

Second Edition

Data Networks

DIMITRI BERTSEKAS

Massachusetts Institute of Technology

ROBERT GALLAGER

Massachusetts Institute of Technology



PRENTICE HALL, Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging-in-Publication Data

Bertsekas, Dimitri P.

Data networks / Dimitri Bertsekas, Robert Gallager. -- 2nd ed.

p. cm.

Includes bibliographical references and index.

ISBN 0-13-200916-1

1. Data transmission systems. I. Gallager, Robert G. II. Title.

TK5105.B478 1992

004.6--dc20

91-35561

CIP

Acquisitions editor: Pete Janzow
Production editor: Bayani Mendoza de Leon
Copy editor: Zeiders & Associates
Cover designer: Butler/Udell
Prepress buyer: Linda Behrens
Manufacturing buyer: Dave Dickey
Editorial assistant: Phyllis Morgan



©1992, 1987 by Prentice-Hall, Inc.
A Simon & Schuster Company
Englewood Cliffs, New Jersey 07632

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-200916-1

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S. A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Simon & Schuster Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

Contents

	PREFACE	xv
1	INTRODUCTION AND LAYERED NETWORK ARCHITECTURE	1
1.1	Historical Overview	1
1.1.1	Technological and Economic Background,	5
1.1.2	Communication Technology,	6
1.1.3	Applications of Data Networks,	7
1.2	Messages and Switching	9
1.2.1	Messages and Packets,	9
1.2.2	Sessions,	11
1.2.3	Circuit Switching and Store-and-Forward Switching,	14
1.3	Layering	17
1.3.1	The Physical Layer,	20
1.3.2	The Data Link Control Layer,	23
	<i>The MAC sublayer,</i>	<i>24</i>
1.3.3	The Network Layer,	25
	<i>The Internet sublayer,</i>	<i>28</i>
1.3.4	The Transport Layer,	29
1.3.5	The Session Layer,	30

1.3.6	The Presentation Layer, 31	
1.3.7	The Application Layer, 31	
1.4	A Simple Distributed Algorithm Problem	32
	Notes and Suggested Reading	35
	Problems	35
2	POINT-TO-POINT PROTOCOLS AND LINKS	37
2.1	Introduction	37
2.2	The Physical Layer: Channels and Modems	40
2.2.1	Filtering,	41
2.2.2	Frequency Response,	43
2.2.3	The Sampling Theorem,	46
2.2.4	Bandpass Channels,	47
2.2.5	Modulation,	48
2.2.6	Frequency- and Time-Division Multiplexing,	52
2.2.7	Other Channel Impairments,	53
2.2.8	Digital Channels,	53
	<i>ISDN,</i>	<i>54</i>
2.2.9	Propagation Media for Physical Channels,	56
2.3	Error Detection	57
2.3.1	Single Parity Checks,	58
2.3.2	Horizontal and Vertical Parity Checks,	58
2.3.3	Parity Check Codes,	59
2.3.4	Cyclic Redundancy Checks,	61
2.4	ARQ: Retransmission Strategies	64
2.4.1	Stop-and-Wait ARQ,	66
	<i>Correctness of stop and wait,</i>	<i>69</i>
2.4.2	Go Back n ARQ,	72
	<i>Rules followed by transmitter and receiver in go back</i>	<i>n, 74</i>
	<i>Correctness of go back n,</i>	<i>76</i>
	<i>Go back n with modulus $m > n$,</i>	<i>78</i>
	<i>Efficiency of go back n implementations,</i>	<i>80</i>
2.4.3	Selective Repeat ARQ,	81
2.4.4	ARPANET ARQ,	84
2.5	Framing	86

- 2.5.1 Character-Based Framing, 86
- 2.5.2 Bit-Oriented Framing: Flags, 88
- 2.5.3 Length Fields, 90
- 2.5.4 Framing with Errors, 92
- 2.5.5 Maximum Frame Size, 93
 - Variable frame length, 93*
 - Fixed frame length, 97*
- 2.6 Standard DLCs **97**
- 2.7 Initialization and Disconnect for ARQ Protocols **103**
 - 2.7.1 Initialization in the Presence of Link Failures, 103
 - 2.7.2 Master–Slave Protocol for Link Initialization, 104
 - 2.7.3 A Balanced Protocol for Link Initialization, 107
 - 2.7.4 Link Initialization in the Presence of Node Failures, 109
- 2.8 Point-to-Point Protocols at the Network Layer **110**
 - 2.8.1 Session Identification and Addressing, 111
 - Session identification in TYMNET, 112*
 - Session identification in the Codex networks, 113*
 - 2.8.2 Packet Numbering, Window Flow Control, and Error Recovery, 114
 - Error recovery, 115*
 - Flow control, 116*
 - Error recovery at the transport layer versus the network layer, 117*
 - 2.8.3 The X.25 Network Layer Standard, 118
 - 2.8.4 The Internet Protocol, 120
- 2.9 The Transport Layer **123**
 - 2.9.1 Transport Layer Standards, 123
 - 2.9.2 Addressing and Multiplexing in TCP, 124
 - 2.9.3 Error Recovery in TCP, 125
 - 2.9.4 Flow Control in TCP/IP, 127
 - 2.9.5 TP Class 4, 128
- 2.10 Broadband ISDN and the Asynchronous Transfer Mode **128**
 - 2.10.1 Asynchronous Transfer Mode (ATM), 132
 - 2.10.2 The Adaptation Layer, 135
 - Class 3 (connection-oriented) traffic, 136*
 - Class 4 (connectionless) traffic, 137*
 - Class 1 and 2 traffic, 137*
 - 2.10.3 Congestion, 138

	Summary	139
	Notes, Sources, and Suggested Reading	140
	Problems	141
3	DELAY MODELS IN DATA NETWORKS	149
3.1	Introduction	149
3.1.1	Multiplexing of Traffic on a Communication Link,	150
3.2	Queueing Models: Little's Theorem	152
3.2.1	Little's Theorem,	152
3.2.2	Probabilistic Form of Little's Theorem,	154
3.2.3	Applications of Little's Theorem,	157
3.3	The $M/M/1$ Queueing System	162
3.3.1	Main Results,	164
	<i>Arrival statistics—the Poisson process,</i>	164
	<i>Service statistics,</i>	165
	<i>Markov chain formulation,</i>	166
	<i>Derivation of the stationary distribution,</i>	167
3.3.2	Occupancy Distribution upon Arrival,	171
3.3.3	Occupancy Distribution upon Departure,	173
3.4	The $M/M/m$, $M/M/\infty$, $M/M/m/m$, and Other Markov Systems	173
3.4.1	$M/M/m$: The m -Server Case,	174
3.4.2	$M/M/\infty$: The Infinite-Server Case,	177
3.4.3	$M/M/m/m$: The m -Server Loss System,	178
3.4.4	Multidimensional Markov Chains: Applications in Circuit Switching,	180
	<i>Truncation of independent single-class systems,</i>	182
	<i>Blocking probabilities for circuit switching systems,</i>	185
3.5	The $M/G/1$ System	186
3.5.1	$M/G/1$ Queues with Vacations,	192
3.5.2	Reservations and Polling,	195
	<i>Single-user system,</i>	196
	<i>Multi-user system,</i>	198
	<i>Limited service systems,</i>	201
3.5.3	Priority Queueing,	203
	<i>Nonpreemptive priority,</i>	203

	<i>Preemptive resume priority</i> , 205	
	3.5.4 An Upper Bound for the $G/G/1$ System, 206	
3.6	Networks of Transmission Lines 209	
	3.6.1 The Kleinrock Independence Approximation, 211	
3.7	Time Reversibility—Burke’s Theorem 214	
3.8	Networks of Queues—Jackson’s Theorem 221	
	<i>Heuristic explanation of Jackson’s Theorem</i> , 227	
	3.8.1 Extensions of Jackson’s Theorem, 229	
	<i>State-dependent service rates</i> , 229	
	<i>Multiple classes of customers</i> , 230	
	3.8.2 Closed Queueing Networks, 233	
	3.8.3 Computational Aspects—Mean Value Analysis, 238	
	Summary 240	
	Notes, Sources, and Suggested Reading 241	
	Problems 242	
	Appendix A: Review of Markov Chain Theory 259	
	3A.1 Discrete-Time Markov Chains, 259	
	3A.2 Detailed Balance Equations, 261	
	3A.3 Partial Balance Equations, 262	
	3A.4 Continuous-Time Markov Chains, 262	
	3A.5 Drift and Stability, 264	
	Appendix B: Summary of Results 265	
4	MULTIACCESS COMMUNICATION	271
4.1	Introduction 271	
	4.1.1 Satellite Channels, 273	
	4.1.2 Multidrop Telephone Lines, 274	
	4.1.3 Multitapped Bus, 274	
	4.1.4 Packet Radio Networks, 275	
4.2	Slotted Multiaccess and the Aloha System 275	
	4.2.1 Idealized Slotted Multiaccess Model, 275	
	<i>Discussion of assumptions</i> , 276	

- 4.2.2 Slotted Aloha, 277
- 4.2.3 Stabilized Slotted Aloha, 282
 - Stability and maximum throughput*, 282
 - Pseudo-Bayesian algorithm*, 283
 - Approximate delay analysis*, 284
 - Binary exponential backoff*, 286
- 4.2.4 Unslotted Aloha, 287
- 4.3 Splitting Algorithms **289**
 - 4.3.1 Tree Algorithms, 290
 - Improvements to the tree algorithm*, 292
 - Variants of the tree algorithm*, 293
 - 4.3.2 First-Come First-Serve Splitting Algorithms, 293
 - Analysis of FCFS splitting algorithm*, 297
 - Improvements in the FCFS splitting algorithm*, 301
 - Practical details*, 302
 - Last-come first-serve (LCFS) splitting algorithm*, 302
 - Delayed feedback*, 303
 - Round-robin splitting*, 304
- 4.4 Carrier Sensing **304**
 - 4.4.1 CSMA Slotted Aloha, 305
 - 4.4.2 Pseudo-Bayesian Stabilization for CSMA Aloha, 307
 - 4.4.3 CSMA Unslotted Aloha, 309
 - 4.4.4 FCFS Splitting Algorithm for CSMA, 310
- 4.5 Multiaccess Reservations **312**
 - 4.5.1 Satellite Reservation Systems, 313
 - 4.5.2 Local Area Networks: CSMA/CD and Ethernet, 317
 - Slotted CSMA/CD*, 317
 - Unslotted CSMA/CD*, 318
 - The IEEE 802 standards*, 320
 - 4.5.3 Local Area Networks: Token Rings, 320
 - IEEE 802.5 token ring standard*, 323
 - Expected delay for token rings*, 324
 - FDDI*, 326
 - Slotted rings and register insertion rings*, 330
 - 4.5.4 Local Area Networks: Token Buses and Polling, 331
 - IEEE 802.4 token bus standard*, 332
 - Implicit tokens: CSMA/CA*, 333
 - 4.5.5 High-Speed Local Area Networks, 333
 - Distributed queue dual bus (IEEE 802.6)*, 335

Expressnet, 339

Homenets, 341

4.5.6 Generalized Polling and Splitting Algorithms, 342

4.6 Packet Radio Networks **344**

4.6.1 TDM for Packet Radio Nets, 346

4.6.2 Collision Resolution for Packet Radio Nets, 347

4.6.3 Transmission Radii for Packet Radio, 349

4.6.4 Carrier Sensing and Busy Tones, 350

Summary **351**

Notes, Sources, and Suggested Reading **352**

Problems **353**

5 ROUTING IN DATA NETWORKS **363**

5.1 Introduction **363**

5.1.1 Main Issues in Routing, 365

5.1.2 Wide-Area Network Routing: An Overview, 367

Flooding and broadcasting, 368

Shortest path routing, 370

Optimal routing, 372

Hot potato (deflection) routing schemes, 372

Cut-through routing, 373

ARPANET: An example of datagram routing, 374

TYMNET: An example of virtual circuit routing, 376

Routing in SNA, 378

Routing in circuit switching networks, 379

5.1.3 Interconnected Network Routing: An Overview, 379

Bridged local area networks, 382

Spanning tree routing in bridged local area networks, 383

Source routing in bridged local area networks, 385

5.2 Network Algorithms and Shortest Path Routing **387**

5.2.1 Undirected Graphs, 387

5.2.2 Minimum Weight Spanning Trees, 390

5.2.3 Shortest Path Algorithms, 393

The Bellman-Ford algorithm, 396

Bellman's equation and shortest path construction, 399

Dijkstra's algorithm, 401

The Floyd-Warshall algorithm, 403

5.2.4 Distributed Asynchronous Bellman-Ford Algorithm, 404

5.2.5	Stability of Adaptive Shortest Path Routing Algorithms, 410	
	<i>Stability issues in datagram networks, 410</i>	
	<i>Stability issues in virtual circuit networks, 414</i>	
5.3	Broadcasting Routing Information: Coping with Link Failures	418
5.3.1	Flooding: The ARPANET Algorithm, 420	
5.3.2	Flooding without Periodic Updates, 422	
5.3.3	Broadcast without Sequence Numbers, 425	
5.4	Flow Models, Optimal Routing, and Topological Design	433
5.4.1	Overview of Topological Design Problems, 437	
5.4.2	Subnet Design Problem, 439	
	<i>Capacity assignment problem, 439</i>	
	<i>Heuristic methods for capacity assignment, 442</i>	
	<i>Network reliability issues, 445</i>	
	<i>Spanning tree topology design, 447</i>	
5.4.3	Local Access Network Design Problem, 448	
5.5	Characterization of Optimal Routing	451
5.6	Feasible Direction Methods for Optimal Routing	455
5.6.1	The Frank–Wolfe (Flow Deviation) Method, 458	
5.7	Projection Methods for Optimal Routing	464
5.7.1	Unconstrained Nonlinear Optimization, 465	
5.7.2	Nonlinear Optimization over the Positive Orthant, 467	
5.7.3	Application to Optimal Routing, 468	
5.8	Routing in the Codex Network	476
	Summary	477
	Notes, Sources, and Suggested Reading	478
	Problems	479
6	FLOW CONTROL	493
6.1	Introduction	493
6.1.1	Means of Flow Control, 494	
6.1.2	Main Objectives of Flow Control, 496	
	<i>Limiting delay and buffer overflow, 496</i>	
	<i>Fairness, 498</i>	

6.2	Window Flow Control	500	
6.2.1	End-to-End Windows,	501	
	<i>Limitations of end-to-end windows,</i>	502	
6.2.2	Node-by-Node Windows for Virtual Circuits,	506	
6.2.3	The Isarithmic Method,	508	
6.2.4	Window Flow Control at Higher Layers,	508	
6.2.5	Dynamic Window Size Adjustment,	510	
6.3	Rate Control Schemes	510	
	<i>Queueing analysis of the leaky bucket scheme,</i>	513	
6.4	Overview of Flow Control in Practice	515	
	<i>Flow control in the ARPANET,</i>	515	
	<i>Flow control in the TYMNET,</i>	517	
	<i>Flow control in SNA,</i>	517	
	<i>Flow control in a Codex network,</i>	518	
	<i>Flow control in the PARIS network,</i>	518	
	<i>Flow control in X.25,</i>	519	
6.5	Rate Adjustment Algorithms	519	
6.5.1	Combined Optimal Routing and Flow Control,	519	
6.5.2	Max-Min Flow Control,	524	
	Summary	530	
	Notes, Sources, and Suggested Reading	530	
	Problems	531	
	REFERENCES		537
	INDEX		552