T_9)=0.286\ 6$, $X(T_{10})=0.123\ 5$, $X(T_{11})=0.082$, $X(T_{13})=0.061\ 7$, $X(T_{14})=0.143\ 9$.

(5)性能指标分析. 在本例中, 变迁 T_2 每实施一次就代表整个组合 Web 服务的调用完成一次, 则变迁的平均激发速率 $f_i=\lambda_i X(T_i)$, 其中 λ_i 是 T_i 的平均实施速率, 则 T_2 的平均激发率 $f_2=0.008\ 22$, 由此得到 FullCongoBuy 的平均实施时间为 121.618 ms. 另外, 根据变迁利用率找到的服务组合性能瓶颈是原子过程 LocateBook 和 SpecifyPaymentMethod, 要提高整个组合服务的效率, 需完善这 2 个原子过程.

5 结束语

本文提出了一种用广义随机 Petri 网对 OWL-S 过程建模、分析、验证及评估的新方法. 它不仅可以在组合服务执行前发现潜在的逻辑错误, 而且可以估算整个组合服务的总实施时间, 找出组合流程中存在的性能瓶颈.

下一步的工作将是开发一种自动完成从 OWL-S 到 GSPN 的转换、分析、验证以及评估等任务的原型系统.

参考文献:

- [1] MARTIN D, BURSTEIN M, NARAYANA S, et al. OWL-S1.1 release [EB/OL]. [2007-11-01]. <http://www.daml.org/services/owl-s/1.1/>.
- [2] NARAYANA S, MCLLRAITH S A. Simulation, verification and automated composition of web services [C] // Proceedings of the 11th International World Wide Web Conference. New York, USA: ACM, 2002: 77-88.
- [3] ANKOLEKAR A, PAOLUCCI M, SYCARA K. Towards a formal verification of OWL-S process models [C] // Proceedings of the 4th International Semantic Web Conference. Heidelberg, Germany: Springer-Verlag, 2005: 37-51.
- [4] 林闯. 随机 Petri 网和系统性能评价 [M]. 北京: 清华大学出版社, 2005.
- [5] 谢文祥, 薛钧义. 状态转移服从指数分布控制系统的 SPN 可靠性分析 [J]. 西安交通大学学报, 1999, 33(2): 11-18.
XIE Wenxiang, XUE Junyi. Stochastic Petri nets applied to reliability of control system with exponential distribution [J]. Journal of Xi'an Jiaotong University, 1999, 33(2): 11-18.
- [6] ZURAWSKI R, ZHOU Mengchu. Petri nets and industrial application: a tutorial [J]. IEEE Transactions on Industrial Electronics, 1994, 41(2): 567-583.

(编辑 苗凌)