

Preface

Networks of queues (= stochastic networks) have been a field of intensive research over the last three decades. The foundation for this research is classical queueing theory, which mainly dealt with single node queueing systems [21], [24], [49]. There is now a well developed theory of stochastic networks accompanied by an unbounded set of open problems which originates directly from applications as well as from theoretical considerations. Most of these open problems are easily stated and put into a theoretical framework, but often require either intricate techniques on an adhoc basis or deep mathematical methods. Often both of these approaches have to be combined to successfully tackle the solution of a quickly formulated problem.

The possibly most influential sources were the books by Kelly [75], Whittle [139], and Walrand [136]. The recent books by Van Dijk [40], Serfozo [121], and Chao, Miyazawa, Pinedo [18] gave a good state-of-the-art survey of the subject.

The development of queueing network theory which provided application areas with solutions, formulas, and algorithms commenced around 1950 in the area of Operations Research, with special emphasis on production, inventory, and transportation. The first breakthrough came with the works of Jackson [69] and Gordon and Newell [50]. Both papers appeared in OPERATIONS RESEARCH. The second breakthrough in queueing network theory was already connected with Computer Science: the celebrated papers by Baskett, Chandy, Muntz, and Palacios [4], which appeared in the JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY, and by Kelly [74].

At the same time the two volumes of Kleinrock's book appeared [81], [82]. From that time on queueing network theory and its application has been intimately connected with performance analysis of complex systems in Computer and Communication Sciences, both on the hardware and on the software level.

The way that production, manufacturing, and transportation is growing together with information processing and communication technology is resulting in more and more complex systems which require even more elaborate models, techniques, and algorithms to better understand their performance behavior and to predict performance and quality of service.

The classical single node queues and their networks are models living on a continuous time scale. Even if for some applications a discrete time scale might be more appropriate, the well established continuous time machinery often serves as an approximation tool. Consequently, the survey articles [24], [137] from 1990 still do not review discrete time models.

However, from then on an astonishing evolution of discrete time stochastic network models can be observed. Usually it is thought that the invention of ATM (Asynchronous Transfer Mode) as the protocol for high speed transmission network technology triggered this development.

At least three (to my knowledge) books appeared recently solely dedicated to theory and applications of discrete time queueing systems and networks, [14], [128], and [142]. The contents of this lecture notes are spread over the same area, but are almost complementary to the above mentioned books. The center of my interest is the question, whether there is in discrete time an analogue to the celebrated *product form calculus* of continuous time stochastic network theory. This calculus lies behind the two breakthroughs mentioned above. Its simplicity and general applicability was its entrance ticket into applications. Its theoretical elegance opened the door to mathematical investigations around stochastic networks.

From a mathematical point of view, for the description of stochastic networks, we deal with stochastic processes in continuous time as well as in discrete time. The theory of stochastic processes in continuous time is much more elaborate than that in discrete time. However, we shall see that we have to pay for turning to the simpler theoretical framework by being burdened with a much more technically involved machinery for arriving at simple methods for performance analysis and simple algorithms. The question of whether there can be a product form calculus in discrete time will be answered moderately positively in the book, but in my opinion the answer is still in a premature status. There seems to be much need for further research and there *are disappointing negative partial answers* in cases where the continuous time counterparts of the question have an easy positive answer. This will be discussed at several instances throughout the presentation.

The text is centered around explicit expressions for the steady state behavior of discrete time queueing networks. Three classes of networks which show a product form equilibrium will be discussed in detail:

(1) Linear networks (closed cycles and open tandems) of single server FCFS Bernoulli nodes.

(2) Networks of doubly stochastic and geometrical queues (which are discrete time analogues of Kelly's symmetric and general servers and of the BCMP nodes):

Customers of different types move through the network governed by a general routing mechanism. Request for service at symmetric servers may be according to general distributions.

(3) Networks with batch movements of customers and batch service. The service and routing mechanism is defined on the basis of an abstract transition scheme.

I further discuss computational algorithms for the standard performance measures and present explicit results on end-to-end-delay distributions and their moments.

The mathematical prerequisites for reading the book are moderate: discrete time Markov chains and their elementary linear algebra, and elementary trans-

form methods (generating functions, z-transform) as can be found in [70], chapter 2, 3, and [71], chapter 10, 11, or [44], chapter 15, or [66], [67]. Knowledge of elementary continuous time queueing networks, [75], [47] would be helpful to understand their discrete time counterparts, but is not necessary.

I tried to compile the relevant literature with respect to discrete time stochastic networks. But, unfortunately enough, due to the enormous increase of publication rates in the field I am quite sure that I have missed some sources. However, I did not try to repeat the compilation in [18] on batch movement networks (with e.g. signals and further specific properties) where in general it is proposed, that the derivations and results hold for discrete time and continuous time systems equally. A number of models, which in my opinion are prototype models, are surveyed in chapter 6 with reference to further readings.

These lecture notes are based on the lecture notes of a one-semester course on “Discrete Time Stochastic Networks” held at the Department of Mathematics at Hamburg University in summer 1997 for students on a graduate level. The audience consisted of members of the Departments of Mathematics and Computer Sciences and of the Department of Communication Networks of the Technical University of Hamburg–Harburg. A preliminary version of these lecture notes was used as material and distributed within a Tutorial on “Discrete Time Queueing Networks: Recent Developments”, which I held at the conference PERFORMANCE '96 in Lausanne (Switzerland) in October 1996.

The text is a report on some parts of the recent research and development in stochastic network theory and some of its applications. It may also serve as a textbook for a course on discrete time stochastic networks or as a companion book in a course on (elementary) stochastic processes.

Acknowledgements. I would like to thank students and colleagues who participated in the lectures on “Discrete Time Stochastic Networks” in summer 1997 for their many valuable comments and remarks on the problems I presented to them. I thank Kristin Betancourt for reading parts of the manuscript. I am grateful to Regina, Mirok Korea, and their friends for their encouragement and continuous support during the time I was writing these lecture notes.