Nonlinear Programming Algorithms for Performance Modelling of Computer Networks

Z. L. KROUGLY, D. A. STANFORD,

Department of Statistical & Actuarial Sciences,
University of Western Ontario, London, Ont., Canada N6A5B9

Keywords: Computer networks, queueing network models, priority approximation, global balance solution.

1 Introduction

A number of nonlinear programming algorithms are proposed for computer networks analysis using queueing network models. The efficient iterative technique is due to the use of sensitivity in closed queueing networks. We compare the approximations with exact results based on the global balance solution. To support the modelling technique in multiclass queueing network models, we developed the software package ZEDNET, a Windows application written in C++. The package includes a graphical interface (input and output), a number of projects (product form solutions, non product form solution for priority and other approximations in closed queueing networks, Markov chain routines, global balance solution), and a set of analytical algorithms implementing theorems from queueing theory (BCMP, mean value analysis, etc.).

2 Background

In [1,3,5] derivatives have been obtained for the performance metrics as a function of the service demands and service rates. Sensitivity analysis can be used to develop efficient algorithms, and to solve optimization problems in queueing networks [3,5]. In [3] was illustrated the technique to substantiate numerical optimization methods of solution of the optimization problem in a closed queueing network of guaranteed convergence.

Consider a closed product form queuing network with M service centers and R customer classes. The number of class v customers ($v = 1, ..., R$) is equal to

n_v. The visit ratio e_{iv} is the solution to the system of linear equations $e_{iv} = \sum_{j=1}^{M} e_{jv} P_{jiv}; i = 0, ..., M; v = 1, ..., R$, where P_{jiv} are transition probabilities. Using e_{iv} we can also compute the relative utilization $x_{iv} = e_{iv}/\mu_{iv}$, where $1/\mu_{iv}$ is the mean service time for a class v customers at service center i. Let's specify the input process through the interarrival time distribution with the given arrival rates Λ_{0v} for class v customers.

The iterative procedure presented below allows for different arrival process definitions.

Algorithm 2.1

Step 0. Initialization: Set initial values of service rates and the mean size of the source $\mu_{0v}^{(0)} = \Lambda_{0v}; L_{0v}^{(0)} = n_v; v = 1, ..., R$.

Step 1. The parameters $\mu_{0v}^{(s)}$ and $L_{0v}^{(s)}$, (steps $s = 0, 1, ...$) are used for calculation of transition probabilities P_{ijv} and service rates $\mu_{iv}(i, j = 0, ..., M; v = 1, ..., R)$.

Step 2. The calculation of the BCMP queuing network is performed.

Step 3. The calculation of $L_{0v}^{(s+1)} = \varphi_1(L_{0v}^{(s)}, \mu_{0v}^{(s)})$ and $\mu_{0v}^{(s+1)} = \varphi_2(L_{0v}^{(s+1)}, \mu_{0v}^{(s)})$ is performed as a function of parameters, defined on the previous steps, where $\mu_{0v}^{(s+1)}$ is given by iterative formula (2.1) or (2.2)

$$\mu_{0v}^{(s+1)} = L_{0v}^{(s+1)}/[U_{0v}(1/\Lambda_{0v} - T_v)] \tag{2.1}$$

$$\mu_{0v}^{(s+1)} = N_v \Lambda_{0v}/U_{0v}, v = 1, ..., R; \tag{2.2}$$

U_{iv} is the utilization at service center i for class v customers; T_v is the mean response time for class v customers,

$$U_{0v} = \gamma_1(L_{0v}^{(s+1)}, \mu_{0v}^{(s)}) \tag{2.3}$$

$$T_v = \sum_{i=1}^{M} (\frac{\mu_{iv} U_{iv}}{\mu_{0v} U_{0v}}) \frac{L_{iv}}{\mu_{iv} U_{iv}} = \frac{n_v - L_{0v}^{(s+1)}}{\lambda_{0v}} = \frac{n_v - L_{0v}^{(s+1)}}{\mu_{0v}^{(s)} U_{0v}}, v = 1, ..., R; \tag{2.4}$$

L_{iv} and λ_{0v} are accordingly mean queue length at center i for class v customers and throughput for class v customers.

Step 4. If $L_{0v}^{(s)}$ and $\mu_{0v}^{(s)}$ not converged, return to Step 1. Otherwise, stop.

If the input stream is given conventionally through μ_{0v} the simpler algorithm based on the system of the nonlinear equations: $L_{0v}^{(s+1)} = \varphi_1(L_{0v}^{(s)}), v = 1, ..., R$.

Notice that the iterative formula (2.2) is superior and provides better algorithm's convergence because partial derivatives $\frac{\partial \mu_{0v}^{(s+1)}}{\partial \mu_{0v}^{(s)}}$ by (2.2) are larger than calculated by (2.1).

The partial derivatives for multiclass closed networks are given by the following [1]:

$$\frac{\partial G(\bar{n})}{\partial x_{iv}} = \frac{G(\bar{n})}{x_{iv}} L_{iv}(\bar{n}) \tag{2.5}$$

$$\frac{\partial \lambda_r(\bar{n})}{\partial x_{iv}} = \frac{\lambda_r(\bar{n})}{x_{iv}} [L_{iv}(\bar{n}-1_r) - L_{iv}(\bar{n})]; i = 0, ..., M; v, r = 1, ..., R \ , \tag{2.6}$$

where $\lambda_r = \frac{G(\bar{n}-1)}{G(\bar{n})}$ is the throughput for class r customers, $(\bar{n}-1_r) = (n_1, ..., n_r - 1, ..., n_R)$ is the population vector with one class r customer less in the network.

In the next section, we make use of these derivatives in context with algorithm 2.1 in order to develop an efficient algorithm for the approximation of priority queueing networks for which no product-form solution exists.

3 Priority Approximation

A major development in the analysis of closed priority queueing networks is Sevcik's shadow approximation for preemptive priority scheduling, described in [4]. The approximate service rate μ'_{iv} of a class v customer at its dedicated shadow center is calculated using the utilization U_{iv} of the higher priority classes. Therefore, the following iteration algorithm is used to find the approximate solution for the μ'_{iv}.

Algorithm 3.1

Step 0. Transform the original model into the shadow model. Initialize: $U_{iv}^{(0)} = 0, v = 1, ..., R-1$.

Step 1. Compute the shadow service rates

$$\mu'_{iv} = \mu_{iv}(1 - \sum_{k=1}^{v-1} U_{ik}^{(s)}), s = 0, 1, ..., \tag{3.1}$$

where μ_{iv} denotes the actual service rate of a class v in the priority center.

Step 2. Find product form solution for BCMP network with $M + R - 1$ service centers. Compute $U_{ik}^{(s+1)}, v = 1, ..., R-1$.

Step 3. If $U_{ik}^{(s+1)}$ have not converged return to Step 1. Otherwise, stop.

We incorporated a number of algorithms in the priority context. We resort an iterative scheme based on the shadow approximation. Another possibility is to introduce an objective function and use a direct-search procedure. Yet another option is to solve this optimization problem with the assistance of derivative information, and this is described extensively below. Lastly, we present in problem 3.2 an approach in which the arrival process is specified in accordance with the alternate representation introduced in section 2. That is, we specify the input process in terms of the given arrival rates Λ_{0v} and the size of the source n_v for class v customers. This is addressed in problem 3.2 below.

Using m-dimensional vector-valued function $F(\bar{c})$ the nonlinear programming problem for priority approximation can be formally stated as

Problem 3.1

$$\min F(\bar{c}) = \sum_{k=1}^{m} f_k^2(\bar{c}) \tag{3.2}$$

subject to $g_i(\bar{c}) > 0;\ i = 1, ..., 2m,$

where: $f_k(\bar{c}) = \varphi_k(\bar{c}) - c_k;\ g_i(\bar{c}) = \begin{cases} c_i;\ i = 1, ..., m; \\ 1 - c_i;\ i = m+1, ..., 2m; \end{cases}$ $\quad \bar{c} = (c_1, ..., c_m)$

is m-dimensional solution vector; $m = R - 1$; $\varphi_k(\bar{c}) = U_{ik}^{(s+1)}$ is the utilization at the shadow center $k, k = 1, ..., m$.

First assume that closed queueing network has two classes of customers, class 1 has preemptive priority over class 2 in the priority center i. Because

$$\frac{\partial x'_{i2}}{\partial \mu'_{i2}} = \frac{\partial (e'_{i2}/\mu'_{i2})}{\partial \mu'_{i2}} = -\frac{e'_{i2}}{(\mu'_{i2})^2} = -\frac{x'_{i2}}{\mu'_{i2}} \tag{3.3}$$

using (2.6) we get

$$\begin{aligned} \frac{\partial U_{i1}^{(s+1)}}{\partial \mu'_{i2}} &= \frac{1}{\mu_{i1}} \frac{\partial \lambda_{i1}}{\partial \mu'_{i2}} = \frac{1}{\mu_{i1}} \frac{\partial \lambda_{i1}}{\partial x'_{i2}} \frac{\partial x'_{i2}}{\partial \mu'_{i2}} = \frac{1}{\mu_{i1}} \frac{\partial (e_{i1}\lambda_{01}/e_{01})}{\partial x'_{i2}} \frac{\partial x'_{i2}}{\partial \mu'_{i2}} \\ &= \frac{1}{\mu_{i1}} \frac{e_{i1}}{e_{01}} \frac{\lambda_{01}}{x'_{i2}} [L_{i2}(n_1 - 1, n_2) - L_{i2}(n_1, n_2)](-\frac{x'_{i2}}{\mu'_{i2}}) \\ &= \frac{U_{i1}^{(s+1)}}{\mu'_{i2}} [L_{i2}(n_1, n_2) - L_{i2}(n_1 - 1, n_2)] \end{aligned} \tag{3.4}$$

And because

$$\frac{\partial \mu'_{i2}}{\partial U_{i1}^{(s)}} = \frac{\partial[\mu_{i2}(1-U_{i1}^{(s)})]}{\partial U_{i1}^{(s)}} = -\mu_{i2} = -\frac{\mu'_{i2}}{1-U_{i1}^{(s)}} \tag{3.5}$$

the iteration function derivatives are

$$\frac{\partial U_{i1}^{(s+1)}}{\partial U_{i1}^{(s)}} = \frac{\partial U_{i1}^{(s+1)}}{\partial \mu'_{i2}} \cdot \frac{\partial \mu'_{i2}}{\partial U_{i1}^{(s)}} = \frac{U_{i1}^{(s+1)}}{1-U_{i1}^{(s)}}[L_{i2}(n_1-1,n_2) - L_{i2}(n_1,n_2)] \tag{3.6}$$

The derivative for objective function (3.2) is

$$\frac{\partial F(\bar{c})}{\partial U_{i1}^{(s)}} = 2(U_{i1}^{(s+1)} - U_{i1}^{(s)})\{\frac{U_{i1}^{(s+1)}}{1-U_{i1}^{(s)}}[L_{i2}(n_1-1,n_2) - L_{i2}(n_1,n_2)] - 1\}. \tag{3.7}$$

Assume that the customers belong to set of R different priority classes. Using (2.6) and (3.1)

$$\frac{\partial \mu'_{ir}}{\partial U_{iv}^{(s)}} = \partial[\mu_{ir}(1-\sum_{k=1}^{v} U_{ik}^{(s)})]/\partial U_{iv}^{(s)} = -\mu_{ir}$$

$$= -\mu'_{ir}/(1-\sum_{k=1}^{v} U_{ik}^{(s)}); v = 1,...,R-1; r = 2,...,R; r > v \tag{3.8}$$

and the iteration function derivatives are

$$\frac{\partial U_{iv}^{(s+1)}}{\partial U_{iv}^{(s)}} = \frac{\partial U_{iv}^{(s+1)}}{\partial \mu'_{i(v+1)}} \frac{\partial \mu'_{i(v+1)}}{\partial U_{iv}^{(s)}} = \frac{U_{iv}^{(s+1)}}{\mu'_{i(v+1)}}[L_{i(v+1)}(\bar{n}) - L_{i(v+1)}(\bar{n}-1_v)][-\frac{\mu'_{i(v+1)}}{(1-\sum_{k=1}^{v} U_{ik}^{(s)})}]$$

$$= [U_{iv}^{(s+1)}/(1-\sum_{k=1}^{v} U_{ik}^{(s)})][L_{i(v+1)}(\overline{n}-1_v) - L_{i(v+1)}(\overline{n})]; v = 1,...,R-1. \tag{3.9}$$

For $v > r$ we get the same solution as for $v = r$:

$$\frac{\partial U_{iv}^{(s+1)}}{\partial U_{ir}^{(s)}} = \frac{\partial U_{iv}^{(s+1)}}{\partial \mu'_{i(v+1)}} \cdot \frac{\partial \mu'_{i(v+1)}}{\partial U_{ir}^{(s)}}$$

$$= \frac{U_{iv}^{(s+1)}}{\mu'_{i(v+1)}}[L_{i(v+1)}(\bar{n}) - L_{i(v+1)}(\bar{n}-1_v)][-\mu'_{i(v+1)}/(1-\sum_{k=1}^{v} U_{ik}^{(s)})]$$

$$= U_{iv}^{(s+1)}/(1-\sum_{k=1}^{v} U_{ik}^{(s)})][L_{i(v+1)}(\overline{n}-1_v) - L_{i(v+1)}(\overline{n})]; v > r \tag{3.10}$$

and

$$\frac{\partial U_{iv}^{(s+1)}}{\partial U_{ir}^{(s)}} = 0; v < r. \tag{3.11}$$

Using (3.9) - (3.11) the partial derivatives for objective function (3.2) are

$$\frac{\partial F(\bar{c})}{\partial U_{iv}^{(s)}} = 2(U_{iv}^{(s)} - U_{iv}^{(s+1)})$$

$$+2\sum_{k=v}^{R-1}\{(U_{ik}^{(s+1)} - U_{ik}^{(s+1)})[U_{ik}^{(s)}/(1-\sum_{l=1}^{k}U_{il}^{(s)})][L_{i(k+1)}(\overline{n}-1_k)$$

$$-L_{i(k+1)}(\overline{n})]\}; v = 1, ..., R. \tag{3.12}$$

At this point, we turn our attention to the alternate specification for the arrival process, in terms of the interarrival-time parameters. The following m-dimensional vector-valued function is used to find the numerical solution.

Problem 3.2

$$\min F(\bar{c}) = \sum_{k=1}^{m} f_k^2(\bar{c}) \tag{3.13}$$

subject to $g_i(\bar{c}) > 0$; $i = 1, ..., m + R$, where: $f_k(\bar{c}) = \varphi_k(\bar{c}) - c_k$; $\bar{c} = (c_1, ..., c_m)$ is m-dimensional solution vector; $m = 2R - 1$; $\varphi_k(\bar{c}) = U_{ik}^{(s+1)}$ is the utilization at shadow priority center $k, k = 1, ..., R - 1$; $\varphi_{R-1+k}(\bar{c}) = \mu_{0k}$ is service rates in the source for priority class $k, k = 1, ..., R$; $g_i(\bar{c})$ are taking into account $2(R-1) + R = (m + R - 1)$ constrains: $c_i > 0, i = 1, ..., R - 1; c_i < 1, i = R, ..., 2(R-1); c_i > 0, i = 2R - 1, ..., m + R - 1$.

Service rates in the source are determined by (2.1) or (2.2).

The following algorithm is used to find the solution for the problem 3.1.

Algorithm 3.2

Step 0. Transform the original model into the shadow model.

Step 1. Initialize: $U_{iv}^{(0)} = 0, v = 1, ..., R - 1$ and $\mu_{0v}^{(0)} = \Lambda_{0v}$; $L_{0v}^{(0)} = n_v$; $v = 1, ..., R$.

Step 2. The parameters $U_{iv}^{(s)}, \mu_{0v}^{(s)}$ and $L_{0v}^{(s)}$ are determined for calculation of transition probabilities P_{ijv} and service rates $\mu_{iv}(i, j = 0, ..., M; v = 1, ..., R)$.

Step 3. Compute the shadow service rates: $\mu'_{iv} = \mu_{iv}(1 - \sum_{k=1}^{v-1} U_{ik}^{(s)})$

Step 4. Find product form solution for BCMP network with $(M+R-1)$ centers. Compute $U_{iv}^{(s+1)}, v = 1, ..., R-1$ and the estimated solution $\mu_{0v}^{(s+1)}, L_{0v}^{(s+1)}$.

Step 5. Convergence test: If $U_{iv}^{(s)}$, $L_{0v}^{(s)}$ and $\mu_{0v}^{(s)}$ have not converged, return to Step 2. Otherwise, stop.

The nonlinear programming technique to find the numerical solution is provided as follows. We can form a new unconstrained objective function $F(\bar{c}) = f(\bar{c}) + H(\bar{c})$ by adding $H(\bar{c})$ as a penalty function $H(\bar{c}) = \sum_{i=1}^{m} \delta_i g_i^2(\bar{c})$, where δ_i and δ_j are zero if the constraints are satisfied, and unity otherwise. Thus the penalty function vanished inside the feasible region.

Using the penalty function, the initial problem can be replaced with an equivalent unconstrained minimization problem:

$$D(\overline{c},\gamma) = F(\overline{c}) + \gamma H(\overline{c}), \tag{3.14}$$

where the weight coefficients γ are positive and establish a monotonically decreasing sequence $\gamma(l)$ such, so $\gamma(l) > \gamma(l+1), l = 0, 1, 2,$ and $\lim_{l \to \infty} \gamma^{(l)} = 0$.

4 Global Balance Solution for Priority Models

Consider the closed tandem queueing network with two classes of customers. Customer population of each class: $N_1 = 4, N_2 = 4$. The service discipline at the server 1 is FCFS. First class have preemptive priority over second class in the server 2. The service times at each server are exponentially distributed with rates μ_{ir} $(r = 1, 2; i = 1, 2)$. The server rates are $\mu_{11} = 1$; $\mu_{12} = 3$; $\mu_{21} = 1/3$; $\mu_{22} = 1$. The notation $(n_1, k_1; n_2, k_2)$ says that there are: n_1 and k_1 of the first class customers at the server 1 and the server 2 accordingly, $n_1 + k_1 = N_1$, $n_2 + k_2 = N_2$, and $\pi(n_1, k_1; n_2, k_2)$ denotes the probability for that state in equilibrium. First denote $\lambda_1 = \mu_{11}$; $\lambda_2 = \mu_{12}$; $\mu_1 = \mu_{21}$; $\mu_2 = \mu_{22}$.

The state transition diagram is shown in Fig. 4.1.

We set the overall flux into a state equal to the overall flux out of the state for each state. Let $\pi = (\pi_1, ..., \pi_{25})$ represents the steady state probability vector π = ((0,4;0,4), (0,4;1,3), (0,4;2,2), (0,4;3,1), (0,4;4,0),(1,3;0,4), (1,3;1,3), (1,3;2,2), (1,3;3,1), (1,3;4,0),(2,2;0,4), (2,2;1,3), (2,2;2,2), (2,2;3,1), (2,2;4,0),3,1;0,4), (3,1;1,3), (3,1;2,2), (3,1;3,1), (3,1;4,0),(4,0;0,4), (4,0;1,3), (4,0;2,2), (4,0;3,1), (4,0;4,0)).

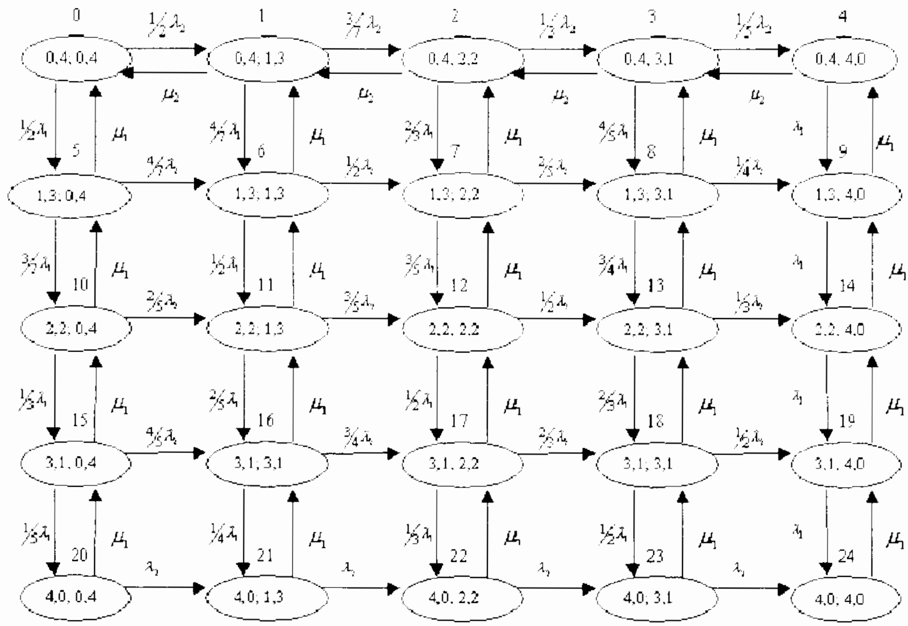


Fig. 4.1. State transition diagram for preemptive priority model

Vector π can be obtained from the system of equations $\pi P = \pi$, where P is the transition probability matrix, defined from the global balance equations:

(1) $p(0,4;0,4)((1/2)\lambda_1 + (1/2)\lambda_2) = p(0,4;1,3)\mu_2 + p(1,3;0,4)\mu_1$

(2) $p(0,4;1,3)((4/7)\lambda_1 + (3/7)\lambda_2 + \mu_2) = p(0,4;0,4)(1/2)\lambda_2 + p(0,4;2,2)\mu_2 + p(1,3;1,3)\mu_1$

...

(24) $p(4,0;3,1)(\lambda_2 + \mu_1) = p(3,1;3,1)(1/2)\lambda_1 + p(4,0;2,2)\lambda_2$

(25) $p(4,0;4,0)\mu_1 = p(3,1;4,0)\lambda_1 + p(4,0;3,1)\lambda_2$

Rewrite this system of equations in the form $\pi Q = 0$, where $Q = P - I$. By solving $\pi Q = 0$ under normalization condition $\sum_{j \in S} \pi_j = 1$, vector π is given by:

0.000101615	0.000263157	0.000752408	0.00216698	0.00613636
4.16882e-005	0.000147798	0.000504313	0.00177914	0.0232453
1.43303e-005	6.86008e-005	0.000290738	0.00131063	0.0732378
3.72216e-006	2.40454e-005	0.000131708	0.000803349	0.221905
5.58324e-007	4.92725e-006	3.66224e-005	0.000328723	0.666701

The marginal probabilities $\pi_i(0)$ that the server i is idle, $i = 1, 2$:
$\pi_1(0) = \pi(0,4;0,4) + \pi(0,4;1,3) + \pi(0,4;2,2) + \pi(0,4;3,1) + \pi(0,4;4,0) = 1 - 0.000101615 + 0.000263157 + 0.000752408 + 0.00216698 + 0.00613636 = 0.009421$;
$\pi_2(0) = 1 - (\pi(0,4;1,3) + \pi(0,4;2,2) + \pi(0,4;3,1) + \pi(0,4;4,0)) = 1 - 0.000263157 + 0.000752408 + 0.00216698 + 0.00613636 = 0.0093189$;

The utilization at server 1 and 2:
$U_1 = 1 - \pi_1(0) = 1 - 0.009421 = 0.990579$;
$U_2 = 1 - \pi_2(0) = 1 - 0.990688 = 0.0093189$.

The throughput at server 1 and 2:
$\lambda_1 = \mu_1 U_1 = (1/3) \cdot 0.990579 = 0.330193$;
$\lambda_2 = \mu_2 U_2 = (1/3) \cdot 0.0093189 = 0.003106$.

5 Numerical Examples

Example 5.1. Consider the closed tandem queuing network with two service centers and two classes of customers. There are four customers in each class. Center 1 operates under a preemptive priority discipline. The service time of customers at each center are exponentially distributed with service rates $\mu_{jv} = 1/s_{jv}(v = 1, ..., R); j = 1, 2$.

The results are presented in Table 5.1. We compare the throughput for priority approximation models with exact results and product form solutions. Exact results were computed for every class using the global balance solution technique, presented in section 4. We observe that while the throughput rates for the high priority class are quite accurate (errors ranging from 1% to 5%), the low priority can be off as much as 30 %.

Example 5.2. A client server system (Fig 5.1) includes k client workstations (center 0) that are connected by Ethernet (CSMA/CD) network (center 1) to a database server. The database server consists of a CPU (center 2) and two disk devices (centers 3 and 4). The client workstations are modelled as infinite center (IS) and submit SQL requests to a database server.

The Ethernet network can be modelled as a load dependent (LD) center [2,4] to represent the effect of network contention. The rate at which the Ethernet delivers packets, given k stations that desire to use the channel, is: $m_p(k) = 1/(L_p/B + S * C(k))$,where $C(k) = (1 - A(k))/A(k)$ denotes the mean number

Table 5.1. Comparison the throughput for priority approximation model with exact result and product form solution

Model #	Class	s_{1v}	s_{2v}	Throughput Approximation	Exact	Product form
Model 1	1	3	3	0.206625	0.215642	0.148148
	2	3	3	0.111717	0.080654	0.148148
Model 2	1	3	1	0.330551	0.330193	0.166650
	2	3	1	0.002782	0.003106	0.166650
Model 3	1	5	2.5	0.193190	0.191647	0.099804
	2	5	2.5	0.006810	0.007962	0.099804

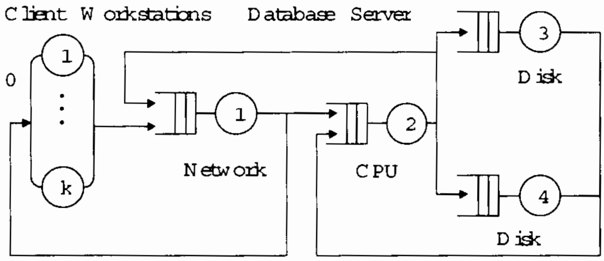
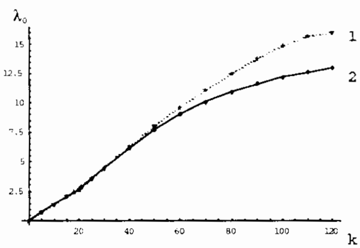


Figure 5.1. Client server queueing network model

Fig. 5.2: Throughput as a function of the number of workstations.

1 - preemptive priority model;
2 - product form model.



of collision per request, and where $A(k) = (1 - 1/k)^{k-1}$ is the probability of a successful transmission.

The other parameters are specified as follows: mean length in bytes per SQL request L_{SQL}=1000 bytes; network bandwidth B= 10 Mbits sec^{-1}; slot duration =51.2 μsec; mean packet length L_p = 1518 bits; maximum length of the data field of a packet L_d=1492 bits; mean number of packets per SQL request $N_{SQL} = 1 + [L_{SQL}/L_d]$=7 packets.

Given k active clients and one database server in the system, there are $k+1$ workstations in the network. But, as the workstation transmits only on request, there are no collisions if there is only one client active. Thus, considering the number of sent packets per SQL request, the service rate of the network, measured in SQL requests per second, is: $\mu_{net}(k) = \begin{cases} \mu_p(1)/N_{SQL}, k = 1 \\ \mu_p(k+1)/N_{SQL}, k > 1. \end{cases}$

The client server model is evaluated using product form model and priority model. For the last model, the first class has preemptive resume priority over second class at CPU (center 2). The transition probabilities and service rates for class 1 and class 2 are as follows: $P_{011} = 1.0; P_{101} = 0.5, P_{121} = 0.5, P_{211} = 0.5, P_{231} = 0.2, P_{241} = 0.3; P_{321} = 1.0; P_{421} = 1.0; \mu_{01} = 0.2$ sec^{-1}, $\mu_{11} = \mu_{net}(k), \mu_{21} = \mu_{CPU} = 32.2$ $sec^{-1}, \mu_{31} = 29.6$ $sec^{-1}, \mu_{41} = 15.5$ $sec^{-1}; P_{012} = 1.0; P_{102} = 0.5, P_{122} = 0.5, P_{212} = 0.15, P_{222} = 0.2, P_{232} = 0.4, P_{242} = 0.25; P_{322} = 1.0; P_{422} = 1.0; \mu_{02} = 0.15$ $sec^{-1}, \mu_{12} = \mu_{net}(k), \mu_{22} = \mu_{CPU} = 25.4$ $sec^{-1}, \mu_{32} = 29.6$ sec^{-1} and $\mu_{42} = 15.5$ sec^{-1}. The number of workstations in class 2 is $k_2 = 20$. The number of workstations in class 1 (denoted k_1) is varying.

The throughput $\lambda_0 = \lambda_{01} + \lambda_{02}$ as a function of the number of workstation $k = k_1 + k_2$ are shown on the Fig. 5.2.

Numerical examples were run to evaluate the performance of a variety of classical computer science applications to address the accuracy of the approximations, and to compare the convergence speed of the various nonlinear approximation techniques. We estimated the execution time for priority approximation models. Analysis demonstrates that optimization methods using derivatives can give significant improvement for convergence speed. We observe that the quasi-Newton method and the iterative algorithm were generally superior at minimizing the $F(\bar{c})$ function. Conjugate gradient method appears to be nearly as satisfactory as the quasi-Newton method. As expected, the search algorithms were slower than the algorithms that used derivatives, but what is interesting is the high ranking of Powell's algorithm.

References

1. K. Gordon, L. Dowdy, The Impact of Certain Parameter Estimation Errors in 0ueueing Network Models, in; Proc. Performance '80, Toronto, Canada printed as Performance Evaluation Review 2(1980), pp 3-9.
2. G. Haring, J. Luthi, S. Majumdar, Mean Value Analysis for Computer Systems with Variabilities in Workload, in; Proc. IEEE Int. Computer Performance and Dependability Symp. (IPDS '96), Urbana-Champaign, September, 1996, pp. 32-41.
3. Z. Krougly, M. Murshtein, Computational Algorithms of Optimization of Closed Queuing Networks. Automation and Remote Control, 7(1990) 926-936.
4. E. Lazowska, J. Zahorjan, G. Graham, K. Sevcik, Quantitative System Performance : Computer System Analysis Using Queueing Network Models, Prentice-Hall, Englewood Cliffs, N.J., 1984.
5. V. Vishnevsky, Z. Krougly, Optimization of Closed Stochastic Networks. Automation and Remote Control. 2(1987) 173-183.