Mean value analysis in multichain queueing network : an iterative approximation

TECHNISCHE HOGESCHOOL EINDHOVEN

Department of Mathematics and Computing Science

Memorandum COSOR 83-18

Mean value analysis in multichain queueing network: an iterative approximation

by

J. van Doremalen

Eindhoven, the Netherlands

September 1983

MEAN VALUE ANALYSIS IN MULTICHAIN QUEUEING NETWORKS: AN ITERATIVE APPROXIMATION

Jan van Doremalen

Abstract. This paper deals with an approximate analysis of multichain queueing networks with FIFO single server queues. Recently, mean value algorithms have been developed to evaluate mean response times, throughputs , mean queue lengths, etcetera in such networks. The complexity and the storage requirements prohibit an exact evaluation of the mean values in large systems and approximate methods have to be used. Several methods have been proposed, e.g. by Schweitzer /7/, Reiser /4/ and Chandy and Neuse /2/. We will describe a method based on a decomposition of the network and mean value arguments.

Zusammenfassung. Wir beschreiben eine approximative Analyse für gemischte Warteschlangennetzen mit FIFO (first-in first-out) Bedieneinheiten. Vor kurzem, sind auf die Mittlerwertanalyse gegründete Algorithmen entwickelt um in solche Netze Verweilzeiten, Durchsätze, Warteschlangelängen, u.s.w. zu berechnen. Die Komplexität und das Speicherplatzbedarf dieser Algorithmen verunmöglichen eine exakte Berechnung der Mittlerwerten für grosse Systemen und approximative Methoden müssen angewendet werden. Verschiedene Methoden sind vorgeschlagen, z.B. von Schweitzer /7/, Reiser /4/ und Chandy und Neuse /2/. Wir werden eine Methode beschreiben die basiert is auf einer Dekomposition des Netzwerkes und Mittlerwertargumenten.

1. Introduction

This paper deals with an iterative approximation for mean residence times, mean queue lengths and throughputs in mixed open and closed multichain queueing networks. We will consider a network with N single server FIFO queues, a set $\mathcal{O}$ of L open chains and a set $\mathcal{C}$ of R closed chains. At queue n, n = 1,2,...,N, the customers have independent exponential service times with mean w_n. An open chain ℓ, $\ell \in \mathcal{O}$, has a Markov routing given by an irreducible substochastic matrix P^{ℓ} and Poisson instream processes with rate $\lambda_{n\ell}$ at queue n, n = 1,2,...,N. A closed chain r, $r \in \mathcal{C}$, has a Markov routing given by an irreducible stochastic matrix P^r and a fixed number of customers K_r. For reasons of presentation only chains with one customer class will be considered.

Mean residence times, mean queue lengths and throughputs in such queueing networks can be evaluated using a mean value oriented algorithm. Reiser describes the method of mean value analysis for closed queueing networks in /4/ and /5/. Lavenberg and Reiser treat the multichain case for closed networks in /6/. The analysis of mixed multichain networks is presented for instance in Zahorjan and Wong /8/ and Krzesinsky, Teunissen and Kritzinger /3/. The mean value analysis is based on Little's formula and two arrival theorems which hold for queueing networks with a product form solution. The first theorem states that a customer of a closed chain sees the system at a jumpmoment as if in equilibrium with one customer of his own chain removed. The second theorem states that a customer of an open chain sees the system at an arrival, jump or departure moment as if in equilibrium.

The computational complexity and the storage requirements of the algorithm grow exponentially with the number of closed chains and an approximate solution, therefore, has to be recommended. In the literature several methods have been proposed, e.g. by Schweitzer /7/, Reiser /4/ and Chancy and Neuse /2/.

We will present an iterative method based on a decomposition idea and mean value arguments. After a short outline of the mean value algorithm for mixed multichain networks in Section 2, we will describe the nature and the behaviour of our method in Section 3. Two examples are given. In Section 4 we will give some tentative conclusions and will glance at points of further research.

2. The mean value algorithm for mixed multichain networks

For lucidity of presentation we will give a short outline of the mean value algorithm for mixed multichain queueing networks with FIFO single server queues.

Let us introduce some notations. For $n = 1,2,\ldots,N$ and $\ell \in \mathcal{C} \cup \mathcal{O}$ we define

$S_{n\ell}$: mean residence time of a chain ℓ customer at queue n
$\Lambda_{n\ell}$: throughput of chain ℓ customers at queue n
$Q_{n\ell}$: mean number of chain ℓ customers at queue n.

Furthermore the population vector K is defined as $K = (K_1,K_2,\ldots K_R)$. The mean values depend on K and will be denoted as $S_{n\ell}(K)$, $\Lambda_{n\ell}(K)$ and $Q_{n\ell}(K)$, if this dependence is important.

For an open chain $\ell \in \mathcal{O}$ the throughputs $\Lambda_{n\ell}$ at the successive queues are the unique solution of a lineair system,

$$(1) \quad \Lambda_{n\ell} = \lambda_{n\ell} + \sum_{m=1}^{N} \Lambda_{m\ell} P_{mn}^{\ell} , \quad n = 1,2,\ldots,N .$$

For a closed chain $r \in \mathcal{C}$ the auxiliary quantities ϑ_{nr} at the successive queues are defined as the unique solution of the linear system

$$(2) \quad \vartheta_{nr} = \sum_{m=1}^{N} \vartheta_{mr} P_{mn}^{r} , \quad n = 1,2,\ldots,N , \quad \sum_{m=1}^{N} \vartheta_{mr} = 1 .$$

Observe that ϑ_{nr} can be interpreted as the fraction of the total number of visits a customer of chain r brings to queue n.

From the arrival theorem for closed chains follows a relation for the mean response time $S_{nr}(K)$ of a closed chain r customer at queue n,

$$(3) \quad S_{nr}(K) = \left(\sum_{s \in C} Q_{ns}(K-e_r) + \sum_{\ell \in O} Q_{n\ell}(K-e_r) + 1 \right) w_n ,$$

where $K-e_r$ denotes the population vector with one customer of chain r removed. The relation, in fact, says that the response time equals the total amount of work a chain r customer sees in front of him upon arrival at queue n plus his own work. Applying the arrival theorem for open chains, we likewise find for the mean response time of an open chain ℓ customer at queue n,

$$(4) \quad S_{n\ell}(K) = \left(\sum_{r \in C} Q_{nr}(K) + \sum_{s \in O} Q_{ns}(K) + 1 \right) w_n .$$

Note that $S_{n\ell}(K)$ is independent of ℓ for $\ell \in O$. Little's formula gives another relation between mean number of customers and mean response time for open chain ℓ customers at queue n, namely

$$(5) \quad Q_{n\ell}(K) = \Lambda_{n\ell}\, S_{n\ell}(K) .$$

Inserting (5) in (4) and using the independence of $S_{n\ell}(K)$ with respect to ℓ we find for open chain ℓ customers at queue n,

$$(6) \quad S_{n\ell}(K) = \left(\sum_{r \in C} Q_{nr}(K) + 1 \right) \tilde{w}_n$$

where $\tilde{w}_n$ is defined as,

$$(7) \quad \tilde{w}_n = w_n \Big/ \left(1 - \sum_{\ell \in O} \Lambda_{n\ell}\, w_n \right) .$$

Inserting (5) and (6) in (3) we find for a closed chain r customer at queue n,

$$(8) \quad S_{nr}(K) = \left(\sum_{s \in C} Q_{ns}(K-e_r) + 1 \right) \tilde{w}_n .$$

Note that the factor $1 - \sum_{\ell \in O} \Lambda_{n\ell}\, w_n$ in $\tilde{w}_n$ can be seen as a fluid dynamic adjustment of the workrate of server n to deal with the influence of open chain work. Equation (8) corresponds with equation (15) in Akyildiz and Bolch /1/.

Applying Little over the network we find a relation for the throughput of closed chain r customers at queue n,

$$(9) \quad \Lambda_{nr}(K) = \vartheta_{nr} K_r / \sum_{m=1}^{N} \vartheta_{mr} S_{mr}(K) .$$

Again applying Little but now on a specific queue we find for the mean number of customers of chain r at queue n,

$$(10) \quad Q_{nr}(K) = \Lambda_{nr}(K) S_{nr}(K) .$$

For the closed chains we can evaluate the mean values from the recursive scheme defined by equations (8), (9) and (10). The equations correspond with equations (3.1), (3.3) and (3.4) in Reiser and Lavenberg /6/ for closed networks.
Afterwards equations (6) and (5) give the mean values for the open chains.

3. An approximation method for closed multichain networks

In a queueing network with many closed chains the evaluation of the mean value scheme becomes problematic. The great complexity and the large storage requirements of the algorithm are caused by the fact that every mutual influence of the chains has to be incorporated. However, it will be clear that many of these influences will be relatively small. Our approximation method is based on this observation. We propose a decomposition of the network such that each chain will be analysed separately. To bind the chains the mutual influence will be approximated using a mean value argument. Only closed multichain networks will be considered. As we have seen in Section 2, this is not a restriction.

Consider a single chain $r \in C$ with K_r customers. To evaluate mean values for this particular chain we introduce a recursive single chain mean value scheme. For $k = 1,2,...,K_r$ we have

$$(11) \quad S_{nr}(k) = (Q_{nr}(k-1) + A_{nr}(k-1) + 1)w_n, \qquad n = 1,2,...,N ,$$

$$(12) \quad \Lambda_{nr}(k) = \vartheta_{nr} k_r / \sum_{m=1}^{N} \vartheta_{mr} S_{mr}(k) , \qquad n = 1,2,...,N ,$$

$$(13) \quad Q_{nr}(k) = \Lambda_{nr}(k) S_{nr}(k) , \qquad n = 1,2,...,N .$$

where we start with $Q_{nr}(0) = 0$. The term $A_{nr}(k-1)$ reflects the mean number of customers of other chains a customer of chain r sees in front of him upon his arrival at queue n if k customers of his own chain are in the system. Using equations (8) through (10) one can verify that the scheme (11) through (13) is exact, if we set

$$(14) \quad A_{nr}(k) = \sum_{s \neq r} Q_{ns}(K - (K_r-k)e_r) , \qquad k = 0,1,...,K_r-1 .$$

However, to compute the scheme this way, one would have to use the exact mean value scheme and that we just wanted to avoid. We propose an approximation for the $\Lambda_{nr}(k)$'s based on the idea that customers of chain r see upon a jumpmoment their own chain as if in equilibrium with one customer removed and the other chains as if in global equilibrium. This idea leads to the formulation of a set of implicit equations, for $r \in C$ and $k = 1,2,\ldots,K_r$ given by,

$$(15)\quad S^*_{nr}(k) = (Q^*_{nr}(k-1) + \sum_{\ell \neq r} Q^*_{n\ell}(K_\ell) + 1)w_n\ , \qquad n = 1,2,\ldots,N\ ,$$

$$(16)\quad \Lambda^*_{nr}(k) = \vartheta_{nr}\, k_r \Big/ \sum_{m=1}^{N} \vartheta^*_{mr}\, S_{mr}(k)\ , \qquad n = 1,2,\ldots,N\ ,$$

$$(17)\quad Q^*_{nr}(k) = \Lambda^*_{nr}(k)\, S^*_{nr}(k)\ , \qquad n = 1,2,\ldots,N\ ,$$

where $S^*_{nr}(K_r)$, $\Lambda^*_{nr}(K_r)$ and $Q^*_{nr}(K_r)$ give the approximation for $S_{nr}(K_r)$, $\Lambda_{nr}(K_r)$ and $Q_{nr}(K_r)$. To evaluate the approximations for the mean values, one has to solve for these implicit equations. A natural way to do so, is to use an iterative method. Starting with initial values for the $Q_{nr}(K_r)$'s the schemes (15) through (17) are repeatedly evaluated until convergence is established.

Up till now we have not been able to show the convergence of the method and the uniqueness of the solution. However, we have tested our method in a series of examples and it turned out that the method converged relatively fast in all cases. To establish convergence we have compared two successive approximations of the throughputs. To obtain a six decimal precision the number of iteration steps varied from 4 up to 20. For large problems the computation time is only a fraction of the time to evaluate the exact mean value scheme.

The approximations, in general, are quite good, especially for larger models with many chains and many customers. But in extreme cases the approximations can be very bad.

As an example of the bad behaviour of the method, consider the following model of a computer system with two terminal groups.

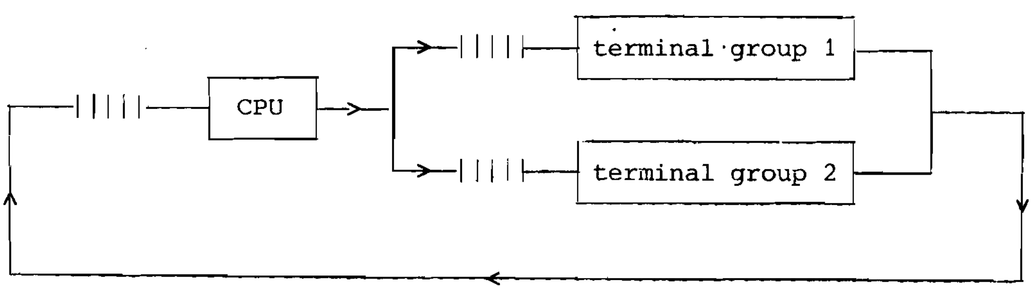

Figure 1. A computer system model.

Terminal group 1 consists of K_1 active terminals each with exponential think time with mean 100 seconds. The K_2 active terminals of group 2 have exponential think times with mean 10 seconds. The jobsizes at the CPU (central processor unit) are exponential with mean 10 seconds. If the service discipline at the CPU is FIFO and all thinktimes and jobsizes are independent then the network has a product form solution and the mean value algorithm can be invoked to evaluate the mean values. In Table 1 we have compared the utilization of the CPU and the response times of the individual terminals of the two groups for the exact and the approximative method. The approximations are very bad because of the fact that the approximation assumption is far from being accurate.

POPULATION		UTILIZATION CPU		RESP. TIME 1		RESP. TIME 2	
K1	K2	EXACT	APPROX	EXACT	APPROX	EXACT	APPROX
10	1	.948	.899	41.44	43.06	31.46	40.10
10	2	.637	.604	56.42	58.01	46.88	54.99
10	3	.583	.558	70.86	72.22	62.55	69.09

Table 1. Results for the computer system model.

Another example, a model of a communication network with window flow control, shows the behaviour of the method for a more complicated network. Consider the network of Figure 2.

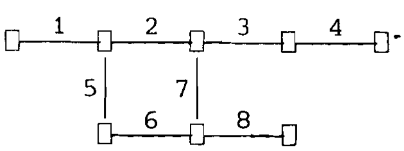

Figure 2. A communication network.

CHAIN	ROUTING
1	1->2->3->4
2	2->7->8
3	6->7->3->4
4	5->6->8
5	8->6->5->1
6	3->2->5

Table 2. Routing table

The arcs, numbered 1 through 8, are the communication channels and the nodes the software interfaces. The channels are modelled as single server FIFO queues, which handle requests in independent exponential times with mean 1 second. The service times at the interfaces are neglegible. A chain is defined by a fixed sequence of channels through which a member of the chain (a message) has to find its way. The routes of the chains are given in Table 2.

The scheduling of the network is as follows. For each chain only a fixed maximum number of customers is allowed in the network. We assume that the input of messages is such that at every moment a message leaves the network a new message of the same chain enters the network. Then the network can be analysed as a closed multichain network.

In Tables 3 and 4 we have pictured the utilizations of the channels and the mean response times of the chains for three control mechanisms. Case i, i = 1,2,3, corresponds with a scheduling where for each chain i customers are allowed in the system.

CHANNEL	UTILIZATION 1		UTILIZATION 2		UTILIZATION 3	
	EXACT	APPROX	EXACT	APPROX	EXACT	APPROX
1	.346	.336	.463	.450	.521	.507
2	.637	.604	.851	.773	.890	.847
3	.583	.558	.766	.732	.850	.815
4	.363	.351	.497	.478	.563	.544
5	.604	.578	.760	.728	.826	.794
6	.567	.548	.743	.717	.827	.800
7	.420	.399	.554	.525	.616	.585
8	.621	.594	.793	.757	.867	.831

Table 3. Utilization of the channels.

CHAIN	RESP. TIME 1		RESP. TIME 2		RESP. TIME 3	
	EXACT	APPROX	EXACT	APPROX	EXACT	APPROX
1	5.56	5.73	8.20	8.47	10.91	11.24
2	4.22	4.49	6.63	7.07	9.15	9.73
3	5.46	5.67	7.92	8.26	10.40	10.84
4	4.58	4.78	7.37	7.67	10.25	10.62
5	6.04	6.18	9.11	9.35	12.19	12.50
6	4.55	4.83	7.42	7.89	10.46	11.06

Table 4. Response times of the chains.

4. Some final remarks

We have described a new method to approximate mean values in multichain queueing networks. The method can be extended in a natural way to networks with FIFO, LIFO, processor-sharing and pure delay or infinite server queues.

The second example shows that the approximations deviate in a certain direction. There is a tendency to overestimate the response times as a consequence of the assumption that arriving customers see the other chains in global equilibrium. We are working on a refinement of the method to deal with the apparent problems. In a forthcoming paper we will analyse the method in more detail and will provide a comparison with other methods.

References

/1/ AKYILDIZ AND BOLCH
ERWEITERUNG DER MITTELWERTANALYSE ZUR BERECHNUNG DER ZUSTANDSWAHRSCHEINLICHKEITEN FUR GESCHLOSSENE UND GEMISCHTE NETZE.
IN: MESSUNG, MODELLIERUNG UND BEWERTUNG VON RECHENSYSTEMEN
EDS. P. KUEHN UND K. SCHULZ. SPRINGER VERLAG BERLIN (1983).

/2/ K.M. CHANDY AND D. NEUSE
LINEAIRISER: A HEURISTIC ALGORITHM FOR QUEUEING NETWORK MODELS OF COMPUTING SYSTEMS
CACM 25(1982): 126-134.

/3/ A. KRZESINSKY, P. TEUNISSEN AND P. KRITINGER
MEAN VALUE ANALYSIS FOR LOAD DEPENDENT SERVERS IN MIXED MULTICLASS QUEUEING NETWORKS
ITR 82-01+00(1982), UNIVERSITY OF STELLENBOSCH, SOUTH AFRICA.

/4/ M. REISER
MEAN VALUE ANALYSIS: A NEW LOOK AT AN OLD PROBLEM
4TH INT. SYMP. ON MODELLING AND PERFORMANCE EVALUATION OF COMPUTER SYSTEMS
VIENNA (1979).

/5/ M. REISER
MEAN VALUE ANALYSIS AND CONVOLUTION METHOD FOR QUEUE DEPENDENT SERVERS IN CLOSED QUEUEING NETWORKS
PERFORMANCE EVALUATION 1(1981): 7-18.

/6/ M. REISER AND S.S. LAVENBERG
MEAN VALUE ANALYSIS OF CLOSED MULTICHAIN QUEUEING NETWORKS
JACM 27(1980): 313-322.

/7/ P. SCHWEITZER
APPROXIMATE ANALYSIS OF MULTICLASS NETWORKS OF QUEUES
PRESENTED AT THE INT. CONF. STOCHASTIC CONTROL AND OPTIMIZATION
AMSTERDAM (1979).

/8/ ZAHORJAN AND E. WONG
THE SOLUTION OF SEPARABLE QUEUEING NETWORKS USING MEAN VALUE ANALYSIS
ACM SIGMETRICS PERF. EV. REV. 3(1981): 80-85.