

CONTROL TECHNIQUES FOR COMPLEX NETWORKS

Power grids, flexible manufacturing, cellular communications: interconnectedness has consequences. This remarkable book gives the tools and philosophy you need to build network models detailed enough to capture essential dynamics but simple enough to expose the structure of effective control solutions and to clarify analysis.

Core chapters assume only prior exposure to stochastic processes and linear algebra at the undergraduate level; later chapters are for advanced graduate students and researchers/practitioners. This gradual development bridges classical theory with the state of the art. The workload model that is the basis of traditional analysis of the single queue becomes a foundation for workload relaxations used in the treatment of complex networks. Lyapunov functions and dynamic programming equations lead to the celebrated MaxWeight policy along with many generalizations. Other topics include methods for synthesizing hedging and safety stocks, stability theory for networks, and techniques for accelerated simulation.

Examples and figures throughout make ideas concrete. Solutions to end-of-chapter exercises are available on a companion Web site.

Sean Meyn is a professor in the Department of Electrical and Computer Engineering and director of the Division and Control Laboratory of the Coordinated Science Laboratory at the University of Illinois. He has served on the editorial boards of several journals in areas of systems and control and applied probability, and he is coauthor with Richard Tweedie of *Markov Chains and Stochastic Stability*, which won the 1994 ORSA/TIMS Best Publication in Applied Probability Award.

Cambridge University Press
978-0-521-88441-9 - Control Techniques for Complex Networks
Sean Meyn
Frontmatter
[More information](#)

CONTROL TECHNIQUES FOR COMPLEX NETWORKS

SEAN MEYN

University of Illinois, Urbana-Champaign



Cambridge University Press
978-0-521-88441-9 - Control Techniques for Complex Networks
Sean Meyn
Frontmatter
[More information](#)

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521884419

© Sean Meyn 2008

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.

First published 2008

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication Data

Meyn, S. P. (Sean P.)
Control techniques for complex networks / Sean Meyn.
p. cm.

Includes bibliographical references and index.
ISBN 978-0-521-88441-9 (hardback)
1. Computer networks. 2. Control theory. I. Title.
TK5105.5.M49 2008
004.6-dc22 2007035250

ISBN 978-0-521-88441-9 hardback

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

Contents

<i>List of Illustrations</i>	<i>page</i> ix
<i>Preface</i>	xiii
<i>Dedication</i>	xvii
1 Introduction	1
1.1 Networks in practice	2
1.2 Mathematical models	7
1.3 What do you need to know to read this book?	10
1.4 Notes	21
Part I: Modeling and Control	23
2 Examples	25
2.1 Modeling the single server queue	25
2.2 Klimov model	31
2.3 Capacity and queueing in communication systems	34
2.4 Multiple-access communication	34
2.5 Processor sharing model	36
2.6 Inventory model	37
2.7 Power transmission network	37
2.8 Optimization in a simple re-entrant line	39
2.9 Contention for resources and instability	43
2.10 Routing model	46
2.11 Braess' paradox	49
2.12 Notes	50
3 The Single Server Queue	52
3.1 Representations	55
3.2 Approximations	58
3.3 Stability	62
3.4 Invariance equations	66
3.5 Big queues	76
3.6 Model selection	80
3.7 Notes	82
Exercises	82

vi	<i>Contents</i>	
4	Scheduling	87
4.1	Controlled random-walk model	89
4.2	Fluid model	97
4.3	Control techniques for the fluid model	103
4.4	Comparing fluid and stochastic models	115
4.5	Structure of optimal policies	119
4.6	Safety-stocks	122
4.7	Discrete review	128
4.8	MaxWeight and MinDrift	131
4.9	Perturbed value function	134
4.10	Notes	138
	Exercises	139
	Part II: Workload	143
5	Workload and Scheduling	145
5.1	Single server queue	146
5.2	Workload for the CRW scheduling model	149
5.3	Relaxations for the fluid model	153
5.4	Stochastic workload models	172
5.5	Pathwise optimality and workload	178
5.6	Hedging in networks	183
5.7	Notes	192
	Exercises	193
6	Routing and Resource Pooling	195
6.1	Workload in general models	198
6.2	Resource pooling	204
6.3	Routing and workload	209
6.4	MaxWeight for routing and scheduling	215
6.5	Simultaneous resource possession	218
6.6	Workload relaxations	221
6.7	Relaxations and policy synthesis for stochastic models	233
6.8	Notes	240
	Exercises	242
7	Demand	246
7.1	Network models	249
7.2	Transients	255
7.3	Workload relaxations	267
7.4	Hedging in a simple inventory model	274
7.5	Hedging in networks	280
7.6	Summary of steady-state control techniques	291
7.7	Notes	292
	Exercises	293

Part III: Stability and Performance	295
8 Foster–Lyapunov Techniques	297
8.1 Lyapunov functions	302
8.2 Lyapunov functions for networks	305
8.3 Discrete review	315
8.4 MaxWeight	319
8.5 MaxWeight and the average-cost optimality equation	325
8.6 Linear programs for performance bounds	328
8.7 Brownian workload model	336
8.8 Notes	342
Exercises	343
9 Optimization	348
9.1 Reachability and decomposibility	352
9.2 Linear programming formulations	354
9.3 Multiobjective optimization	362
9.4 Optimality equations	365
9.5 Algorithms	375
9.6 Optimization in networks	381
9.7 One-dimensional inventory model	385
9.8 Hedging and workload	391
9.9 Notes	402
Exercises	404
10 ODE Methods	407
10.1 Examples	412
10.2 Mathematical preliminaries	416
10.3 Fluid limit model	419
10.4 Fluid-scale stability	423
10.5 Safety stocks and trajectory tracking	431
10.6 Fluid-scale asymptotic optimality	437
10.7 Brownian workload model	443
10.8 Notes	448
Exercises	450
11 Simulation and Learning	452
11.1 Deciding when to stop	458
11.2 Asymptotic theory for Markov models	461
11.3 The single-server queue	465
11.4 Control variates and shadow functions	470
11.5 Estimating a value function	483
11.6 Notes	498
Exercises	499

Cambridge University Press
978-0-521-88441-9 - Control Techniques for Complex Networks
Sean Meyn
Frontmatter
[More information](#)

viii

Contents

<i>Appendix</i>	Markov Models	505
<i>Bibliography</i>		537
<i>Index</i>		559

List of Illustrations

1.1	Control issues in a production system	<i>page</i> 2
1.2	The Internet is one of the most complex man-made networks	3
1.3	Control issues in the Internet	3
1.4	California power grid	5
1.5	Demand is periodic and shows high variability	6
1.6	Simulating the random walk	13
1.7	Simulation of the M/M/1 queue during a transient	20
1.8	Simulation of the M/M/1 queue in steady-state	20
2.1	Single server queue	26
2.2	Sample paths of the M/M/1 queue and its fluid model	30
2.3	Single-station scheduling problem	31
2.4	Multiple-access communication system	35
2.5	Achievable rates in communication models	35
2.6	Models for the processor-sharing model	36
2.7	Simple make-to-stock system	37
2.8	Power model with two suppliers providing power to a single utility	38
2.9	Simple re-entrant line	39
2.10	Optimal control of the simple fluid re-entrant line	42
2.11	Control of the simple re-entrant line	43
2.12	The Kumar–Seidman–Rybko–Stolyar model	44
2.13	Sample paths of the KSRS model under a priority policy	45
2.14	Simple routing model	46
2.15	Trajectories for the simple routing model	47
2.16	Starvation and instability in the simple routing model	48
2.17	A routing model with delay	49
2.18	Additional choices may lead to increased delay	49
2.19	Delay is doubled when additional choices are presented	50
3.1	Skorokhod representation for the single server queue	57
3.2	Trajectory of the queue length process conditioned on a large value	77
3.3	Log moment generating function for three distributions	81
4.1	A time optimal trajectory for the fluid model	99
4.2	Velocity space for the processor sharing model	101
4.3	Tandem queues	110
4.4	Value function for the tandem queues	111

x	<i>List of Illustrations</i>	
4.5	Time optimal trajectories for the simple re-entrant line	113
4.6	Value function for the simple fluid re-entrant line	114
4.7	Average cost optimal policy for the tandem queues in Case 1	120
4.8	Average cost optimal policy for the tandem queues in Case 2	121
4.9	Discounted cost optimal policy for the tandem queues in Case 1	121
4.10	Discounted cost optimal policy for the tandem queues in Case 2	121
4.11	Optimal control of the simple CRW re-entrant line	122
4.12	A comparison of four policies	124
4.13	Estimates of the steady-state customer population in the KSRS model	125
4.14	Optimal policy for the processor sharing model	126
4.15	The five-buffer model of Dai and Wang	127
4.16	Average cost for the model of Dai and Wang	128
4.17	Level sets of the perturbed linear cost function for the tandem queues	136
4.18	Criss-cross network	140
5.1	Velocity space for the processor-sharing model	155
5.2	Three-station network	156
5.3	Level sets of the effective cost for the KSRS model	162
5.4	Level sets of the value function \hat{J} for the KSRS model	162
5.5	Effective cost with buffer constraints	164
5.6	Effective cost and monotone region for the simple re-entrant line	165
5.7	Effective cost and monotone region with buffer constraints	166
5.8	Optimal trajectories for the fluid model and its relaxation	173
5.9	Cost trajectories for workload-relaxations in the simple re-entrant line	173
5.10	Extreme points in the Klimov model optimization problem	181
5.11	Workload model and the height process	185
5.12	Average cost using an affine policy: Case II (a)	188
5.13	Average cost using an affine policy: Case II (b)	188
5.14	Average cost using an affine policy in Case III	189
5.15	Average cost for the simple re-entrant line with buffer constraints	191
6.1	Network routing	196
6.2	Minimal draining time and velocity space	201
6.3	Redundant server	206
6.4	Velocity space for the simple routing model	208
6.5	Velocity spaces in two versions of the ALOHA communication model	208
6.6	Single-user routing model	209
6.7	Single-user routing model with min-cut	212
6.8	Simultaneous resource possession in aircraft scheduling	218
6.9	Simultaneous resource possession	219
6.10	Input-queued switch and a matching of inputs to outputs	220
6.11	Velocity space \hat{V} is an unbounded relaxation of V	222
6.12	Velocity space for the simple routing model and its relaxation	223
6.13	One-dimensional relaxations in the ALOHA communication model	223
6.14	Coupling of optimal solutions in the simple routing model	229
6.15	Input-queued switch and a matching of inputs to outputs	235
6.16	Starvation and instability in the simple routing model	237
6.17	Optimal policy for the simple routing model	238
6.18	Monotone region under MaxWeight with and without hedging	240
6.19	Simulation using MaxWeight with and without hedging	240

List of Illustrations

xi

7.1	Demand-driven model with routing, scheduling, and rework	247
7.2	Optimal policies for the complex production system	249
7.3	Conversion from pull to push	251
7.4	Velocity space V_0 in the arrival-free model	252
7.5	Simple re-entrant line with demand	253
7.6	Simple re-entrant line with demand under the GTO policy	256
7.7	Demand-driven network with two sources of demand	257
7.8	Individual buffer levels under the GTO policy	258
7.9	Cumulative buffer levels under the GTO policy	261
7.10	Cumulative buffer levels for a normal-priority hot-lot	263
7.11	Subsystem serving the hot-lot	263
7.12	Cumulative buffer levels for a high-priority hot-lot	264
7.13	GTO-B and GTO-M policies	267
7.14	Optimal policy for a 1-d relaxation	269
7.15	Effective cost for a 1-d relaxation of the 16-buffer model	270
7.16	Effective cost for the simple re-entrant line with demand	272
7.17	Effective cost for the 2-d relaxation of the 16-buffer model	273
7.18	R^* contains the monotone region, and R^{GTO} contains R^*	274
7.19	Effective cost for the 3-d relaxation of the 16-buffer model	274
7.20	Ancillary service ramps up at maximal rate when $Q(t) < \bar{q}^a$	285
7.21	Average cost for the CRW and CBM power distribution models	286
7.22	Optimal policies in a power distribution system	287
7.23	Optimal switching-curves for a relaxation in Case II	289
8.1	Smoothed piecewise linear Lyapunov function	312
8.2	Fluid value functions for the tandem queue	314
8.3	Feasibility of the linear test for stability in the KSRS model	331
9.1	Extreme points in average-cost optimization	363
9.2	Extreme points in multi-objective average-cost optimization	363
9.3	Construction of the relative value function	370
9.4	Convergence of the VIA with quadratic initialization	379
9.5	Convergence of the VIA initialized with the fluid value function	379
9.6	Sample paths of the two workload processes	401
10.1	Service-aware scheduling	415
11.1	Monte Carlo estimates for the M/M/1 queue	455
11.2	Monte Carlo estimates for the M/M/1 queue with c exponential	456
11.3	Estimates of the steady-state customer population for the KSRS model	457
11.4	Log moment generating function and rate function	460
11.5	Log moment generating function and rate function for the CRW queue	465
11.6	Monte Carlo estimates for the CRW queue	474
11.7	Estimates obtained using the smoothed estimator	474
11.8	Smoothed estimates for the M/M/1 queue	475
11.9	Estimates obtained using the smoothed estimator for the KSRS model	482
11.10	LSTD estimates for value function approximation in the M/M/1 queue	495
11.11	LSTD estimates for optimizing shadow functions in the M/M/1 queue	498
A1.1	$V(X(t))$ is decreasing outside of the set S	521

Preface

A representative of a major publishing house is on her way home from a conference in Singapore, excited about the possibility of a new book series. On the flight home to New York she opens her blackberry organizer, adding names of new contacts, and is disappointed to realize she may have caught the bug that was bothering her friend Alex at the café near the conference hotel. When she returns home she will send Alex an email to see how she's doing and to make sure this isn't a case of some new dangerous flu.

Of course, the publisher is aware that she is part of an interconnected network of other business men and women and their clients: Her value as an employee depends on these connections. She depends on the transportation network of taxis and airplanes to get her job done and is grateful for the most famous network today that allows her to contact her friend effortlessly even when separated by thousands of miles. Other networks of even greater importance escape her consciousness, even though consciousness itself depends on a highly interconnected fabric of neurons and vascular tissue. Communication networks are critical to support the air traffic controllers who manage the airspace around her. A supply chain of manufacturers makes her book business possible, as well as the existence of the airplane on which she is flying.

Complex networks are everywhere. Interconnectedness is as important to business men and women as it is to the viruses who travel along with them.

Much of the current interest in networks within physics and the biological sciences is phenomenological. For example, given a certain degree of connectivity between individuals, what is the likelihood that a virus will spread to the extinction of the planet? Degree and mode of connectivity in passive agents can combine to form images resembling crystals or snowflakes [463].

The main focus within our own bodies is far more utilitarian. Endocrine, immune, and vascular systems adjust chemical reactions to maintain equilibria in the face of ongoing attacks from disease and diet. In biology this is called *homeostasis*. In this book, the regulation of a network is called *control*.

It is not our goal to take on biology, computer science, communications, and operations research in a single volume. Rather, the intended purpose of this book is an introduction to a rapidly evolving engineering discipline. The examples come from applications in which complexity is real, but less daunting than that found in the human

brain. We describe methods to model networks in order to capture essential structure, dynamics, and uncertainty. Based on these models we explore ways to visualize network behavior so that effective control techniques can be synthesized and evaluated.

Modeling and control. The operator of an electric power grid hopes to find a network model that will help form predictions of supply and demand to maintain stability of the power network. This requires the expertise of statisticians, economists, and power engineers. The resulting model may provide useful simulations for forecasting, but will fail entirely for our purposes. This book is about control, and for this it is necessary to restrict to models that capture essential behavior, but no more.

Modeling for the purposes of control and the development of control techniques for truly complex networks has become a major research activity over the past two decades. Breakthroughs obtained in the stochastic networks community provide important tools that have had real impact in some application areas, such as the implementation of MaxWeight scheduling for routing and scheduling in communications. Other breakthroughs have had less impact due in part to the highly technical and mathematical language in which the theory has developed. The goal of this book is to expose these ideas in the simplest possible setting.

Most of the ideas in this book revolve around a few concepts.

- (i) The *fluid model* is an idealized deterministic model. In a communication network a unit of “fluid” corresponds to some quantities of packets; in a power network this might correspond to a certain number of megawatts of electricity.

A fluid model is often a starting point to understand the impact of topology, processing rates, and external arrivals on network behavior. Based on the fluid model we can expose the inherent conflict between short-sighted control objectives, longer-range issues such as recovery from a singular external disruption, and truly long-range planning such as the *design* of appropriate network topology.

- (ii) Refinements of the fluid model are developed to capture variability in supply, demand, or processing rates. The *controlled random walk* model favored in this book is again a highly stylized model of any real network, but contains enough structure to give a great deal of insight and is simple enough to be tractable for developing control techniques.

For example, this model provides a vehicle for constructing and evaluating *hedging* mechanisms to limit exposure to high costs, and to ensure that valuable resources can operate when needed.

- (iii) The concept of *workload* is developed for the deterministic and stochastic models. Perhaps the most important concept in this book is the *workload relaxation* that provides approximations of a highly complex network by a far simpler one. The approximation may be crude in some cases, but its value in attaining intuition can be outstanding.
- (iv) Methods from the stability theory of Markov models form a foundation in the treatment of stochastic network models. Lyapunov functions are a basis of

dynamic programming equations for optimization, for stability and analysis, and even for developing algorithms based on simulation.

What's in here? The book is divided into three parts. The first part, entitled Modeling and Control, contains numerous examples to illustrate some of the basic concepts developed in the book, especially those topics listed in (i) and (ii) concerning the fluid and CRW models. Lyapunov functions and the dynamic programming equations are introduced; based on these concepts we arrive at the MaxWeight policy along with many generalizations.

Workload relaxations are introduced in Part II. In these three chapters we show how a cost function defined for the network can be “projected” to define the *effective cost* for the relaxation. Applications to control involve first constructing a policy for the low-dimensional relaxation, and then translating this to the original physical system of interest. This translation step involves the introduction of hedging to guard against variability.

Most of the control techniques are contained in the first two parts of the book. Part III, entitled Stability and Performance, contains an in-depth treatment of Lyapunov stability theory and optimization. It contains approximation techniques to explain the apparent solidarity between control solutions for stochastic and deterministic network models. Moreover, this part of the book develops several approaches to performance evaluation for stochastic network models.

Who's it for? The book was created for several audiences. The gradual development of network concepts in Parts I and II was written with the first-year graduate student in mind. This reader may have had little exposure to operations research concepts, but some prior exposure to stochastic processes and linear algebra at the undergraduate level.

Many of the topics in the latter chapters are at the frontier of stochastic networks, optimization, simulation, and learning. This material is intended for the more advanced graduate student, as well as researchers and practitioners in any of these areas.

Acknowledgments

This book has been in the making for 5 years and over this time has drawn inspiration and feedback from many. Some of the ideas were developed in conjunction with students, including Mike Chen, Richard Dubrawski, Charuhas Pandit, Rong-Rong Chen, and David Eng. In particular, the numerics in Section 7.2 are largely taken from Dubrawski's thesis [151], and the *diffusion heuristic* for hedging is based on a paper with Chen and Pandit [103]. Section 9.6 is based in part on research conducted with Chen [105].

My collaborators are a source of inspiration and friendship. Many of the ideas in this book revolve around stochastic Lyapunov theory for Markov processes, which is summarized in the appendix. This appendix is essentially an abridged version of my

book coauthored with Richard Tweedie [368]. Vivek Borkar's research on Markov decision theory (as summarized in [70]) has had a significant influence on my own view of optimization. My interest in networks was sparked by a lecture presented by P. R. Kumar when he was visiting the Australian National University in 1988 while I resided there as a postdoctoral fellow. He became a mentor and a coauthor when I joined the University of Illinois the following year. I learned of the beauty of simulation theory from the work of Peter Glynn and his former student Shane Henderson. More recent collaborators are Profs. In-Koo Cho, David Gamarnik, Ioannis Kontoyiannis, and Eric Moulines, who have provided inspiration on a broad range of topics. I am grateful to Devavrat Shah and Damon Wischik for sharing insights on the "input-queued switch," and for allowing me to adapt a figure from their paper [435] that is used to illustrate workload relaxations in Section 6.7.1. Pierre L'Ecuyer shared his notes on simulation from his course at the University of Montréal, and Bruce Hajek at the University of Illinois shared his lecture notes on communication networks.

Profs. Cho, Kontoyiannis, Henderson, and Shah have all suggested improvements on exposition, or warned of typos. Sumit Bhardwaj, Jinjing Jiang, Shie Mannor, Eric Moulines, and Michael Veatch each spent significant hours pouring through selected chapters of draft text and provided valuable feedback. Early input by Veatch moved the book toward its present organization, with engineering techniques introduced first, and harder mathematics postponed to later chapters.

Any remaining errors or awkward prose are, of course, my own.

It would be impossible to write a book like this without financial support for graduate students and release-time for research. I am sincerely grateful to the National Science Foundation, in particular, the Division of Electrical, Communications & Cyber Systems, for ongoing support during the writing of this book. The DARPA ITMANET initiative, the Laboratory for Information and Decision Systems at MIT, and United Technologies Research Center provided support in the final stages of this project during the 2006–2007 academic year.

Equally important has been support from my family, especially during the last months, when I have been living away from home. Thank you Belinda! Thank you Sydney and Sophie! And thanks also to all the poodles at South Harding Drive.

Dedication

It was a sad day on June 7, 2001, when Richard Tweedie died at the peak of his career. A brief survey of his contributions to applied probability and statistics can be found in [154].

In memory of his friendship and collaboration, and in honor of his many contributions to our scientific communities, this book is dedicated to Richard.

Cambridge University Press
978-0-521-88441-9 - Control Techniques for Complex Networks
Sean Meyn
Frontmatter
[More information](#)

CONTROL TECHNIQUES FOR COMPLEX NETWORKS