

An iterative approximation for closed queueing networks with two-phase servers

Citation for published version (APA):

van Doremalen, J. B. M., & Wessels, J. (1983). *An iterative approximation for closed queueing networks with two-phase servers*. (Memorandum COSOR; Vol. 8312). Technische Hogeschool Eindhoven.

Document status and date:

Published: 01/01/1983

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics and Computing Science

Memorandum COSOR 83 - 12

An iterative approximation for closed
queueing networks with two-phase servers

by

J. van Doremalen

J. Wessels

Eindhoven, the Netherlands

August 1983

AN ITERATIVE APPROXIMATION FOR CLOSED
QUEUEING NETWORKS WITH TWO-PHASE SERVERS

by

J. van Doremalen

J. Wessels

Abstract.

This paper deals with an iterative method to obtain approximations for mean values in closed queueing networks with two-phase servers. The servers are particular in a sense that the first phase, a preparatory one, can be done while no customers are at the server.

1. Introduction

The study of queueing networks using exact and computationally attractive methods like the convolution algorithm (see: Reiser and Kobayashi [1975]) and the mean-value algorithm (see: Reiser and Lavenberg [1980] and Reiser [1981]) is restricted to a very special class of networks, the networks with a product-form solution. However, we have the feeling that the use of iterative approximations gives a good tool to study networks not fitting in the exact methods mentioned above. This paper deals with an example, an iterative method to approximate steady-state quantities in a closed queueing network with a special kind of two-phase servers.

The system to be considered is essentially a Gordon and Newell network. However, some of the queues are special. Such a special queue is a single server FCFS queue where the service falls apart in two independent negative exponentially distributed phases. The first phase is a preparatory one and can be executed while no customers are present at the queue, the second phase can be executed only if the customer is present. No more than one preparatory phase can be done in advance. This feature arises in a natural way in several examples. For us the impetus to study this type of model came from the analysis of a container terminal for sea-bearing ships as performed by K.M. van Hee. In this example the cranes are the two-phase servers.

With the two-phase servers the network can still be analyzed as a continuous-time Markov-process on a finite state space. However, the solution no longer has the product-form property and a direct mean-value approach is not possible. An exact solution of the corresponding set of equilibrium equations is very unattractive from a computational point of view.

It seems possible, however, to incorporate the first order effects of the two-phase server in the mean service times and, therefore, we will try to develop a method to approximate the original network with two-phase servers by a Gordon and Newell network with adjusted mean service times. The method will be iterative and partly based on mean-value arguments. For simplicity of reasoning we will restrict ourselves to single chain networks with single server FCFS queues. An extension to more general networks is straight-forward.

In Section 2 we give the notations and the mean-value scheme to evaluate mean values in a single chain network with single server FCFS queues of the product-form class. In Section 3 it is demonstrated how to approximate with an iterative method mean residence times, throughputs and mean queue lengths in the network with two-phase servers. Section 4 is devoted to some numerical experiments and Section 5 to a few concluding remarks.

2. Notations and the mean-value scheme

We consider a closed queueing network with N single server FCFS queues, where all queues are two-phase servers of the kind described in the introduction. The routing through the network is Markovian with routing matrix P and there is one customer class of size K . The following notations are introduced,

- (1) $S_n(K)$: mean residence time (waiting plus service) of a customer at queue n , if K customers are in the system.
- (2) $\Lambda_n(K)$: throughput of queue n , if K customers are in the system.
- (3) $L_n(K)$: mean queue length (waiting plus service) at queue n , if K customers are in the system.
- (4) w_n : mean service time at queue n , if queue n is a two-phase server then $w_n = w_{n_1} + w_{n_2}$ with w_{n_i} the mean service time for the i -th phase.

Furthermore, we define the auxiliary quantities ϑ_n , $n = 1, 2, \dots, N$, as the unique solution of

$$(5) \quad \vartheta_n = \sum_{m=1}^N \vartheta_m P_{mn}, \quad n = 1, 2, \dots, N \quad \text{and} \quad \sum_{n=1}^N \vartheta_n = 1.$$

Observe that ϑ_n can be interpreted as the fraction of the total number of visits a customer brings to queue n .

The problems in analyzing the system arise in two ways. Firstly, the service time of a customer is not necessarily negative exponentially distributed. Secondly, the effective service time distribution is in general different for the first customer of a busy cycle and worse, is even unknown as it is not clear on forehand how much work has been done on the preparatory phase in the preceding idle period.

However, if we assume that $w_{n_1} = 0$, for $n = 1, 2, \dots, N$, then every server is an ordinary exponential one. In that case it is well-known (confer Reiser [1979], Reiser and Lavenberg [1980] and Reiser [1981]) that mean residence times, throughputs and mean queue lengths can be evaluated using a recursive mean value scheme. This scheme is based on two principal arguments: Little's formula and an arrival theorem, which states that in such networks as described above, a customer sees upon a jumpmoment the system in equilibrium as if one customer less was in the system. For later use we develop the scheme below.

Using the arrival theorem one finds, still assuming that $w_{n_1} = 0$ for $n = 1, 2, \dots, N$, a relation for the mean residence time $S_n(K)$ at queue n ,

$$(6) \quad S_n(K) = L_n(K-1)w_n + w_n, \quad n = 1, 2, \dots, N.$$

Using Little's formula over the complete system we find for the throughput $\Lambda_n(K)$ at queue n ,

$$(7) \quad \Lambda_n(K) = \vartheta_n K / \sum_{m=1}^N \vartheta_m S_m(K) \quad , \quad n = 1, 2, \dots, N.$$

Eventually applying Little's formula to a single queue n we have for the mean queue length $L_n(K)$ at queue n ,

$$(8) \quad L_n(K) = \Lambda_n(K) S_n(K) \quad , \quad n = 1, 2, \dots, N.$$

Starting with $L_n(0) = 0$, $n = 1, 2, \dots, N$, the mean values can be evaluated recursively using the relations (6) through (8).

3. An iterative approximation for the network with two-phase servers

In this section we will describe an iterative method to approximate the network with two-phase servers by a Gordon and Newell network with single-server FCFS queues. We are interested in mean values only and will try to derive relations like (6) through (8) to evaluate these values. It should be noted that relations (7) and (8) are based on Little's formula and therefore will hold in the network with two-phase servers also. It is relation (6) which is violated by the introduction of two-phase servers, and consequently, we have to look for an approximation of this relation.

We note that a customer arriving at a two-phase server queue n can find the preparatory phase executed already if he is the first one in a busy period. This particular customer has an expected service time of w_{n2} only. It is

appealing to introduce an adjusted average workload \tilde{w}_n at queue n as

$$(9) \quad \tilde{w}_n = (1 - a_n)w_{n_1} + w_{n_2},$$

and then to approximate the original network by a Gordon and Newell network with these adjusted loads. To find appropriate values for the a_n 's, $n = 1, 2, \dots, N$, is the item of the rest of this section.

Observe that by introducing relation (6) with the adjusted average workloads \tilde{w}_n , implicitly an "arrival theorem"-like argument is introduced.

An interpretation of a_n is that it reflects the probability that an arriving customer finds phase one already executed. As only the first customer of a busy cycle can find phase one executed we can write a_n as $a_n = b_n c_n$, where b_n denotes the probability that an arriving customer is the first one of a busy cycle and where c_n denotes the probability that phase one has been executed upon arrival, if the customer is the first one of a busy cycle.

Assume that we have a set a_n , $n = 1, 2, \dots, N$. Using the mean-value idea that an arriving customer sees the system in equilibrium as if one customer less was in the system, we can approximate b_n by b'_n given as

$$(10) \quad b'_n = 1 - \Lambda_n(K-1)\tilde{w}_n, \quad n = 1, 2, \dots, N,$$

where $\tilde{w}_n = (1 - a_n)w_{n_1} + w_{n_2}$ and where the throughput of queue n $\Lambda_n(K-1)$ is evaluated by scheme (6) through (8) with workload \tilde{w}_n .

Note that the mean-value idea in general need not be true.

In the idle period preparatory work is done and to determine c_n we consequently need information on the idle period length distribution. If we assume that the idle period length is negative exponentially distributed with mean I' , then c_n can be approximated by c'_n as

$$(11) \quad c'_n = \frac{w_{n1}^{-1}}{w_{n1}^{-1} + I'} \quad , \quad n = 1, 2, \dots, N.$$

To determine I' we observe that the fraction of time there are no customers at queue n is given by $1 - \Lambda_n(K)\tilde{w}_n$. The number of idle periods per unit time is equal to the number of customers per unit time who are the first ones of a busy period and can be approximated by $\Lambda_n(K)b'_n$. For I' we thus have the approximation

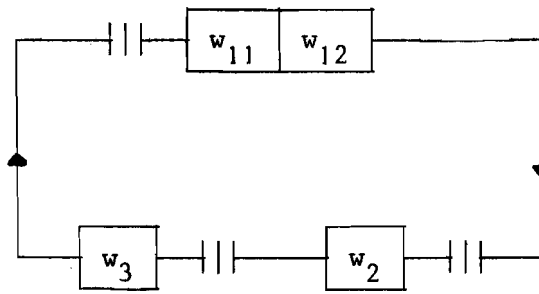
$$(12) \quad I' = \frac{1 - \Lambda_n(K)\tilde{w}_n}{\Lambda_n(K)b'_n}.$$

Consequently, once we have a set a_n , $n = 1, 2, \dots, N$, we can evaluate new values $a'_n = b'_n c'_n$. Thus we have obtained an iterative method to evaluate a sequence $(a_n^{(v)})_{v=0}^{\infty}$. The hope is that this sequence converges for every n , $n = 1, 2, \dots, N$, to a value which gives a good approximation for the original queueing network.

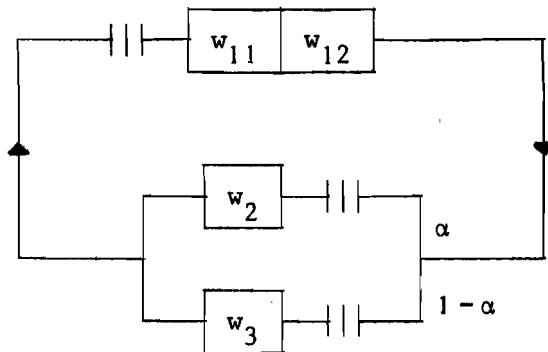
4. Numerical results

We present the iterative approximation in two illustrative situations. The exact results were obtained solving for the set of equilibrium equations.

The first system (Picture 1) is a cyclic system with a two-phase server and two exponential servers in series. The second system (Picture 2) has the two exponential servers in parallel.



Picture 1. A cyclic system with a two-phase server.



Picture 2. A cyclic system with a two-phase server.

The numerical results for the two systems are given in Tables 1 and 2. For both systems are evaluated the fractions of time customers are present ρ , the throughput Λ and the mean queue length L at the two-phase server. Furthermore the cycle time C is evaluated.

		exact results				approximations			
w_2	w_3	ρ	Λ	L	C	ρ	Λ	L	C
2	2	.594	.357	1.07	11.38	.548	.351	1.07	11.38
8	2	.155	.124	.18	32.19	.155	.124	.18	32.19
2	8	.155	.124	.18	32.19	.155	.124	.18	32.19
.25	.25	.999	.500	3.74	8.00	.999	.500	3.72	8.00

System 1. $K = 4$, $w_{11} = w_{12} = 1$

Table 1.

		exact results				approximations			
w_2	w_3	ρ	Λ	L	C	ρ	Λ	L	C
2	2	.920	.477	2.50	8.39	.901	.473	2.45	8.46
8	8	.269	.194	.36	20.57	.267	.194	.35	20.57
.25	.25	1.000	.500	3.87	8.00	1.000	.500	3.87	8.00
10	1	.364	.198	.45	20.22	.277	.199	.38	20.12
20	1	.140	.100	.17	40.03	.120	.100	.14	40.01

System 2. $K = 4$, $w_{11} = w_{12} = 1$, $\alpha = .5$

Table 2.

The convergence of the method is very fast, for all cases in the order of 5 to 10 iteration steps to obtain three decimals precision in the throughput approximation. Again, note that a converged approximation scheme does not imply that the exact result is good approximated.

We observe that the approximations are very good but for the extreme cases in the system with parallel exponential servers. This is caused by the great influence of the feedback in the system, i.e. a customer too much influences his own future. Consequently, we are convinced that especially in larger systems with more customers the approximations will be even better than in the comparatively small systems discussed above.

5. Conclusions

We have described an iterative method to obtain approximations for mean values in a queueing network with a special kind of two-phase servers. To show the development of such a method was our main purpose. It is our conviction that the iterative aspect in approximations is crucial in order to describe the influence of feedback phenomena. One-step approximations too much neglect these phenomena. Furthermore, we have shown that the method gives good results in two illustrative examples.

6. References

Reiser, M. and Kobayashi [1975]: "Queueing Networks with Multiple Closed Chains: Theory and Computational Algorithms", IBM J. of Res. and Dev. 19: 283 - 294.

Reiser, M. [1979]: "Mean Value Analysis: a new look at an old problem", 4th International Symposium on Modelling and Performance Evaluation of Computer Systems, Vienna.

Reiser, M. and S. Lavenberg [1980]: "Mean Value Analysis of Closed-Multichain Queueing Networks", JACM 22: 313 - 322.

Reiser, M. [1981]: "Mean Value Analysis and Convolution Method for Queue Dependent Servers in Closed Queueing Networks", Performance Evaluation 1: 7 - 18.