PERFORMANCE EVALUATION OF DEPENDENT TWO-STAGE SERVICES

Werner Sandmann
Department of Information Systems and Applied Computer Science
University of Bamberg
Feldkirchenstr. 21
D-96045, Bamberg, Germany
E-mail: werner.sandmann@wiai.uni-bamberg.de

KEYWORDS
Two-Stage Services, Tandem Queue, Dependent Service Times, QoS, Waiting Times, Simulation

ABSTRACT

In many scenarios services are provided in successive stages. While tandem queues appropriately reflect the structure of such scenarios, the typical assumption that service times at different stages are independent often does not fit to reality. We examine, via simulation, the impact of dependencies among service times on expected customer waiting times. The usual network simulation overhead caused by event list handling is avoided by an extension of the Lindley recursion to the two-stage case. Numerical results are presented for exponentially and uniformly distributed service times with different types of dependencies and varying server utilizations.

INTRODUCTION

Providing services in successive stages is a common feature of many service facilities. Consider for instance manufacturing operations where different steps are processed by different machines, airplane maintenance and refueling, ordering and delivery of goods or services, or supermarkets where the first stage is the self service of customers collecting their items and the second stage is the payment at the check out. In tandem queues, customers arrive at the first queueing node, successively pass in fixed order through a number of (not necessarily only two) nodes and then leave the system. Hence, tandem queues are particularly suited for modeling services that are delivered in successive stages.

The investigation of tandem queues has a long tradition starting at the latest with the work by (Reich 1957). Already important for the analysis of tandem queues, (Burke 1956) studied departures from Markovian single server queues and proved his famous theorem that the output process of an M/M/1 queue is a Poisson process with the same rate as the arrival process. Obviously, in a tandem queue the output process of the first queue constitutes the arrival process to the second queue. Based upon Burke's theorem it is shown in (Weber 1979) that in case of single server nodes with exponentially distributed service times and Poisson arrival process interchanging the nodes preserves the behavior of all nodes and the overall behavior of the original system. An overview of known results up to the late 1970s and a collection of approximation methods can be found in (Newell 1979).

Over the years tandem queues have been further studied since the provision of two-stage services in real life as well as in technical systems is still customary. Many new and emerging domains have significantly renewed current interest in tandem queues. For instance modeling and performance evaluation of the Internet has become a vital application area where it is known that despite their relative simplicity, tandem queues are suitable models for the structure of a variety of network components such as packet switches, routers and many more, see for example (Gomez-Corral 2002), (Ryoki et al 2002), (Palmowski et al 2003), (Mandjes 2004), (Buchholz 2005), (Klimenok et al 2005), (Denteneer 2006). In these recent applications, due to the requirements of Internet traffic modeling the assumption of exponentially distributed times has been dropped. However, this significantly increases the difficulty to analyze such systems, and one has to apply, e.g., matrix-geometric methods approximation and bounding techniques, or simulation. Of course, many other application areas can be considered, too, and we will not be focused on Internet services.

In any case, be it Internet services or any other type of service, providing Quality of Service (QoS) is not possible without assuring some desired level of performance and QoS guarantees should rely on proper studies of the performance. Thus, one should aim at important insights to various performance aspects, in particular delays or waiting times. The Markovian assumption, that is exponentially distributed interarrival and service times, may be also questionnable but seems to be better justified in most applications other than the Internet. In particular, exponentially distributed interarrival times consti-

Proceedings 21st European Conference on Modelling and Simulation
Ivan Zelinka, Zuzana Oplatková, Alessandra Orsoni ©ECMS 2007
ISBN 978-0-9553018-2-7 / ISBN 978-0-9553018-3-4 (CD)

tuting Poisson processes well reflect many real-world arrival processes.

The aforementionned classical works mostly considered Markovian tandem queues where service times of any customer at the successive stages are independent. Unfortunately, this limits the usefulness of the results not only for Internet applications. It seems to be evident that the assumption of independent service times at the stages does not appropriately reflect reality. For example an airplane that requires long maintenance time probably tends to require long refueling time since typically both times depend on the airplane size. Similarly, a supermarket customer who needs a long time to collect his items probably collects quite a lot of items and thus requires a longer time at the check out than a customer with fewer items who needs also less time for item collection.

Only relatively few studies of tandem queues with dependent service times were reported. (Wolff 1982) investigates light traffic asymptotics for expected waiting times in tandem queues with r stages, that means the system is considered for server utilizations approaching zero. However, though it is stated in Wolff's paper that light traffic results are important in some settings, we believe that they are of limited interest for most practical situations. (Pinedo and Wolff 1982) consider two node tandem queues where customers have equal service times at both nodes but detailed results are only provided for Markovian queues. Thus, the case of dependent service times deserves further investigation, in particular for more general dependencies than just equal service times and for non-exponential service time distributions.

In this paper we consider two-node tandem queues with Poissonian arrivals to the systems where the service times of customers at the two stages are dependent. The service times at the first stage is either exponentially or uniformly distributed. In addition to equal service times we examine cases where the service time at the second stage is a deterministic linear function (increasing or decreasing) of the service time at the first stage, which naturally extends the case of equal service times where this function is the identity. Furthermore, we introduce some kind of "Gaussian noise" affecting the service time at the second stage, that is after a deterministic function has been applied a normally distributed time is added.

Since dependencies among service times do not affect the behavior of the first stage that thus can still be treated like a standard single server queue, we are interested in the impact of the dependencies on the waiting time at the second stage. For this purpose, simulation studies are performed for the described types of dependencies, and waiting times are compared to that of the independent case where the complete range of possible (stability providing) server utilizations is considered, thus neither restricting to light traffic nor to heavy traffic. To avoid significant overhead due to event list handling that is usually present in queueing network simulations we make use of an extension of the Lindley recursion to the two-node tandem queue. Our results show that the effects of dependencies are substantially different for the chosen types of dependencies and varying utilizations.

In the remainder of this paper we first introduce our terminology and notation providing a formal model specification and we present the waiting time recursion. Then, the simulation methodology and the experimental settings are described, followed by numerical results. Finally, we conclude the paper and give directions for further research.

MODEL SPECIFICATION

In a two-node tandem queue, as depicted in Fig. 1, customers arrive at the first queueing node, receive service there and then proceed to the second node and leave the system after receiving service at that node. No arrivals from outside occur at the second node, and no customers leave the system before passing successively through both nodes. In the following we introduce the notation used throughout the paper and present the recursion that has been applied for simulating waiting times. In short-hand notation similar to the Kendall notation tandem queues are described by $\mathrm{A/B}_1/c_1 \rightarrow /\mathrm{B}_2/c_2$ where A denotes the arrival process to the system and B_1, B_2, c_1, c_2 the service time distributions and the queue capacities. Hence, for the systems that we consider we have $\mathrm{M/M/1} \rightarrow /\mathrm{M/1}$ and $\mathrm{M/U/1} \rightarrow /\mathrm{U/1}$ but with dependencies among the service times at the first and second stage not covered in this notation.

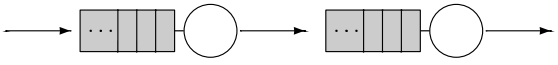

Fig. 1. Two-Node Tandem Queue

The interarrival time between the arrivals of the n-th and the $(n+1)$-th customer is denoted by T_n, and $S_{n,1}$, $S_{n,2}$ are the service times of the n-th customer at stage 1 and 2, respectively. $W_{n,1}$ and $W_{n,2}$ denote the waiting times, that means the times from the arrival of the n-th customer till service starts at stage 1 and at stage 2. The pure waiting (queueing) times of the n-th customer in the queues at stages 1 and 2 are denoted by $Q_{n,1}$ and $Q_{n,2}$. Hence,

$$Q_{n,1} = W_{n,1} \quad \text{and} \quad Q_{n,2} = W_{n,2} - S_{n,1} - W_{n,1}.$$

The according steady-state properties are denoted by $T, S_1, S_2, W_1, W_2, Q_1, Q_2$ and $\lambda = 1/E[T]$, $E[S_1]$, $E[S_2]$ are the mean (expected) steady state arrival rate and service times, respectively. The server utilizations are $\rho_i = \lambda E[S_1]$, $\rho_2 = \lambda E[S_2]$ and the system is stable iff $\rho_1, \rho_2 < 1$.

WAITING TIME RECURSION

The famous Lindley recursion (Lindley 1957) for the successive waiting times of customers in a single server queue is given by

$$W_{n+1} = \max(W_n + S_n - T_n,\ 0),$$

where of course superscripts denoting the stage are redundant and thus dropped.

To intuitively understand how this can be extended to tandem queues bear the meaning of the two components of the maximum in mind. Actually it is the differentiation of two cases as explained in what follows. If the $(n+1)$-th customer must wait then the preceding customer has not yet completed his service. Thus, the waiting time of the $(n+1)$-th is simply the sum of the waiting and service time of the preceding customer minus the time the $(n+1)$-th customer arrived later. In case of an empty system upon arrival of the $(n+1)$-th customer this becomes negative. However, arriving to an empty system means no wait and thus in this case $W_{n+1} = 0$.

With much the same reasoning an extension to waiting times in tandem queues can be obtained. Obviously, $W_{n,1}$ is exactly the waiting time in a single server queue and thus the Lindley recursion directly applies to $W_{n+1,1}$. Now, consider how long it takes until the $(n+1)$-th customer starts his service at stage 2, again differentiating two cases. If the customer must wait, then the preceding customer has not yet completed service at stage 2, and the waiting time of the $(n+1)$-th customer is again given by the sum of the waiting time and the service time of the n-th customer minus the time the (n+1)-th customer arrived later at the system. In case that the $(n+1)$-th customer need not wait at stage 2 his overall waiting time until service starts at stage 2 is given by his sojourn time in stage 1, that is the sum of the waiting time and the service time. Altogether,

$$\begin{aligned}
W_{n+1,2} &= \max\left(W_{n,2} + S_{n,2} - T_n,\ W_{n+1,1} + S_{n+1,1}\right) \\
&= \max\Big(\ W_{n,2} + S_{n,2} - T_n, \\
&\qquad \max\left(W_{n,1} + S_{n,1} - T_n,\ 0\right) + S_{n+1,1}\Big) \\
&= \max\Big(\ W_{n,2} + S_{n,2} - T_n, \\
&\qquad \max\left(W_{n,1} + S_{n,1} - T_n + S_{n+1,1}\right),\ S_{n+1,1}\Big) \\
&= \max\Big(\ W_{n,2} + S_{n,2} - T_n, \\
&\qquad\quad W_{n,1} + S_{n,1} - T_n + S_{n+1,1},\ S_{n+1,1}\Big).
\end{aligned}$$

Then the waiting time at the second stage can be obtained by $Q_{n+1,2} = W_{n+1,2} - S_{n+1,1} - W_{n+1,1}$.

SIMULATION METHODOLOGY

The extended Lindley recursion allows us to avoid any event list handling. Instead, the waiting times of the customers are simulated by successive random variate generation of the involved random times $T_1, T_2, \ldots, S_{1,1}, S_{1,2}, S_{2,1}, S_{2,2}, \ldots$ and computing the according waiting times recursively. As stated before, we are interested in the mean waiting times at the second stage. These may be estimated via the unbiased sample mean of $Q_{1,2}, Q_{2,2}, \ldots, Q_{N+M,2}$ from one single simulation of length $N+M$ where the first M waiting times have to be discarded to account for the initialization bias. However, that way no information about the accuracy or robustness of the results is available since variance estimation via the sample mean requires independent and identically distributed realizations.

Therefore, we applied the classical replication/deletion approach as described for example in (Law and Kelton 2000), that is independent simulation runs with a sufficiently large warm-up period to determine point estimates and confidence intervals. More specifically, we performed independent runs, where the observation period for each run was of length (= number of served customers) 10^7 to form 99% confidence intervals with a relative half width less than 1%. All simulations have been implemented in C++, and to omit exhausting a random number generator's cycle length we did not use the standard C++ random number generator but the one described in (L'Ecuyer et al 2002).

EXPERIMENTAL SETTINGS

We study tandem queues with Poissonian arrival processes with arrival rate $\lambda > 0$ and either exponentially or uniformly distributed service times at the first stage. Hence, the first queue is either an M/M/1 queue or an M/U/1 queue. Accordingly, for the independent cases the systems M/M/1$\rightarrow$ /M/1 and M/U/1$\rightarrow$ /U/1 are considered and compared to the dependent cases. Note that for M/M/1$\rightarrow$ /M/1 due to Burke's theorem the second stage also behaves like a single server M/M/1 queue and thus the results for this system need not be simulated but can be computed exactly. The mean service time $E[S_1]$ has been fixed to one, which means that for M/M/1 queues service times are exponentially distributed with rate $\mu_1 = 1$ and for M/U/1 queues service times are uniformly distributed on the interval $(0, 2)$.

Two basic types of dependency, as outlined in the introduction, are deterministic functions and Gaussian noise. Applying three simple linear deterministic functions we set the service time of any customer at the second stage as equal to the service time at the first stage, half of the service time at the first stage and twice the service time at the first stage.

Hence, formally expressed for the customer service times and the mean service times

$$S_{n,2} := S_{n,1} \quad \Rightarrow \quad E[S_2] = E[S_1] = 1,$$

$$S_{n,2} := \frac{1}{2} S_{n,1} \quad \Rightarrow \quad E[S_2] = \frac{1}{2} E[S_1] = \frac{1}{2},$$

$$S_{n,2} := 2S_{n,1} \quad \Rightarrow \quad E[S_2] = 2E[S_1] = 2.$$

To let the comparison of dependent and independent service times make sense all of the described dependent cases are of course compared to the independent ones with the same mean service time, that is for exponentially distributed service times with rates $1, 2, \frac{1}{2}$ and with service times uniformly distributed on the intervals $(0,2), (0,1)$ and $(0,4)$, respectively. Additionally accounting for the second type of dependency, we apply in the first step the above deterministic functions and then add a normally distributed time. More specifically, the standard normal distribution, appropriately truncated at both tails such that no negative service times are obtained and the mean service time has the required value. This assures that the mean of the truncated normal distribution still equals zero and the added distribution remains symmetric.

NUMERICAL RESULTS

Tables 1–6 show the obtained numerical results for a wide range of possible values for the arrival rate λ in the different described settings. All tables contain the estimated mean waiting times at the second stage denoted by $\overline{Q}^{(1)}$ for the independent case, $\overline{Q}^{(2)}$ for dependence due to a deterministic function and $\overline{Q}^{(3)}$ for dependence due to a deterministic function with additional Gaussian noise. Besides, the ratios of the mean waiting times for dependent service times to the mean waiting times for independent service times are given where a ratio less than one means that dependence decreases the mean waiting time and a ratio greater than one means that dependence increases the mean waiting time. Note that because of $E[S_1] = 1$ the utilization of the first server equals the arrival rate, that is $\rho_1 = \lambda$. For $E[S_2] = E[S_1] = 1$ this also holds for the utilization of the second server, that is $\rho_2 = \lambda$. For $E[S_2] = E[S_1]/2 = 1/2$ and $E[S_2] = 2E[S_1] = 2$ the utilization of the second server is given by $\rho_2 = \lambda/2$ and $\rho_2 = 2\lambda$, respectively. Consequently, to meet the stability conditions $\rho_1, \rho_2 < 1$ for $\rho_2 = \lambda/2 = \rho_1/2$ only $\rho_2 < 0.5$ is considered and for $\rho_2 = 2\lambda = 2\rho_1$ only $\rho_1 = \lambda = 2\rho_1 < 0.5$ is considered. Thus, giving the value of λ in lieu of ρ_1 and/or ρ_2 is appropriate and sufficient.

The results indicate that the impact of dependencies is not so obvious as to provide a simple general statement of the kind that dependencies increase mean waiting times in relatively light traffic and decreases mean waiting times in relatively heavy traffic.

Table 1: M/M/1 → /M/1 with $E[S_2] = E[S_1]$

λ	$\overline{Q}^{(1)}$	$\overline{Q}^{(2)}$	$\frac{\overline{Q}^{(2)}}{\overline{Q}^{(1)}}$	$\overline{Q}^{(3)}$	$\frac{\overline{Q}^{(3)}}{\overline{Q}^{(1)}}$
0.10	0.111	0.182	1.640	0.205	1.847
0.20	0.250	0.381	1.522	0.433	1.732
0.30	0.429	0.600	1.400	0.689	1.606
0.40	0.667	0.848	1.271	0.985	1.477
0.50	1.000	1.135	1.135	1.344	1.344
0.60	1.500	1.481	0.987	1.805	1.203
0.70	2.333	1.920	0.823	2.454	1.052
0.80	4.000	2.537	0.634	3.551	0.888
0.90	9.000	3.605	0.401	6.342	0.705
0.95	19.000	4.714	0.248	11.463	0.603

Table 2: M/M/1 → /M/1 with $E[S_2] = \frac{1}{2}E[S_1]$

λ	$\overline{Q}^{(1)}$	$\overline{Q}^{(2)}$	$\frac{\overline{Q}^{(2)}}{\overline{Q}^{(1)}}$	$\overline{Q}^{(3)}$	$\frac{\overline{Q}^{(3)}}{\overline{Q}^{(1)}}$
0.10	0.026	0.050	1.901	0.065	2.472
0.20	0.055	0.097	1.748	0.126	2.270
0.30	0.088	0.138	1.568	0.183	2.080
0.40	0.125	0.175	1.400	0.236	1.888
0.50	0.167	0.206	1.234	0.284	1.701
0.60	0.214	0.233	1.089	0.328	1.533
0.70	0.269	0.256	0.952	0.366	1.361
0.80	0.333	0.275	0.826	0.400	1.200
0.90	0.409	0.290	0.709	0.429	1.049
0.95	0.452	0.297	0.657	0.442	0.978

Table 3: M/M/1 → /M/1 with $E[S_2] = 2E[S_1]$

λ	$\overline{Q}^{(1)}$	$\overline{Q}^{(2)}$	$\frac{\overline{Q}^{(2)}}{\overline{Q}^{(1)}}$	$\overline{Q}^{(3)}$	$\frac{\overline{Q}^{(3)}}{\overline{Q}^{(1)}}$
0.05	0.222	0.311	1.401	0.327	1.473
0.10	0.500	0.679	1.358	0.714	1.428
0.15	0.857	1.126	1.314	1.186	1.384
0.20	1.333	1.691	1.269	1.785	1.339
0.25	2.000	2.445	1.223	2.585	1.293
0.30	3.000	3.529	1.176	3.741	1.247
0.35	4.667	5.275	1.130	5.608	1.202
0.40	8.000	8.678	1.085	9.256	1.157
0.45	18.000	18.715	1.040	20.040	1.113

Table 4: M/U/1 → /U/1 with $E[S_2] = E[S_1]$

λ	$\overline{Q}^{(1)}$	$\overline{Q}^{(2)}$	$\frac{\overline{Q}^{(2)}}{\overline{Q}^{(1)}}$	$\overline{Q}^{(3)}$	$\frac{\overline{Q}^{(3)}}{\overline{Q}^{(1)}}$
0.10	0.054	0.067	1.248	0.101	1.857
0.20	0.120	0.138	1.152	0.214	1.782
0.30	0.202	0.212	1.051	0.347	1.717
0.40	0.309	0.290	0.941	0.509	1.647
0.50	0.456	0.375	0.821	0.716	1.568
0.60	0.676	0.465	0.688	1.004	1.485
0.70	1.042	0.566	0.543	1.454	1.396
0.80	1.777	0.681	0.383	2.307	1.298
0.90	4.003	0.817	0.204	4.771	1.192
0.95	8.477	0.900	0.106	9.609	1.134

Table 5: M/U/1 → /U/1 with $E[S_2] = \frac{1}{2}E[S_1]$

λ	$\overline{Q}^{(1)}$	$\overline{Q}^{(2)}$	$\frac{\overline{Q}^{(2)}}{\overline{Q}^{(1)}}$	$\overline{Q}^{(3)}$	$\frac{\overline{Q}^{(3)}}{\overline{Q}^{(1)}}$
0.10	0.011	0.014	1.318	0.025	2.336
0.20	0.022	0.027	1.252	0.050	2.289
0.30	0.034	0.040	1.179	0.073	2.173
0.40	0.046	0.051	1.106	0.096	2.074
0.50	0.060	0.062	1.032	0.117	1.963
0.60	0.074	0.071	0.962	0.137	1.851
0.70	0.090	0.080	0.892	0.155	1.726
0.80	0.107	0.088	0.822	0.173	1.612
0.90	0.127	0.096	0.753	0.189	1.489
0.95	0.138	0.099	0.717	0.196	1.420

Table 6: M/U/1 → /U/1 with $E[S_2] = 2E[S_1]$

λ	$\overline{Q}^{(1)}$	$\overline{Q}^{(2)}$	$\frac{\overline{Q}^{(2)}}{\overline{Q}^{(1)}}$	$\overline{Q}^{(3)}$	$\frac{\overline{Q}^{(3)}}{\overline{Q}^{(1)}}$
0.05	0.134	0.161	1.197	0.182	1.354
0.10	0.303	0.355	1.171	0.401	1.324
0.15	0.521	0.596	1.145	0.676	1.299
0.20	0.814	0.912	1.121	1.035	1.272
0.25	1.229	1.347	1.094	1.533	1.248
0.30	1.859	1.996	1.074	2.274	1.223
0.35	2.926	3.079	1.052	3.510	1.200
0.40	5.091	5.259	1.033	5.996	1.178
0.45	11.673	11.858	1.016	13.506	1.157

Such relations could be suggested by results given in (Pinedo and Wolff 1982). The only common features that we observe in all experiments is that the ratio of the mean waiting time for dependent service times to the mean waiting time for independent service times decreases with increasing utilization, and that this ratio is greater for service times that are subject to an additional Gaussian noise than for service times without that noise. Taking a look at Table 1 where the setting of equal service times in M/M/1 → /M/1 queues considered in (Pinedo and Wolff 1982) is included, we indeed obtain results consistent with that in (Pinedo and Wolff 1982). In particular, between utilizations of 0.50 and 0.60 the mean waiting times for equal service times become smaller than for independent service times. By further simulation experiments the cross-over point has been determined to be approximately 0.58.

However, Table 1 also shows that in the same setting with additional Gaussian noise this cross-over point is reached only between utilizations of 0.70 and 0.80. Moreover, as presented in Table 2, the cross-over point for $S_1 = S_2/2$ is between 0.60 and 0.70 and with additional Gaussian noise between 0.90 and 0.95. Table 3 shows similar effects where now in all dependent settings the mean waiting times are greater than for independent service times. Note again that the given values of λ correspond to utilizations $\rho_2 \in \{0.1, \ldots, 0.9\}$.

Tables 4–6 show similar effects for M/U/1 → /U/1 but there are also significant differences to M/M/1 → /M/1. It is remarkable that the ratio of the mean waiting time for equal service times to the mean waiting time for independent service times here for all utilizations is smaller than the corresponding ratio in the M/M/1 → /M/1 setting. However, this does not hold for the other types of dependency.

Given that the impact of dependencies is not obvious even for relatively simple types of dependency, we claim that more complicated dependencies will possess even significantly more complicated impacts on the mean waiting time. In the light of complex dependencies that may arise in real-world settings this should motivate further investigations of models with dependent service times.

CONCLUSION

We investigated tandem queues where service times of customers at successive services stages are dependent. Simulation results obtained via an extension of the Lindley recursion have been presented for M/M/1 → /M/1 and M/U/1 → /U/1 with different types of dependency. The results show that dependencies significantly affect the mean waiting time at the second stage. Compared to the independent case, in some settings the mean waiting time is increased

only for moderately light traffic and decreased for heavier traffic. In other settings the mean waiting time is always increased. Thus, a simple general relation between mean waiting times for independent and dependent service times seems not to exist and further investigations are desirable to better understand the impact of dependencies on system performance and QoS.

Further research aims at more simulation results as well as analytical and numerical results. Besides more complex types of dependency, many more interarrival and service time distributions should be considered. With regard to Internet services as special applications, heavy-tailed distributions must be incorporated. Studying multiple service stages, that is a series of more than two queues, is also desirable. In particular, the Lindley recursion can be further extended to multiple stages and utilized for simulations similarly to the two-stage case.

Concerning the performance properties, it is also of interest to estimate the waiting time distribution or to study variances, higher moments and quantiles of the waiting time and probabilities of extreme or rare events such as the probability of excessive waiting times. The estimation of the latter is very difficult even in the independent case and can be expected to be more complicated in the dependent case. Hence, despite their relative simplicity, tandem queues still pose various great challenges for research, all of which are of high practical relevance.

REFERENCES

Buchholz, P. 2006. "Bounding Stationary Results of Tandem Networks with MAP Input and PH Service Time Distributions". In *Proceedings of SIGMETRICS/Performance'06*, 191–202.

Burke, P. J. 1956. "The Output of a Queueing System". *Operations Research* 4, 699–704.

Denteneer, D. 2006. "Models for Data Transmission Delay in Cable Networks". *Statistica Neerlandica* 60, No. 1, 12–45.

Gomez-Corral, A. 2002. "A Tandem Queue with Blocking and Markovian Arrival Process". *Queueing Systems* 41, 343–370.

Klimenok, V. I.; L. Breuer; G. V. Tsarenkov; and A. Dudin. 2005. "The BMAP/G/1/N→PH/1/M System with Losses". *Performance Evaluation* 61, 17–40.

Law, A. M. and W. D. Kelton. 2000. *Simulation Modeling and Analysis*. 3rd ed., McGraw Hill.

L'Ecuyer, P.; R. Simard; E. J. Chen; and W. D. Kelton. 2002. "An Object-Oriented Random-Number Package with Many Long Streams and Substreams". *Operations Research* 50, No. 6, 1073–1075.

Lindley, D. V. 1952. "The Theory of Queues with a Single Server". *Proceedings of the Cambridge Philosophical Society* 48, 277–289.

Mandjes, M. 2004. "Packet Models Revisited: Tandem and Priority Systems". *Queueing Systems* 47, No. 4, 363–377.

Newell, G. F. 1979. *Approximate Behavior of Tandem Queues. Lecture Notes in Economics and Mathematical Systems* 171, Springer.

Palmowski, Z.; S. Schlegel; and O. Boxma. 2003. "A Tandem Queue with Gate Mechanism". *Queueing Systems* 43, No. 4, 349–364.

Pinedo, M. and R. W. Wolff. 1982. "A Comparison between Tandem Queues with Dependent and Independent Service Times". *Operations Research* 30, No. 3, 464–479.

Reich, E. 1957. "Waiting Times when Queues are in Tandem". *The Annals of Mathematical Statistics* 28, No. 3, 768–773.

Ryoki, N.; K. Kawhara; T. Ikenaga; and Y. Oie. 2002. "Performance Analysis of Queue Length Distributions of Tandem Routers for QoS Measurement". In *Proceedings of the Symposium on Applications and the Internet (SAINT)*, 82–87.

Weber, R. 1979. "The Interchangeability of /M/1 Queues in Series". *Journal of Applied Probability* 16, 690–695.

Wolff, R. W. 1982. "Tandem Queues with Dependent Service Times in Light Traffic". *Operations Research* 30, No. 4, 619–635.

AUTHOR BIOGRAPHY

WERNER SANDMANN studied Computer Science and Mathematics at the University of Bonn, Germany, where he received his diploma degree in Computer Science (Dipl.–Inform.) and his doctoral degree (Dr. rer. nat.) in 1998 and 2004, respectively. From 1998–2003 he was a Research and Teaching Assistant at the Computer Science Department of the University of Bonn. Since 2004 he is an Assistant Professor of Computer Science at the University of Bamberg, Germany. His research interests are in applied probability and stochastic modeling, including computer systems performance evaluation, reliability, quality of services, computational biology, analytical and numerical solution techniques, and simulation. His email address is `werner.sandmann@wiai.uni-bamberg.de`.