

## Communication Networks

*Communication Networks* blends control, optimization, and stochastic network theories with features that support student learning to provide graduate students with an accessible, modern approach to the design of communication networks.

- Covers a broad range of performance analysis tools, including important advanced topics that are made accessible to graduate students for the first time.
- Taking a top-down approach to network protocol design, the authors begin with the deterministic model and progress to more sophisticated models.
- Network algorithms and protocols are tied closely to the theory, engaging students and helping them understand the practical engineering implications of what they have learnt.
- The background behind the mathematical analyses is given before the formal proofs and is supported by worked examples, enabling students to understand the big picture before going into the detailed theory.
- End-of-chapter exercises cover a range of difficulties; complex problems are broken down into several parts, many with hints to guide students. Full solutions are available to instructors.

**R. Srikant** is the Fredric G. and Elizabeth H. Nearing Endowed Professor of Electrical and Computer Engineering, and a Professor in the Coordinated Science Laboratory, at the University of Illinois at Urbana-Champaign, and is frequently named in the university's list of teachers ranked as excellent. His research interests include communication networks, stochastic processes, queueing theory, control theory, and game theory. He has been a Distinguished Lecturer of the IEEE Communications Society, is a Fellow of the IEEE, and is currently the Editor-in-Chief of the *IEEE/ACM Transactions on Networking*.

**Lei Ying** is an Associate Professor in the School of Electrical, Computer and Energy Engineering at Arizona State University, and former Northrop Grumman Assistant Professor at Iowa State University. He is a winner of the NSF CAREER Award and the DTRA Young Investigator Award. His research interests are broadly in the area of stochastic networks, including wireless networks, P2P networks, cloud computing, and social networks.

# Communication Networks

AN OPTIMIZATION, CONTROL, AND STOCHASTIC  
NETWORKS PERSPECTIVE

R. SRIKANT

University of Illinois at Urbana-Champaign

LEI YING





Shaftesbury Road, Cambridge CB2 8EA, United Kingdom  
One Liberty Plaza, 20th Floor, New York, NY 10006, USA  
477 Williamstown Road, Port Melbourne, VIC 3207, Australia  
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India  
103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,  
a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of  
education, learning and research at the highest international levels of excellence.

[www.cambridge.org](http://www.cambridge.org)  
Information on this title: [www.cambridge.org/9781107036055](http://www.cambridge.org/9781107036055)

© Cambridge University Press & Assessment 2014

This publication is in copyright. Subject to statutory exception and to the provisions  
of relevant collective licensing agreements, no reproduction of any part may take  
place without the written permission of Cambridge University Press & Assessment.

First published 2014

*A catalogue record for this publication is available from the British Library*

*Library of Congress Cataloging-in-Publication data*

Srikant, R. (Rayadurgam)

Communication networks : an optimization, control, and stochastic networks perspective /  
R. Srikant, University of Illinois at Urbana-Champaign, Lei Ying, Arizona State University.  
pages cm

Includes bibliographical references and index.

ISBN 978-1-107-03605-5 (Hardback)

1. Telecommunication systems. I. Ying, Lei (Telecommunication engineer) II. Title.  
TK5101.S657 2013  
384–dc23 2013028843

ISBN 978-1-107-03605-5 Hardback

Additional resources for this publication at [www.cambridge.org/srikant](http://www.cambridge.org/srikant)

Cambridge University Press & Assessment has no responsibility for the persistence  
or accuracy of URLs for external or third-party internet websites referred to in this  
publication and does not guarantee that any content on such websites is, or will  
remain, accurate or appropriate.

v

To Amma, Appa, Susie, Katie, and Jenny  
*RS*  
  
To my parents, Lingfang and Ethan  
*LY*

“This book by Srikant and Ying fills a major void – an analytical and authoritative study of communication networks that covers many of the major advances made in this area in an easy-to-understand and self-contained manner. It is a must read for any networking student, researcher, or engineer who wishes to have a fundamental understanding of the key operations of communication networks, from network dimensioning and design to congestion control, routing, and scheduling. Throughout the book, the authors have taken pains to explain highly mathematical material in a manner that is accessible to a beginning graduate student. This has often required providing new examples, results, and proofs that are simple and easy to follow, which makes the book attractive to academics and engineers alike. A must have networking book for one’s personal library!”

**Ness B. Shroff**

*The Ohio State University*

“*Communication Networks* provides a deep, modern and broad yet accessible coverage of the analysis of networks. The authors, who made many original contributions to this field, guide the readers through the intuition behind the analysis and results. The text is ideal for self-study and as a basis for a graduate course on the mathematics of communication networks. Students in networking will benefit greatly from reading this book.”

**Jean Walrand**

*University of California, Berkeley*

“*Communication Networks*, by Srikant and Ying, provides a mathematically rigorous treatment of modern communication networks. The book provides the essential mathematical preliminaries in queueing theory, optimization and control, followed by a rigorous treatment of network architectures, protocols and algorithms that are at the heart of modern-day communication networks and the Internet. It is the best textbook on communication networks from a theoretical perspective in over 20 years, filling a much needed void in the field. It can be an excellent textbook for graduate and advanced undergraduate classes, and extremely useful to researchers in this rapidly evolving field.”

**Eytan Modiano**

*Massachusetts Institute of Technology*

“This book presents a view of communication networks, their architecture and protocols, grounded in the theoretical constructs from optimization and queueing theory that underpin the modern approach to the design and analysis of networks. It is a superb introduction to this approach.”

**Frank Kelly**

*University of Cambridge*

“This textbook provides a thoughtful treatment of network architecture and network protocol design within a solid mathematical framework. Networks are required to provide good stable behavior in random environments. This textbook provides the tools needed to make this happen. It provides needed foundations in optimization, control, and probabilistic techniques. It then demonstrates their application to the understanding of current networks and the design of future network architectures and protocols. This is a ‘must’ addition to the library of graduate students performing research in networking, and engineers researching future network architectures and protocols.”

**Donald F. Towsley**

*University of Massachusetts Amherst*

CONTENTS

<i>Preface</i>	<i>page xi</i>
1 Introduction	1
<b>I Network architecture and algorithms</b>	<b>5</b>
2 Mathematics of Internet architecture	7
2.1 Mathematical background: convex optimization	7
2.1.1 Convex sets and convex functions	7
2.1.2 Convex optimization	11
2.2 Resource allocation as utility maximization	15
2.2.1 Utility functions and fairness	17
2.3 Mathematical background: stability of dynamical systems	19
2.4 Distributed algorithms: primal solution	21
2.4.1 Congestion feedback and distributed implementation	24
2.5 Distributed algorithms: dual solution	26
2.6 Feedback delay and stability	27
2.6.1 Linearization	29
2.7 Game-theoretic view of utility maximization	30
2.7.1 The Vickrey–Clarke–Groves mechanism	31
2.7.2 The price-taking assumption	34
2.7.3 Strategic or price-anticipating users	35
2.8 <i>Summary</i>	41
2.9 <i>Exercises</i>	42
2.10 <i>Notes</i>	47
3 Links: statistical multiplexing and queues	49
3.1 Mathematical background: the Chernoff bound	49
3.2 Statistical multiplexing and packet buffering	51
3.2.1 Queue overflow	52
3.3 Mathematical background: discrete-time Markov chains	55
3.4 Delay and packet loss analysis in queues	64
3.4.1 Little’s law	64
3.4.2 The Geo/Geo/1 queue	67
3.4.3 The Geo/Geo/1/B queue	69
3.4.4 The discrete-time G/G/1 queue	70

Contents

3.5	Providing priorities: fair queueing	72
3.5.1	Key properties	76
3.6	<i>Summary</i>	78
3.7	<i>Exercises</i>	79
3.8	<i>Notes</i>	85
4	<b>Scheduling in packet switches</b>	<b>86</b>
4.1	Switch architectures and crossbar switches	87
4.1.1	Head-of-line blocking and virtual output queues	88
4.2	Capacity region and MaxWeight scheduling	90
4.2.1	Intuition behind the MaxWeight algorithm	96
4.3	Low-complexity switch scheduling algorithms	96
4.3.1	Maximal matching scheduling	96
4.3.2	Pick-and-compare scheduling	102
4.3.3	Load-balanced switches	102
4.4	<i>Summary</i>	105
4.5	<i>Exercises</i>	106
4.6	<i>Notes</i>	109
5	<b>Scheduling in wireless networks</b>	<b>110</b>
5.1	Wireless communications	110
5.2	Channel-aware scheduling in cellular networks	114
5.3	The MaxWeight algorithm for the cellular downlink	116
5.4	MaxWeight scheduling for ad hoc P2P wireless networks	122
5.5	General MaxWeight algorithms	125
5.6	Q-CSMA: a distributed algorithm for ad hoc P2P networks	129
5.6.1	The idea behind Q-CSMA	129
5.6.2	Q-CSMA	130
5.7	<i>Summary</i>	134
5.8	<i>Exercises</i>	135
5.9	<i>Notes</i>	140
6	<b>Back to network utility maximization</b>	<b>142</b>
6.1	Joint formulation of the transport, network, and MAC problems	142
6.2	Stability and convergence: a cellular network example	151
6.3	Ad hoc P2P wireless networks	155
6.4	Internet versus wireless formulations: an example	157
6.5	<i>Summary</i>	159
6.6	<i>Exercises</i>	160
6.7	<i>Notes</i>	163
7	<b>Network protocols</b>	<b>165</b>
7.1	Adaptive window flow control and TCP protocols	166
7.1.1	TCP-Reno: a loss-based algorithm	167
7.1.2	TCP-Reno with feedback delay	170

Contents

7.1.3	TCP-Vegas: a delay-based algorithm	171
7.2	Routing algorithms: Dijkstra and Bellman–Ford algorithms	175
7.2.1	Dijkstra’s algorithm: link-state routing	176
7.2.2	Bellman–Ford algorithm: distance-vector routing	179
7.3	IP addressing and routing in the Internet	182
7.3.1	IP addressing	183
7.3.2	Hierarchical routing	184
7.4	MAC layer protocols in wireless networks	186
7.4.1	Proportionally fair scheduler in cellular downlink	187
7.4.2	MAC for WiFi and ad hoc networks	188
7.5	<i>Summary</i>	191
7.6	<i>Exercises</i>	192
7.7	<i>Notes</i>	194
8	Peer-to-peer networks	195
8.1	Distributed hash tables	195
8.1.1	Chord	196
8.1.2	Kademlia	202
8.2	P2P file sharing	207
8.2.1	The BitTorrent protocol	208
8.3	Structured P2P streaming	210
8.4	Unstructured P2P streaming	215
8.5	The gossip process	219
8.6	<i>Summary</i>	221
8.7	<i>Exercises</i>	222
8.8	<i>Notes</i>	225
II	Performance analysis	227
9	Queueing theory in continuous time	229
9.1	Mathematical background: continuous-time Markov chains	229
9.2	Queueing systems: introduction and definitions	237
9.3	The M/M/1 queue	239
9.4	The M/M/s/s queue	241
9.4.1	The PASTA property and blocking probability	242
9.5	The M/M/s queue	242
9.6	The M/GI/1 Queue	243
9.6.1	Mean queue length and waiting time	246
9.6.2	Different approaches taken to derive the P-K formula	247
9.7	The GI/GI/1 queue	249
9.8	Reversibility	251
9.8.1	The M/M/1 queue	253
9.8.2	The tandem M/M/1 queue	254
9.9	Queueing systems with product-form steady-state distributions	254



Contents

9.9.1	The Jackson network	255
9.9.2	The multi-class M/M/1 queue	256
9.10	Insensitivity to service-time distributions	258
9.10.1	The M/M/1-PS queue	259
9.10.2	The M/GI/1-PS queue	259
9.11	Connection-level arrivals and departures in the internet	263
9.12	Distributed admission control	267
9.13	Loss networks	269
9.13.1	Large-system limit	271
9.13.2	Computing the blocking probabilities	274
9.13.3	Alternative routing	275
9.14	Download time in BitTorrent	276
9.15	<i>Summary</i>	280
9.16	<i>Exercises</i>	282
9.17	<i>Notes</i>	289
10	Asymptotic analysis of queues	290
10.1	Heavy-traffic analysis of the discrete-time G/G/1 queue	291
10.2	Heavy-traffic optimality of JSQ	294
10.3	Large deviations of i.i.d. random variables: the Cramer–Chernoff theorem	302
10.4	Large-buffer large deviations	307
10.5	Many-sources large deviations	312
10.6	<i>Summary</i>	317
10.7	<i>Exercises</i>	318
10.8	<i>Notes</i>	321
11	Geometric random graph models of wireless networks	323
11.1	Mathematical background: the Hoeffding bound	323
11.2	Nodes arbitrarily distributed in a unit square	325
11.3	Random node placement	328
11.4	<i>Summary</i>	335
11.5	<i>Exercises</i>	336
11.6	<i>Notes</i>	339
	<i>References</i>	340
	<i>Index</i>	349

## PREFACE

### Why we wrote this book

Traditionally, analytical techniques for communication networks discussed in textbooks fall into two categories: (i) analysis of network protocols, primarily using queueing theoretic tools, and (ii) algorithms for network provisioning which use tools from optimization theory. Since the mid 1990s, a new viewpoint of the architecture of a communication network has emerged. Network architecture and algorithms are now viewed as slow-time-scale, distributed solutions to a large-scale optimization problem. This approach illustrates that the layered architecture of a communication network is a natural by-product of the desire to design a fair and stable system. On the other hand, queueing theory, stochastic processes, and combinatorics play an important role in designing low-complexity and distributed algorithms that are viewed as operating at fast time scales.

Our goal in writing this book is to present this modern point of view of network protocol design and analysis to a wide audience. The book provides readers with a comprehensive view of the design of communication networks using a combination of tools from optimization theory, control theory, and stochastic networks, and introduces mathematical tools needed to analyze the performance of communication network protocols.

### Organization of the book

The book has been organized into two major parts. In the first part of the book, with a few exceptions, we present mathematical techniques only as tools to design algorithms implemented at various layers of a communication network. We start with the transport layer, and then consider algorithms at the link layer and the medium access layer, and finally present a unified view of all these layers along with the network layer. After we cover all the layers, we present a brief introduction to peer-to-peer applications which, by some estimates, form a significant portion of Internet traffic today.

The second part of the book is devoted to advanced mathematical techniques which are used frequently by researchers in the area of communication networks. We often sacrifice generality by making simplifying assumptions, but, as a result, we hope that we have made techniques that are typically found in specialized texts in mathematics more broadly accessible. The collection of mathematical techniques relevant to communication networks is vast, so we have perhaps made a personal choice in the selection of the topics. We have chosen to highlight topics in basic queueing theory, asymptotic analysis of queues, and scaling laws for wireless networks in the second part of the book.

We note that two aspects of the book are perhaps unique compared to other textbooks in the field: (i) the presentation of the mathematical tools in parallel with a top-down view of communication networks, and (ii) the presentation of heavy-traffic analysis of queueing models using Lyapunov techniques.

The background required to read the book

Graduate students who have taken a graduate-level course in probability and who have some basic knowledge of optimization and control theory should find the book accessible. An industrious student willing to put in extra effort may find the book accessible even with just a strong undergraduate course in probability. Researchers working in the area of communication networks should be able to read most chapters in the book individually since we have tried to make each chapter as self contained as possible. However, occasionally we refer to results in earlier chapters when discussing the material in a particular chapter, but this overlap between chapters should be small. We have provided a brief introduction to the mathematical background required to understand the various topics in the book, as and when appropriate, to aid the reader.

How to use the book as an instructor

We have taught various graduate-level courses from the material in the book. Based on our experience, we believe that there are two different ways in which this book can be used: to teach either a single course or two courses on communication networks. Below we provide a list of chapters that can be covered for each of these options.

- A two-course sequence on communication networks.
  - Course 1 (modeling and algorithms): Chapters 1–6 except Section 3.5, and Sections 7.1, 7.2, 7.4.1, and Chapter 8. The mathematical background, Sections 2.1 and 2.3, can be taught as and when necessary when dealing with specific topics.
  - Course 2 (performance analysis): Chapters 9 (cover Section 8.2 before Section 9.14), 10, and 11. We recommend reviewing Chapter 3 (except Section 3.5), which would have been covered in Course 1 above, before teaching Chapter 10.
- A single course on communication networks, covering modeling, algorithms, and performance analysis: Chapters 1–6 except Section 3.5, Sections 9.1–9.10 of Chapter 9, and Chapters 10 and 11.

Acknowledgements

This book grew out of courses offered by us at the University of Illinois at Urbana-Champaign, Iowa State University, and Arizona State University. The comments of the students in these courses over the years have been invaluable in shaping the material in the book. We would like to acknowledge Zhen Chen, Javad Ghaderi, Juan Jose Jaramillo, Xiaohan Kang, Joohwan Kim, Siva Theja Maguluri, Chandramani Singh, Weina Wang, Rui Wu, Zhengyu Zhang, and Kai Zhu in particular, who read various parts of the book carefully and provided valuable comments. We also gratefully acknowledge collaborations and/or discussions with Tamer Başar, Atilla Eryilmaz, Bruce Hajek, Frank Kelly, P. R. Kumar, Sean Meyn, Sanjay Shakkottai, Srinivas Shakkottai, Ness Shroff, and Don Towsley over the years, which helped shape the presentation of the material in this book.