Editorial introduction

Venkat Anantharam

School of Electrical Engineering, 313 Phillips Hall, Cornell University, Ithaca, NY 14853-3801, USA

Ralph L. Disney

Department of Industrial Engineering, 337 Zachry Engineering Center, Texas A&M University, College Station, TX 77843-3131, USA

Jean Walrand

Department of Electrical Engineering and Computer Science, University of California – Berkeley, Berkeley, CA 94720, USA

Received 21 December 1992

Queueing network theory has a long history, starting with the early attempts to model telephone networks. The subject has grown into a rich area being fed both by new applications and by the use of increasingly sophisticated tools from the theory of probability and stochastic processes. Thus one finds applications to diverse technologies such as packet switched and circuit switched communication networks, computer networks, personal communication networks, job shops, flexible manufacturing systems, and to the timing analysis of distributed software systems. The mathematical tools being used include general point process and marked point process theory, Markov decision theory, martingales, renewal and regenerative processes, diffusion processes, perturbation analysis, and techniques from the theory of (max, +) algebras. This special issue of QUESTA was developed with the aim of demonstrating the vitality of the field to the extent possible within the framework of survey articles by some of the prominent researchers. We hope it will have the effect of catalyzing further developments in the area.

The papers of Harrison and Nguyen and of Dai and Wang deal with the area of Brownian models of queueing networks. The former paper is a survey of the area. It discusses how one can write down a Brownian approximation to queueing networks in heavy traffic based on heuristic arguments, and justifies these arguments in the case of single class networks and multiclass networks with feedforward routing. It focuses on the analysis of the distribution of sojourn times using Brownian approximations and offers a couple of numerical examples to illustrate how this approach can be used for performance analysis in practice. The latter paper is devoted to a counterexample to the natural conjecture that the Brownian approximations of Harrison and Nguyen are correct in general. Multiclass queueing networks are demonstrated where the Brownian models in fact do not exist. This paper draws on the work of Reiman and Williams on conditions for the existence of semimartingale reflected Brownian motions. The existence of such counterexamples is a source of considerable interest and work on understanding what is going on is currently in full swing.

Kelly and Laws discuss several problems on routing and sequencing in open queueing networks, focusing especially on simplifications that result in the heavy traffic limit. Through an analysis of several examples of increasing intricacy, they bring out the role of the topology of the underlying network in the kinds of aggregates that take place in the heavy traffic limit. In particular they demonstrate the importance of a class of generalized cut constraints. This paper then studies the performance of some natural control strategies in the heavy traffic limit in networks where one of the generalized cut constraints is close to being tight. The formal justification of this approach via weak convergence theorems is posed as an important challenge.

Kumar introduces to queueing network theory a new class of networks encountered in the important areas of computer chip manufacturing and thin film processing. Though more classical systems models have allowed an occasional node to which a part may return (e.g. the retrial problem in telephony or the feedback problem in industrial rework) such re-entrant occurrences are probabilistically determined by some aspect of the network (e.g. a full queue causes an arrival to retry or a defective part causes a part to be reworked). In the Kumar model these re-entrants are a designed part of the manufacturing system; the systems are cyclic. A basic question in the area is how best to schedule parts at various stages of completion through a given production center.

The paper starts with a standard BCMP model and asks if the system's performance can be optimized by a choice of a schedule other than the FCFS inherent in the BCMP model. Several alternatives are suggested, analyzed, and simulated in an attempt to determine an optimal schedule depending on what system's performance is to be optimized. The sections that follow include a stochastic control approach for machine failures and set up times in re-entrant problems. The author presents several intriguing open problems scattered throughout.

In his paper Marbukh studies a communication network with a finite number of nodes and multiple, two way trunks joining pairs of nodes. Incoming traffic on a trunk is Poisson, but its parameter may depend on the origin–destination pair. Call lengths are exponentially distributed. If all routes joining the origin–destination nodes of an incoming call are occupied, the call is lost. The Kolmogorov differential equations can be set up to define the time dependent state probabilities for the finite numbers of states, but as usual a direct attack on these, even for moderately sized problems, does not appear feasible. To get to a solution, the simplifying assumption that occupancy on the branches of the network graph are independent random variables is made. This approach leads to a much smaller set of differential equations and these with additional assumptions lead eventually to the problem to be analyzed in detail. This line of development leads to the introduction and analysis of a new state-dependent, dynamic routing strategy which is then compared to the more commonly discussed routing procedures. Of particular interest to a queueing theorist is the author's study of the bifurcation graph of some of the differential equations.

Using a structure reminiscent of the Kelly or Whittle setups, Serfozo studies first the Markov queue length process where the service intensities at a given node may depend on the vector of the number of customers at other nodes (the "dependent nodes" case of the title). Then, after establishing a relation between these intensities similar to the Kolmogorov condition for reversibility, the paper presents an equilibrium distribution for the network and demonstrates, from this, results for well known networks (e.g. Jackson, Kelly, Whittle). Following this discussion, Serfozo allows the transition intensities to depend on the state of the network as before as well as the number of arrivals and departures at the nodes. From this structure he obtains both the dependent node and concurrent movement condition of the paper's title. This more general network is explored as to its reversibility and implications of this, partially balanced networks of this type, quasi-reversibility, and weak coupling. The paper concludes with a brief discussion of some of the measures of effectiveness for the networks previously discussed.

The paper of Massey and Whitt studies a model called the Poisson arrival location model that is motivated by the emerging field of mobile cellular telecommunications. In this model, customers arrive according to a nonhomogeneous Poisson process and move independently through a general state space according to a location stochastic process after arrival. The familiar networks of infinite exponential serve queues are a special case of this model. A number of time dependent traffic characteristics can be easily calculated based on an underlying Poisson random measure representation. In particular, a time dependent product form result is also obtained.

It is known that quasi-reversible networks obey an analog of Norton's theorem of electrical circuit theory. In such a network it is possible to replace a subset of the network by a single queue such that the equilibrium distribution of the rest of the network remains unchanged. This theorem is useful in a hierarchical approach to network analysis. If a network can be considered as a global interconnection of local clusters or subnetworks and one is only interested in global characteristics, one can study these by replacing each local cluster by a single station. The paper of Boucherie and Van Dijk discusses extensions of the analog of Norton's theorem to product form networks with state depending routing. A general framework for aggregation and decomposition of such networks is provided.

Stidham and Weber survey the results achieved on problems of optimal control of networks of queues using Markov decision theory. Problems of service rate control, routing and admission control and scheduling are discussed. While a number of interesting qualitative properties of optimal controls have been identified, many open problems remain in this area. This paper should be a valuable guide to the state of the field.