第 18 卷 第 3 期
Vol. 18 No. 3

控 制 与 决 策
Control and Decision

2003 年 5 月
May 2003

文章编号: 1001-0920(2003)03-0348-03

闭排队网络基于并行仿真的灵敏度估计和优化算法

殷保群, 代桂平, 周亚平, 谭小彬, 奚宏生
(中国科学技术大学 自动化系, 安徽 合肥 230027)

摘 要: 基于 Markov 性能势理论, 对一类闭排队网络的灵敏度估计和优化, 建立了一种行之有效的并行仿真算法。采用公共随机数, 使所有的处理器使用相同的样本轨道, 以减少各个处理器之间的通讯时间。在一台 SPMD 并行计算机上的仿真实例表明, 该并行仿真算法对于闭排队网络的优化能显著地提高运算速度。

关键词: 灵敏度估计; 闭排队网络; 性能势; 并行仿真; 优化

中图分类号: TP202　　文献标识码: A

Sensitivity estimates and optimization algorithms based on parallel simulation for a class of closed queuing networks

YIN Bao-qun, DAI Gui-ping, ZHOU Ya-ping, TAN Xiao-bin, XI Hong-sheng
(Department of Automation, University of Science and Technology of China, Hefei 230027, China)

Abstract: Based on Markov performance potential, an efficient parallel simulation algorithm is presented for sensitivity estimates and optimization of a class of closed queuing networks. The Common Random Number is applied to make all processors generate the same sample path, which removes the large broadcasting cost at the price of only adding a little workload. The simulation experiments on an SPMD parallel computer show that these algorithms can achieve nearly linear speedup for optimization of a class of closed queuing networks.

Key words: Sensitivity estimate; Closed queuing networks; Performance potential; Parallel simulation; Optimization

1 引　　言

在随机离散事件动态系统(DEDS)的性能灵敏度分析方面, 文献[1]已提出了一个新理论: Markov 性能势理论, 给出了稳态性能关于无穷小矩阵摄动的灵敏度公式, 公式中涉及到的量可通过分析 Markov 过程的一条样本轨道获得。文献[2,3]将该理论应用于一类闭排队网络的性能灵敏度分析, 给出了稳态性能关于网络参数的灵敏度公式。

针对 DEDS 仿真计算量大、耗时长的特点, 采用多处理器并行仿真来加快仿真速度显得尤为重要。本文基于性能势理论, 对一类闭排队网络的性能灵敏度估计和优化, 提出了一种效率较高的并行仿真算法。

2 性能势和灵敏度估计

考虑一个具有 M 个服务节点和 N 个顾客的闭排队网络, 每个服务节点具有无限容量的缓冲器, 并服从先到先服务的规则。系统的状态空间为

收稿日期: 2002-01-11; 修回日期: 2002-03-25。
基金项目: 国家自然科学基金资助项目(69974037); 安徽省自然科学基金资助项目(01042308)。
作者简介: 殷保群(1962—), 男, 安徽全椒人, 副教授, 博士, 从事非线性系统展开、DEDS 性能分析、优化及其应用等研究; 代桂平(1977—), 女, 山东平度人, 硕士生, 从事 DEDS 性能灵敏度估计与优化算法等研究。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

$$\Phi = \{n = (n_1, n_2, \cdots, n_M): \sum_{i=1}^{M} n_i = N\}$$

其中: n_i 表示在服务节点 i 处的顾客数, 共有 $K = \begin{pmatrix} N+M-1 \\ N \end{pmatrix}$ 个状态, 路径转移概率矩阵为 $Q = [q_{ij}]$, 服务节点 i 在系统状态为 n 时的服务率为 $\mu_{i,n}$, 则这样一个网络的状态过程可用一个正常返、不可约的Markov 过程 $Y = \{N_t; t \geq 0\}$ 来描述。记该过程的无穷小矩阵为 A, $\pi(n)$ 是状态 n 的稳态概率, π 是稳态概率向量, 则 $Ae = 0, \pi A = 0, \pi e = 1$, 其中 $e = (1, 1, \cdots, 1)'$ 为 K 维列向量。

定义 $v(n) = (\mu_{1,n}, \mu_{2,n}, \cdots, \mu_{M,n})$ 为状态 n 时的一个允许策略, 记 $v = \{v(n)\}_{n\in\Phi}$ 为闭排队网络的一个策略, U 为所有允许策略集, 假定 U 是 M 维实Euclid空间 R^M 中的一个有界闭集。设 $f: \Phi \to R$ 为一个性能函数, 且假设 f 也与策略 v 有关, 即 $f = f(n, v(n))$, $n \in \Phi$, 且关于 v 是可微的。定义稳态性能为

$$J(v) = \sum_{n\in\Phi} \pi(n) f(n, v(n)) = \pi f \tag{1}$$

有关实现因子 d_{nm}, 实现矩阵 D, 性能势向量 g 等概念及其相关结论, 可参看文献[2, 3]。通过分析网络过程的一条样本轨道给出性能势的估计量, 可获得性能测度关于服务率导数的估计。设 $\{T_l\}_{l=0}^{\infty}$ 是网络过程 Y 的状态转移时刻, 令 $X_l = N_{T_l+}$, 则 $\{X_l\}_{l=0}^{\infty}$ 是过程 Y 的嵌入Markov 链。设 $\epsilon^n(X_l) = 1$, $X_l = n$; $\epsilon^n(X_l) = 0$, $X_l \neq n$。又记 $s_l = T_{l+1} - T_l$, 则可得稳态概率 $\pi(n)$ 和稳态性能 $J(v)$ 的无偏估计分别为

$$\hat{\pi}(n) = \frac{1}{T}\sum_{l=0}^{C-1} \epsilon^n(X_l) s_l, \quad \hat{J}(v) = \hat{\pi} f \tag{2}$$

其中: T 是仿真时间, C 是在时间 $[0, T]$ 上状态转移的次数。记 $S^{\{m\}}(n, v)$ 为系统从初始状态 m 出发首次到达状态 n 的时间, $S_j^{\{m\}}(n, v)$ 是系统第 $j(j = 1, 2, \cdots)$ 次从状态 m 出发首次到达状态 n 的时间, 而在这段时间内网络的性能测度记为 $R_j^{\{m\}}(n)$, 则 $E\{S^{\{m\}}(n, v)\}$, $E\{\int_0^{S_j^{\{m\}}(n,v)} f(N_t^{\{m\}}, v(N_t^{\{m\}}))\,dt\}$ 的无偏估计分别为

$$\hat{E}\{S^{\{m\}}(n, v)\} = \frac{1}{L^{\{m\}}(n)}\sum_{j=1}^{L^{\{m\}}(n)} S_j^{\{m\}}(n, v) \tag{3}$$

$$\hat{E}\{\int_0^{S^{\{m\}}(n,v)} f(N_t^{\{m\}}, v(N_t^{\{m\}}))\,dt\} = \frac{1}{L^{\{m\}}(n)}\sum_{j=1}^{L^{\{m\}}(n)} R_j^{\{m\}}(n) \tag{4}$$

其中 $L^{\{m\}}(n)$ 是事件"过程从状态 m 出发首次到达状态 n"发生的次数。从而实现因子和性能势的无偏估计分别为

$$\hat{d}_{nm} = \hat{E}\{\int_0^{S^{\{m\}}(n,v)} f(N_t^{\{m\}}, v(N_t^{\{m\}}))\,dt\} - \hat{E}\{S^{\{m\}}(n, v)\}\hat{J}(v) \tag{5}$$

$$\hat{g} = \hat{D}^{\tau}\hat{\pi}^{T} \tag{6}$$

基于性能势的性能导数估计为[2, 3]

$$\frac{\partial \hat{J}(v)}{\partial \mu_{i,n}} = \hat{\pi}\frac{\partial A}{\partial \mu_{i,n}}\hat{g} + \hat{\pi}\frac{\partial f}{\partial \mu_{i,n}} \tag{7}$$

3　并行仿真算法

在一般的矩阵并行运算中, 数据通讯的开销占并行运算耗时的比例很大, 本文在所有的处理器上使用公共随机数序列, 同时作样本轨道仿真, 这样不需要在处理器之间传送样本轨道的数据。灵敏度估计的并行仿真算法如下:

1) 给定初始状态 $n = n_0$, 在所有处理器中用同样的随机数种子初始化伪随机数发生器, 这样各个处理器就可产生公共随机数序列;

2) 设当前状态为 $n = (n_1, n_2, \cdots, n_M)$, 取一随机向量 $\xi = (\xi_1, \xi_2, \cdots, \xi_M)$, ξ_k 为相互独立、服从[0, 1) 均匀分布的随机数, 以指数分布率计算服务节点 k 中当前顾客的服务时间

$$t_k = \begin{cases} -\dfrac{1}{\mu_{k,n}}\ln(1-\xi_k), & n_k \neq 0 \\ \infty, & n_k = 0 \end{cases}$$
$$k = 1, 2, \cdots, M \tag{8}$$

3) 找出 $t_i = \min(t_1, t_2, \cdots, t_M)$, 知系统在当前状态 n 停留了时间 t_i 后, 服务节点 i 中有一顾客被服务完, 将要向某一服务节点转移;

4) 取一个服从[0, 1) 均匀分布的随机数 ζ, 如 ζ 满足 $\sum_{l=1}^{j-1} q_{i,l} \leq \zeta < \sum_{l=1}^{j} q_{i,l}$ (其中 $\sum_{l=1}^{M} q_{i,l} = 1$, $q_{i,l}$ 是服务节点 i 到节点 l 的路径转移概率), 在服务节点 i 服务完的顾客则转移到节点 j, 系统则到达状态 $(n_1, n_2, \cdots, n_i - 1, \cdots, n_j + 1, \cdots, n_M)$;

5) 在有序状态表中用折半查找法找出状态序号 $m = (n_1, n_2, \cdots, n_i - 1, \cdots, n_j + 1, \cdots, n_M)$;

6) 所有处理器并行计算参数矩阵;

7) 回到 2) 进行下一次状态转移, 直到达到指定的仿真次数;

8) 使用通常的矩阵并行计算方法计算 $\hat{J}$, $\hat{\pi}$, $\hat{D}$ 和 $\hat{g}$;

9) 所有处理器并行计算 $\nabla \hat{J} = \{\partial \hat{J}(v)/\partial \mu_{i,n}\}$。

4　优化算法和算例

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

考虑 2 中描述的闭排队网络, 设系统状态为 n, 且在服务率为 $v(n)$ 下的瞬时代价函数为 $f(n, v(n)) = c(n, v(n)) + h(n, v(n))$, 其中 $c(n, v(n)) = \sum_{i=1}^{M} c_i(n, v(n))$ 为系统的运行代价, $c_i(n, v(n)) \geq 0$ 表示服务率为 $\mu_{i,n}$ 时的单位时间代价; $h(n, v(n)) = \sum_{i=1}^{M} h_i(n, v(n))$ 为系统的维持代价, $h_i(n, v(n)) \geq 0$ 表示系统内有 n_i 个顾客时的代价, 假设 $f(n, v(n))$ 关于 $v \in U$ 是可微的。

在策略 v 下, 定义系统的平均代价为

$$J(v) = \lim_{T\to\infty} E\left\{\frac{1}{T}\int_0^T f(N_t, v(N_t))\,\mathrm{d}t\right\} \qquad (9)$$

易证式(9) 与式(1) 是等价的。则系统的优化问题为: 寻找策略 $v^* \in U$, 使 $J(v^*) = \min\limits_{v\in U} J(v)$。

在基于串行仿真的优化算法中[4], CPU 运行时间集中在两部分计算上, 第一部分是通过分析样本轨道得到导数估计, 第二部分是在获得的梯度方向上寻找最优的服务率, 而 97% 以上的时间集中在第一部分, 因此并行化的重点也在此。

4.1 优化算法

1) 对初始策略 $v^{(0)}$, 用式(2) ～式(4) 计算 $\pi^{(0)}$, $J(v^{(0)})$, $E\{S^{\{m\}}(n, v^{(0)})\}$ 和 $E\left\{\int_0^{S^{(m)}\{n, v^{(0)}\}} f(N_t^{\{m\}}, v^{(0)})\,\mathrm{d}t\right\}$;

2) 根据式(5) 和式(6), 并行计算 $D^{(v^{(0)})}$ 和 $g^{(v^{(0)})}$;

3) 根据式(7), 所有处理器并行计算矩阵 $\nabla J(v^{(0)})$;

4) 考虑一个随机梯度下降算法

$$v^{(k+1)} = \varphi_{k+1}(v^{(k)} - \alpha_{k+1}\hat{\beta}_{k+1})$$

其中: $\varphi_{k+1}(\cdot)$ 由

$$\varphi_{k+1}(x) = \begin{cases} x, & x \in U \\ v^{(k)}, & 其他 \end{cases}$$

定义; $\hat{\beta}_k$ 是 $\nabla J(v^{(k)})$ 的一个估计; $\{\alpha_k\}$ 是一个正的随机向量序列, 满足:

① $\{\alpha_k\}$ 是非增的;

② 存在常数 $\delta, \bar{A}, \tilde{A}$, $0.5 < \delta \leq 1$, $\bar{A} < \infty$, $\tilde{A} > 0$, 对每个 k, 有 $\tilde{A}K^{-\delta} \leq \alpha_k \leq \bar{A}K^{-\delta}$;

③ 存在有限常数 B_α, 对所有 k, 有 $\frac{1}{\alpha_{k+1}} - \frac{1}{\alpha_k} \leq B_\alpha$。

5) 选定 $\epsilon_1, \epsilon_2 > 0$, 如果

$$\| \hat{J}(v^{(k+1)}) - \hat{J}(v^{(k)}) \|_\infty < \epsilon_1$$

且

$$\| \nabla \hat{J}(v^{(k+1)}) \|_\infty < \epsilon_2$$

则结束计算, 否则, 令 $k + 1 = k$, 返回 2)。

4.2 算 例

设各服务节点的服务率依赖于各自当前的顾客数, 这是服务率依赖于系统状态的一种特殊情况。$\mu_{i,j}$ 为服务节点 i 中顾客数为 j 时的服务率, 设

$$v = \begin{bmatrix} \mu_{1,1} & \mu_{1,2} & \mu_{1,3} & \mu_{1,4} \\ \mu_{2,1} & \mu_{2,2} & \mu_{2,3} & \mu_{2,4} \\ \mu_{3,1} & \mu_{3,2} & \mu_{3,3} & \mu_{3,4} \end{bmatrix}$$

$$Q = \begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.5 & 0.2 & 0.3 \\ 0.3 & 0.2 & 0.5 \end{bmatrix}$$

并行优化仿真结果见表 1。为便于比较, 只列出系统中有1个、5个和15个服务节点时各自仿真的 CPU 时间, 3 种情况的样本轨道长度均为 100 000。

表 1 并行优化仿真结果

初始服务率				优化服务率				$J(v^*)$	节点数	CPU 运行时间 / s
1	1	1	1	1.400 29	0.778 055	0.549 697	0.491 385		1	13 640
1	1	1	1	1.388 03	0.776 002	0.551 021	0.492 249	0.995 458	5	3 117
1	1	1	1	1.393 87	0.779 364	0.546 001	0.488 178		15	1 379
0.1	0.1	0.1	0.1	1.452 63	0.785 6	0.553 298	0.491 38		1	11 540
0.1	0.1	0.1	0.1	1.433 38	0.792 714	0.549 514	0.493 973	0.996 43	5	2 696
0.1	0.1	0.1	0.1	1.422 46	0.787 529	0.549 636	0.492 436		15	1 236
0.40	0.45	0.30	0.60	1.338 65	0.774 777	0.550 921	0.493 243		1	13 328
0.20	0.80	0.45	0.40	1.395 96	0.772 052	0.547 12	0.492 609	0.994 963	5	3 148
0.65	0.20	0.75	0.35	1.368 43	0.785 786	0.550 264	0.492 502		15	1 468
0.27	0.32	0.73	0.54	1.421 23	0.778 004	0.555 088	0.489 476		1	17 000
0.43	0.86	0.23	0.64	1.430 71	0.776 871	0.548 977	0.492 654	0.996 2	5	3 871
0.34	0.77	0.24	0.86	1.439 48	0.782 415	0.549 409	0.493 309		15	1 703

(下转第 354 页)

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

法相比,本文所提出的方法更适合于分析长度较短的周期波动序列。

4.2 主要周期分量的分辨及其周期长度准确性的比较

用一般的谱估计方法分析周期波动序列,可以分辨的主要周期分量的个数就是序列谱估计曲线的峰值点的个数,主要周期分量的周期长度就是相应的峰值所对应的周期长度。按照该种方法分析 CL_1,可知其主周期分量的个数为 2 个,而用各主要周期分量单独分辨的方法,却依次分辨出 5 个周期分量并确定了它们的准确值。可见用一般的谱分析方法很难将各主要周期分量同时分辨出来,也无法确定周期长度的准确值。因此本文提出的依次分辨各主要周期分量,并通过三角函数拟合确定周期长度准确值的方法,对于经济波动的谱分析是很有必要的。

5 结 语

本文针对一般谱分析方法研究经济波动时存在的问题,在对古典的加窗周期图谱估计和现代的最大熵谱估计进行比较分析的基础上,提出一种比前者分辨率更高、比后者更为可靠的适合于我国经济周期波动特点的谱估计方法——两者相结合的谱估计方法,较好地解决了我国经济周期波动序列长度较短、采样间隔不能随意改变的问题,并提出各主要周期分量单独分辨,通过对序列进行三角函数拟合确定周期长度准确值的方法,较好地解决了各主要周期分量难以同时准确分辨的问题。通过对天津经济波动的实证分析,将本文方法与一般的谱分析方法进行比较,说明了该方法的可行性和有效性。但由于经济波动的规律极为复杂,单纯从统计数据的表面分析是不够的,这就要求我们要着眼于我国的现实情况,将定性分析与定量分析相结合来分析和研究经济波动。

参考文献(References):

[1] 陈文海,顾岚. 谱分析在我国经济波动理论中的应用[J]. 数理统计与管理,1994,13(1):1-8.

[2] 陈述云. 谱分析在经济预测中的应用[J]. 预测,1995,14(6):49-52.

[3] 张吉峰. 谱分析在测定时间序列周期中的应用[J]. 预测,1994,13(4):40-45.

[4] Granger C W J, Hantanaka M. *Spectral Analysis of Economic Time Series*[M]. Princeton University Press, 1964.

[5] 杨位钦,顾岚. 时间序列分析与动态数据建模[M]. 北京:北京理工大学出版社,1988.

[6] 潘士先. 谱估计和自适应滤波[M]. 北京:北京航空航天大学出版社,1991.

(上接第 350 页)

5 结 语

根据 Markov 性能势理论及其在闭排队网络中的应用,本文设计了一种高效的并行仿真算法,并将其应用于闭排队网络优化问题中。试验结果表明,本文的并行算法具有很好的加速比和较高的效率。

参考文献(References):

[1] Cao X R, Chen H F. Perturbation realization, potentials, and sensitivity analysis of Markov processes[J]. *IEEE Trans Automatic Control*, 1997, 42(10): 1382-1393.

[2] 殷保群,周亚平,奚宏生,等. 闭排队网络当性能函数与参数相关时的性能灵敏度分析[J]. 控制理论与应用,2002,19(2):311-312.
(Yin Baoqun, Zhou Yaping, Xi Hongsheng, et al. Sensitivity analysis of performance with parameter-dependent performance functions in closed queueing networks[J]. *Control Theory & Application*, 2002, 19(2): 311-312.)

[3] Yin Baoqun, Zhou Yaping, Xi Hongsheng, et al. Sensitivity formulas of performance in two-server cyclic queueing networks with phase-type distributed service times[J]. *Int Trans Operation Research*, 1999, 6(6): 649-663.

[4] 周亚平,殷保群,奚宏生,等. 一类闭排队网络基于性能势的优化算法[J]. 中国科学技术大学学报,2000,30(2):151-157.
(Zhou Yaping, Yin Baoqun, Xi Hongsheng, et al. Algorithms of decentralized optimization for a class of closed queueing networks by using performance potentials[J]. *J of China University of Science and Technology*, 2000, 30(2): 151-157.)

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net