网络体系结构建模和性能评价的形式化方法——随机进程代数

冯烟利[1,2], 余镇危[1], 潘 耘[1], 刘克俭[1]

(1. 中国矿业大学机电与信息工程学院(北京校区), 北京 100083)
(2. 山东工商学院信息工程学院, 山东 烟台 264005)

摘要: 随机进程代数是从经典进程代数发展起来, 用于并发系统的定性和定量分析的形式化方法, 特别适合在计算机网络和通信协议的建模和性能评价. 首先介绍了 SPA 的语法和操作语义, 并以令牌环局域网模型为例来说明其应用. 进一步讨论了 SPA 的三种等价关系, 以期解决模型状态空间爆炸问题

关键词: 形式化方法; 随机进程代数; 性能评价

1 引 言

计算机网络体系结构是指导计算机通信协议设计和实现的逻辑框架, 它包含了模型化的通信功能, 网络资源的分配及各模块或构件的交互等. 网络系统是具有分布和并发特性的复杂系统, 需要用形式化方法进行精确的描述和验证. 所谓形式化方法就是具有坚实的数学理论基础并由此给出一个形式化规范语言来描述系统行为的方法[1]. 形式化方法为计算机网络的性能评价提供了一种有效的、抽象的、低廉的模型工具. 计算机网络的性能可分为定性指标和定量指标. 定性指标包括可靠性、有无死锁和活锁、活性及公平性等等. 定量指标包括系统的吞吐量、用户的响应时间和资源的利用率等. 相应地可以把形式化方法分为两类: 一类只描述并发系统的功能特性, 对系统进行定性分析. 有代表性的方法是有限状态机[2]、Petri 网[3]和进程代数(PA)[4,5]等. 一类是增加了定量分析的参数(如时间和概率), 例如排队网络[6]、随机 Petri 网[7]和随机进程代数[8](SPA)等. 由于系统的性能建模与系统设计紧密相连, 只考虑系统的定性指标已满足不了需求, 具有定量性能分析的形式化方法成为更加有力的工具

2 从 PA 到 SPA

进程代数是基于代数形式建立并发系统的模型并提供对模型结构和行为进行推理的工具, 包括 CCS[4]和 CSP[5]. 在进程代数中系统可模型化为实体的集合, 这些实体称为进程(process), 进程之间可以进行交互和通信. 进程执行一些原子活动(actions), 每个活动只联系一个实施类型, 没有时间的概念, 适用于系统的定性分析. 在 TCCS(Temporal CCS)[10]中, 活动增加了取值为自然数的时间域, 一个进程不仅可观察到它执行的活动类型, 亦可观察到执行该活动的时间延迟. PCCS(probabilistic CCS)[11]扩展了进程代数, 将不确定的选择用概率选择来代替, 从而量化了这种不确定性

收稿日期: 2003-11-21
基金项目: 本课题得到教育部博士点基金(20030290003)和山东省教育厅科研发展计划(03P09)资助

SPA 通过将每一个活动联系一个表达执行时间的随机变量扩充了经典的进程代数, 这些随机变量满足指数分布, 由此建立了与 CTMC (Continuous Time Markov Chain) 之间的联系, SPA 也就具备了坚定的数学理论基础. 目前有代表性的 SPA 包括 PEPA[12], TIPP[13] 和 MPA[14], 相应的模型工具如: TIPPTOOL, PEPAWORKBENCH 和 TWOTOWERS 等.

3　SPA 的语法和语义

为了形式化描述系统的进程的行为以及相互之间的交互, SPA 定义了基本元素和一套完整的语法和语义, 包括活动和操作符. 活动描述了系统的相关行为, 由序偶 (α, λ) 表示, 其中 α 为活动名字, λ 为活动的执行速率, 满足指数分布. 活动亦可是被动的, 即其执行速率由与其同步的共享活动的速率来确定, 被动活动记为 $(\alpha, -)$. 活动还可以是内部不可见的活动, 可用 τ 表示. 活动的全体可表示为集合: $ACT := Com \cup \{\tau\}$. SPA 语法[15]描述如下:

$$P ::= (\alpha, \gamma).P \mid P + P \mid P \| P \mid P/L \mid A$$

· 前缀 $(\alpha, \gamma).P$: 前缀用来描述系统组件的顺序行为. 进程 $(\alpha, \gamma).P$ 先执行活动 (α, γ), 然后表现为进程 P 的行为.

· 选择, $P + Q$: 进程 $P + Q$ 表示其行为为 P 或者为 Q, P 和 Q 的活动都可实施, 最先完成的行为决定进程后续的行为. 选择操作用来描述系统中的不确定性, 采用的机制就是竞争策略. 概率分布的连续性能保证 P 与 Q 同时完成活动的概率为零.

· 并发, $P \|_S Q$: 并发操作根据活动是否属于 S 分为两种情况, 这里 $S \subseteq COM$, 称为同步集, 若活动名字不属于 S, 则 P 和 Q 独立进行或者并发进行. 若活动名字属于集合 S, 则 P 和 Q 必须同步或协同参与该活动. 只有当该活动在 P 和 Q 都可实施时, $P \|_S Q$ 才可实施. 该活动的实施速率可定义为两个进程中相对较慢的速率. 特别地当 $S = \Phi$ 时可简记为 $P \| Q$.

· 隐藏, $P/L: L \subseteq Act$: 隐藏操作提供了一种抽象机制, 使得进程行为的某些部分对外部环境不可见. 进程 P/L 除了在 L 上的活动隐藏外, 表现为进程 P 一样的行为. 隐藏的活动可由名字 τ 来表示, 但该活动的速率不受影响, 可以看成组件的内部延迟. 隐藏操作可以使描述系统行为时尽量详细, 而在分析系统时可以抽象掉不必要的细节. 在性能测量时不必考虑隐藏为内部活动的影响. 隐藏操作也可使同步操作不受其他进程的干涉. 隐藏机制是进程代数方法的主要优点之一.

· 常量, $A := P$: 常量 A 是一个进程, 由方程 $A := P$ 来定义的, 表明 A 具有进程 P 一样的行为. 可利用这种定义为进程赋上名字. 同时可用它来定义系统的递归行为, 而不必单独定义递归操作符, 此时, A 重复出现在方程右边.

4　模型分析

本节以令牌环局域网为例, 说明如何利用 SPA 对该局域网行为进行抽象描述, 建立模型及进行模型分析的方法. 从中进一步增强对 SPA 的语法和语义的理解. 在此过程中亦可发现 SPA 作为实际系统模型工具的特点.

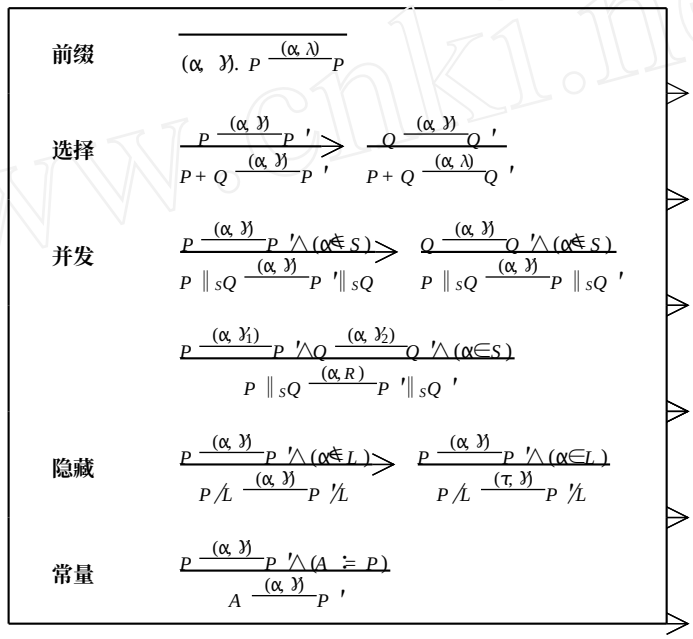
用户的行为可描述为或者有消息发送, 或者空闲. 有消息发送的行为可描述为有消息到达, 等待令牌和发送消息. Arrive 表示用户消息的到达; start 表示可以发送消息, 即该用户已获得令牌; send 表示发送信息; Empty 表示用户空闲, 没有消息发送; Move 表示令牌继

续传递到达下一个用户.

User ≒ (Arrive, λ). (Start, -). (Send, μ). User+ (Empty, γ). User

其中 Start 活动是被动活动, 其速率由同步的其它组件确定

表 1 SPA 的操作语义



令牌的行为可描述为到达某一个用户时, 用户接受询问. 若有消息在等待, 则传送该消息. 否则令牌移动到下一个用户, 在整个局域网中循环往复.

Token ≒ (Start, η). (Send, -). (Move, δ). Token+ (Empty, -). (Move, δ). Token 简单起见, 现在假设用户每次只处理一条消息, 一个用户与令牌组成的系统可以描述为:

Sys1 ≒ User ‖ s Token 其中 S= {Start, Send, Empty}

根据 SPA 的语义规则及前面所述的DG 图的定义, 可得到图 1.

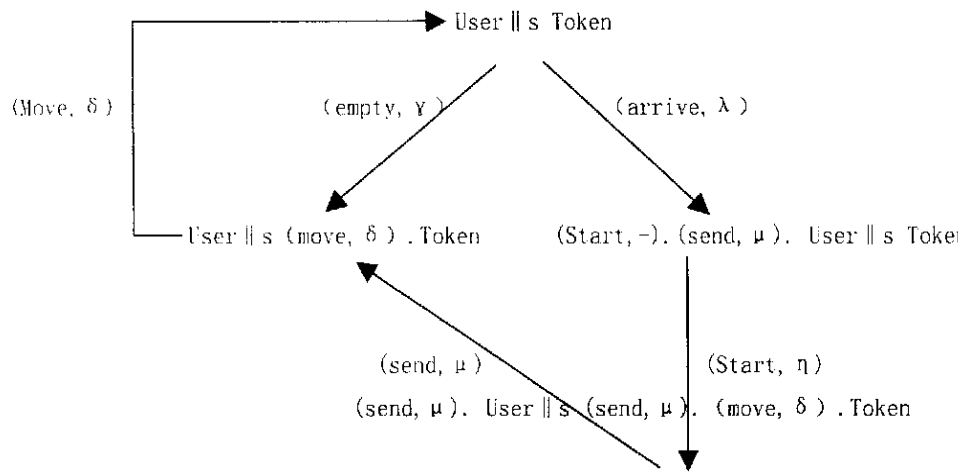


图 1 Sys1 的DG 图

从DG 图中可以得到模型的一些定性的特性如没有死锁和活锁　更进一步, 可从DG 图中直接得到CTMC　图中每个节点对应马尔可夫链的一个状态　图中弧上标注的速率与状态之间的变迁相关, 去掉活动类型, 得到该模型的状态转移速率图如图 2 所示:

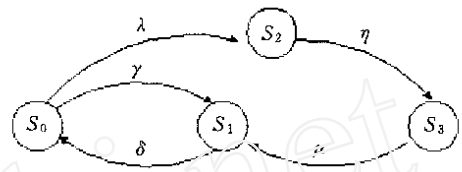


图 2　状态转移速率图

由图 2 可直接得到生成器矩阵为:

$$Q=\begin{bmatrix} -(\gamma+\lambda) & \gamma & \lambda & 0 \\ \delta & -\delta & 0 & 0 \\ 0 & 0 & -\eta & \eta \\ 0 & \mu & 0 & -\mu \end{bmatrix}$$

求解矩阵方程:

$$\Pi Q = 0$$
$$\sum \Pi = 1$$

得到稳定状态概率向量 Π　矩阵Q 和向量 Π 是求解模型各种性能评价指标的基础　通过CTMC 的报酬(reward)结构, 就可以定义许多性能测量如吞吐量和利用率等.

现假设系统共有N 个用户, 令牌在N 个用户中依次循还, 则系统可描述如下:

$User_i \triangleq (Arrive, \lambda).(Start_i, -).(Send, \mu).User_i + (Empty, \gamma).User_i$

$Token_1 \triangleq (Start_1, \eta).(send, -).(Move, \delta).Token_2 + (Empty, -).(Move, \delta).Token_2$

$Token_i \triangleq (Start_i, \eta).(send, -).(Move, \delta).Token_i + 1 + (Empty, -).(Move, \delta).Token_i + 1$

$Token_n \triangleq (Start_n, \eta).(send, -).(Move, \delta).Token_1 + (Empty, -).(Move, \delta).Token_1$

$$i = 1, 2 \cdots N$$

假设系统开始时令牌在第一个用户中, 则:

$$SYS \triangleq (User_1 \| User_2 \| \cdots \| User_n) \| sToken_1 \quad S = \{Start_i, Send, Empty\}$$

5　等价与合并

形式化模型性能评价的一个重要问题时状态空间的爆炸问题　在 SPA 的DG 图中的状态结点, 会随着模型的规模和复杂性的增加而成指数增长的趋势, 这使得实际系统的性能评价不可能　SPA 非常引人注目的优点是它的合并特性　复杂系统可先从独立的小的组件开始建模 , 每个组件尽量详细地描述分析, 然后再考虑它们之间的交互, 逐步组合成完整的系统　等价关系可认为是进程代数语义的一部分, 它可用来区分系统实体与实体之间的三种等价类: 系统与模型之间的等价, 模型与模型之间的等价和状态与状态之间的等价. 在进程代数中, 所有实体—系统、模型和状态都可表示为进程　如果定义的等价关系满足一致性(congruence), 就可用简单的模型或组件(如乘积形式解)代替复杂的模型或组件, 由此简化系统模型　如果定义的等价关系满足汇集性(lumpability), 一个模型中的具有相同行为的状态可以不加区分合并在一起, 以减少状态空间　SPA 首先在模型化简(model simplification)和状态聚集(state aggregation)两个层次上提供了性能建模的合并方法　下面将讨论 SPA 中定义的三种等价关系或称为马尔可夫相似(Markavian bisimulation). 通过等价关系相应的代数公理, 就可以在语义一级上应用这些公理对系统模型进行化简, 而不必在低级

的DG 图上进行.

5.1 强相似(strong bisimulation)

进程代数的等价关系的定义只考虑组件功能的相同. 在 SPA 中就需要考虑其功能和时间两方面的因素, 既在功能和时间上都一致. 一个活动 $(\alpha, 2\gamma)$ 与两个同构可实施的活动 (α, γ) 可以不加区分, 即是强相似的.

定义 5.1.1 等价关系 $R \subseteq S \times S$ 是强相似的且 $(P, Q) \in R$ 当且仅当对所有的活动 α 和所有的等价类 $C \in S/R$ 有:

$$\sum_{\gamma_P \in I(P, \alpha, C)} \gamma_P = \sum_{\gamma_Q \in I(Q, \alpha, C)} \gamma_Q$$

这里 $I(X, \alpha, C) := \{ |\gamma| X \xrightarrow{(\alpha, \gamma)} C' \wedge C' \in C \}$, 称进程 P 和 Q 是强相似的, 记为 $P = Q$.

上述强相似的定义满足一致性和汇集性. 由此定义可在语义一级得到许多等价公式, 这些公式为语义重写算法奠定了基础, 使得在语义一级上的组件简化成为可能, 等价公式见[15]. 利用这种方法可大大减少状态的数量. 假设 N 为用户的个数, 则模型状态空间的复杂性由 $O(e^N)$ 减少为 $O(N^2)$.

5.2 弱相似

前面已提到过 SPA 利用隐藏操作符使某些活动外部不可见, 这意味着模型可执行一系列内部活动 τ. 在进行不同的性能测量计算时, 将不含有性能信息的一些活动隐藏起来, 对应于这些活动不赋予报酬, 当满足一定的语义条件时, 弱相似关系可使一系列的隐藏活动 τ 由单一的隐藏活动代替, 而速率相加, 由此可以合并一些中间状态, 利用 SPA 这种抽象机制, 可使模型更加简化.

5.3 马尔可夫观察一致性(Markovian observational congruence)

TIPP 中增加了瞬时活动用来描述那些时间可以忽略的行为. 这可以理解为活动速率为无穷大. 消去的过程是基于马尔科夫观察一致性的关系, 在建立马尔科夫链的同时消去瞬时活动, 消去的瞬时活动为内部活动, 不参与外部的同步.

6 结 论

随机进程代数如 TIPP 和 PEPA 为系统建模和性能分析提供了一个强有力的形式化工具, 它扩展了进程代数只能进行功能上的分析, 通过给模型中的每个活动联系一个指数分布的执行时间随机变量, 可以将语义模型转换为连续时间的马尔科夫链, 由此可利用已有的求解算法来进行许多性能的测量, 执行时间的指数分布的假设大大简化了系统的分析. 对于实际系统而言, 虽然存在不能由基于传统的马尔科夫模型进行评价的情形如自相似模型, 但大多数情况下的活动执行时间是接近指数分布的. 随着随机进程代数理论的发展及相应的实现工具的进一步完善, 加上其自身的优点如合并、抽象和等价的思想, 将成为模型分析设计和自动建模技术的一种有效的形式化方法.

参考文献:

[1] Bichard Brandau, Tony Confrey, Alin D'Silva, Chirstoper Matheus, and Robert weihmayer.

[2] Arthur Gill Introduction to the theory of Finite-State Machines[M]. New York, Mcgraw-Hill, 1962.

[3] Petri C. Communication with Automata[M]. Technical Report RADC-TR-65-377, Rome Air Dev. Center, New

York, NY, 1966.
[4] Milner R. Communication and Concurrency Prentice Hall[M]. London, 1989.
[5] Hoare C A R. Communicating Sequential Processes Prentice-Hall[M]. Englewood Cliffs, NJ, 1985.
[6] Hock N C. Queuing Modeling Fundamentals[M]. John Wiley & Sons Ltd, 1997.
[7] 林闯　随机 Petri 网和系统性能评价[M]. 清华大学出版社, 2000.
[8] Hillston J. A Compositional Approach to Performance Modeling[M]. PhD thesis CST-107-94, Computer Science Department, University of Edinburgh, 1994.
[9] Plotkin G. A Structural Approach to Operational Semantics Report DAIMI FN-19[M]. Computer Science Department, Aarhus University, September 1981.
[10] Moller F, Tofts C. A Temporal Calculus of Communicating Systems[M]. In CONCUR90, page33—56, springer, Berlin, LNCS458, 1990.
[11] Giacalone A, Jou C-C, Smolka S. Algebraic Reasoning for Probabilistic Concurrent Systems[M]. In Proceedings of working conference on programming concepts and methods IFIP TC2, Sea of Galilee, Israel, 1990.
[12] Hermanns H, Herzog U, Hillston J. Stochastic process algebras-a formal approach to performance modeling 1995. http://www.home.cs.utwente.nl/~hermanns/mypoper.html
[13] Gotz N, Herzog U, Rettelbach M. TIPP-Introduction and Application to Protocol Performance Analysis[M]. Technical Report, University of Erlangen, 1993.
[14] Buchholz P. Markovian Process Algebra: Composition and Equivalence[M]. Proceedings of process Algebra and performance modeling workshop 1994.
[15] Hermanns H, Rettelbach M. Sytax, Semantics, Equivalence and Axioms for MTIPP In Proc. of the 2nd workshop on Process Algebras and Performance Modeling[M], 71—88. Regensberg/Erlangen, July 1994.

Formal Method on Network Architecture Modeling and Performance Evaluation-Stochastic Process Algebra

FENG Yan-li[1,2], YU Zheng-wei[1], PAN Yun[1], LIU Ke-jian[1]

(1. China University of Mining Techenology, Beijing 100083, China)
(2. Shandong Insititute of Business and Techenology, Shandong Yantai 264005, China)

Abstract: Stochastic process algebra (SPA) has been developed from classical process algebra as a formal method to qualitative and quantitative analysis of concurrence system, especially in computer network and communication protocol. In this paper, we will first introduce the syntax and operational semantics of SPA. As a application, we analyze Token ring LAN model. Furthermore, three kinds of equivalent relations have been discussed in order to solve state space explosion problem.

Keywords: formal methods; stochastic process algebra; performance evaluation