

Contents

Preface	v
Acknowledgments	ix
1 Introduction	1
1.1 Role of Modeling and Analysis	1
1.1.1 Modeling and Analysis in Science and Engineering	1
1.1.2 Historical Perspective in Modeling and Analysis	5
1.2 Performance Measures	7
1.3 Modeling Approaches	9
1.3.1 Decomposition and Hierarchical Modeling	9
1.3.2 Analytic Models	12
1.3.3 Simulation Models	13
1.3.4 Measurement and Empirical Models	15
1.4 Examples of Performance Modeling	16
1.4.1 Cellular Radio Network	17
1.4.2 Window-based Flow Control Protocols	19
1.4.3 Call Blocking in Communication Networks	22
1.4.4 Multiplexer Models and Performance Analysis	24
1.4.5 Web Server Systems	29
1.4.6 Internet Traffic Characterization	32
1.4.7 Polling and Token Passing in Local Area Networks	35
1.5 Discussion and Further Reading	37
I Basic Queueing and Loss Models	39
2 Basic Queueing Models	41
2.1 Introduction	41
2.1.1 Shorthand Notation for Queueing and Loss Models	44
2.1.2 Discrete-Time vs. Continuous-Time Models	45
2.1.3 Long-Range Dependent Traffic Models	46
2.2 Little's Formula and its Generalization	47
2.2.1 Formula $L = \lambda W$	47
2.2.2 Formula $H = \lambda G$	51
2.3 Birth-and-Death Processes	53
2.3.1 Poisson Process	53
2.3.2 Properties of Poisson Process	59
2.3.3 Birth-and-Death (BD) Process	63

2.4	Birth-and-Death Queueing Models	68
2.4.1	M/M/1: The Simplest Queueing System	68
2.4.2	M/M/ ∞ and M/G/ ∞ : Infinite Servers	77
2.4.3	M/M/ m : Multiple Server Queue	79
2.4.4	M(K)/M/1 (or M/M/1/ K/K): Machine Servicing Model	83
2.5	Discussion and Further Reading	93
3	Basic Loss Models	95
3.1	Erlang Loss Model: M/M/ $m(0)$ or M/M/ m/m	95
3.1.1	Blocking Probability vs. Loss Probability	97
3.1.2	Calculation of Erlang's Loss Formula	97
3.1.3	Carried Load	99
3.2	Engset Loss Model: M(K)/M/ $m(0)$ or M/M/ $m/m/K$	102
3.2.1	Probability Distribution seen by Arrivals	105
3.2.2	Dimensioning Curves	107
3.3	Insensitivity and Generalization of Loss Models	111
3.4	Application Example: Blocking Analysis of Cellular Communication Systems	112
3.4.1	Frequency Division Multiple Access (FDMA)	114
3.4.2	Time Division Multiple Access (TDMA)	115
3.4.3	Code Division Multiple Access (CDMA)	117
3.4.4	Evaluation of Soft Blocking Probability:	119
3.5	Discussion and Further Reading	122
4	Non-Markovian Queues	124
4.1	Renewal Process and Residual Life	124
4.2	Representations of General Distributions	129
4.2.1	Gamma, Erlangian, and Hyperexponential Distributions	129
4.2.2	Generalized Stage and Cox's Representations	134
4.3	M/G/1 Queueing Model	138
4.3.1	Embedded Markov Chain Technique	138
4.3.2	Queue Distribution in M/G/1	140
4.3.3	Pollaczek-Khintchine Mean Value Formula	145
4.3.4	Waiting Time Distribution in M/G/1	147
4.3.5	Busy Periods in M/G/1	151
4.4	G/M/1 Queueing Model	158
4.4.1	Transition Matrix of Embedded Markov Chain	158
4.4.2	Asymptotic Representation and Characteristic Root	160
4.4.3	Distributions of $N(t)$ and Waiting Time	162
4.4.4	Generalization to G/M/ m Queueing Model	164
4.5	G/G/1 Queue and Waiting Time Distribution	168
4.6	Discussion and Further Reading	174

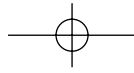
5 Quasi-reversibility and Queues with Product-form Solutions 175

- 5.1 Markov Processes and Markov Chains 175
 - 5.1.1 Discrete-time Markov Chain (DTMC) 176
 - 5.1.2 Continuous-Time Markov Chain (CTMC) 180
 - 5.1.3 Balance Equations 184
- 5.2 Departure Processes, Reversibility, and Quasi-Reversibility 186
 - 5.2.1 Departure Processes 186
 - 5.2.2 Reversible Markov Processes 189
 - 5.2.3 Quasi-reversible Queue 193
 - 5.2.4 Scheduling and Queue Management 198
 - 5.2.5 Symmetric Queues 199
- 5.3 M/G/∞ Queueing Model 206
- 5.4 M/G/1 with Processor Sharing 210
 - 5.4.1 Quasi-reversibility of M/G/1 with PS 211
 - 5.4.2 Conditional Mean Response Time of M/G/1 with PS 212
- 5.5 M/G/1 with LCFS 216
- 5.6 M/G/1 with LCFS-PR 218
- 5.7 Multiple Customer Classes and Quasi-Reversible Stations 218
 - 5.7.1 Multiclass M/M/1 with Queue-dependent Service Rate 219
 - 5.7.2 Multiclass Symmetric Queue 221
 - 5.7.3 Multiclass M/G/∞ 223
 - 5.7.4 Generalized Multiclass M/G/∞ 225
 - 5.7.5 Multiclass M/G/1 with Processor Sharing 228
 - 5.7.6 Multiclass M/G/1 with LCFS-PR 229
- 5.8 Discussion and Further Reading 231

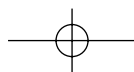
II Queueing and Loss Networks 233

6 Queueing Networks 235

- 6.1 Introduction 235
- 6.2 Jackson Networks: Queueing Networks with Exponential Servers 237
 - 6.2.1 Open Jackson Networks 237
 - 6.2.2 Closed Jackson Networks 247
 - 6.2.3 Convolution Algorithm for Jackson Networks 251
 - 6.2.4 Norton's Theorem and Flow-equivalent Server 261
 - 6.2.5 Multiple Routing Chains 267
- 6.3 Networks with Quasi-Reversible Stations and Multiple Classes of Customers 275
 - 6.3.1 Open Queueing Networks with Multiple Classes of Customers 277
 - 6.3.2 Closed Queueing Networks with Multiple Classes of Customers 283
 - 6.3.3 Mixed Queueing Network with Multiple Closed and Open Chains 285



6.4	Higher-Order Markov Routing	286
6.5	Discussion and Further Reading	294
7	Loss Networks and Generalized Loss Models	296
7.1	Introduction	296
7.2	Generalized Loss Stations	297
7.2.1	Classical Erlang Loss Model Revisited	297
7.2.2	State-Dependent Arrival and Service Rates	298
7.2.3	Engset Loss Model with State-Dependent Service Rate	299
7.3	Loss Networks With Fixed Routing	309
7.3.1	Open Loss Networks	309
7.3.2	Closed and Mixed Loss Networks	311
7.3.3	Link State Distribution	319
7.3.4	Probability Generating Functions of Stationary Distributions	322
7.3.5	Generating Function of Normalization Constants	325
7.4	Queueing-Loss Networks	333
7.5	Discussion and Further Reading	335
8	Computational Algorithms for Product-Form Networks	337
8.1	Introduction	337
8.2	Computational Algorithms for Queueing Networks	338
8.2.1	Application of Pólya Enumeration to Closed Networks	338
8.2.2	Mean Value Analysis (MVA)	341
8.2.3	Integral Representation and Asymptotic Expansion	346
8.3	Parametric Decomposition of a Queueing Network	354
8.3.1	Parametric Decomposition of an Open Network	352
8.3.2	Parametric Decomposition of a Closed Network	354
8.4	Computational Algorithms for Loss Networks	355
8.4.1	Recursive Algorithms for Classical Loss Stations	356
8.4.2	Recursive Algorithms for Generalized Loss Models	358
8.4.3	Recursive Algorithms for Loss Networks	361
8.4.4	Numerical Evaluation of Inversion Integral for Loss Networks	362
8.4.5	Asymptotic Approximation of the Inversion Integral	365
8.4.6	Approximations based on Large-Scale Asymptotics	374
8.5	Reduced Load Approximation of a Loss Network	375
8.6	Discussion and Further Reading	378
III	Advanced Queueing Models	381
9	Conservation Laws, Priority Queues, and Polling Models	383
9.1	Introduction	383
9.2	Work-Conserving Queue Disciplines and Conservation Laws	383



9.2.1	Work-conserving Queue Disciplines	383
9.2.2	Conservation Law for G/G/1	385
9.2.3	Conservation Law for M/G/1	388
9.2.4	Pollaczek–Khinchine Formula	389
9.3	M/G/1 Priority Queues	390
9.3.1	Conservation Law for M/G/1 with Nonpreemptive Priority	391
9.3.2	M/G/1 with Nonpreemptive Priority Scheduling	392
9.3.3	M/G/1 with Preemptive Priority Scheduling	394
9.4	M/G/1 with Server Vacations and Polling Models	396
9.4.1	M/G/1 with Server Vacations	397
9.4.2	Polling and Token-passing Models	402
9.5	Rate Conservation Law (RCL)	404
9.6	Discussion and Further Reading	409
10	Phase-type Process and Matrix-geometric Method	410
10.1	Introduction	410
10.2	Phase-type (PH) Distribution	411
10.2.1	Kronecker Product and Kronecker Sum	415
10.3	Phase-Type (PH) Renewal Process	417
10.4	PH Renewal Service Process	424
10.5	PH/PH/1 Queue and Quasi-Birth-and-Death Process	427
10.6	Stationary Distribution of QBD and Performance Measures	431
10.6.1	Stability Condition	434
10.7	Algorithms for the Rate Matrix \mathbf{R}	435
10.8	Discussion and Further Reading	438
11	Discrete-time Queues	441
11.1	Geo/G/1	441
11.1.1	Geo ^(B) /G/1: Geo/G/1 with Bulk Arrivals	443
11.1.2	Geo/G/1 with Processor Sharing	444
11.1.3	Application of Geo/G/1 to Statistical TDM	446
11.2	Geo/Geo/1: Discrete-time M/M/1	449
11.3	Discrete-time M/G/∞	453
11.4	Discrete-time G/G/1 System	454
11.5	Discussion and Further Reading	455
12	Traffic Modeling	457
12.1	Introduction	457
12.2	Second-Order Properties of Arrival Processes	458
12.2.1	Indices of Dispersion: IDI and IDC	458
12.2.2	Peakedness	460
12.2.3	Modified Peakedness	463
12.3	Markovian Traffic Models	466
12.3.1	On–Off Source Model	466
12.3.2	Markovian Arrival Process (MAP)	469

12.3.3	Markov Modulated Poisson Process (MMPP)	471
12.3.4	MMPP/G/1 Model	473
12.3.5	Batch Markov Arrival Process (BMAP)	477
12.4	Long-Range Dependent (LRD) Traffic Models	478
12.4.1	Long-term Trends and Hurst Parameter	479
12.4.2	Long-range Dependent and Self-similar Processes	482
12.4.3	Fractional Brownian Motion Traffic	487
12.5	Discussion and Further Reading	490
13	Fluid Models	492
13.1	Fluid Approximation	492
13.2	Markov Fluid Model of Statistical Multiplexer	493
13.2.1	Equilibrium-state Solutions	496
13.3	Two Limiting Cases: Buffer Overflows and Infinite Sources	507
13.3.1	Buffer Overflow	507
13.3.2	Infinite Source Model	508
13.4	Markov Fluid Model with Multiple Types of Sources	511
13.5	Discussion and Further Reading	515
14	Approximation and Bounding Techniques	517
14.1	Inequalities and Bounds	517
14.1.1	Markov's Inequality and Kolmogorov's Inequality	517
14.1.2	Chernoff's Bound	519
14.1.3	Large Deviation Theory	521
14.2	Exponential Bounds On G/G/1 Waiting Time Distribution	525
14.3	Bounds for Mean Waiting Time in G/G/1	538
14.4	Bounds for Mean Waiting Time in G/G/m	540
14.5	Heavy-Traffic Approximation for G/G/1 Queue	542
14.6	Diffusion Process Approximations for G/G/1 Queue	544
14.6.1	Approximation of the Queue Process	544
14.6.2	Equilibrium-state Solution	548
14.6.3	Time-dependent Solution	551
14.6.4	General Diffusion Equation	553
14.7	Diffusion Approximation of a Queueing Network	558
14.7.1	Diffusion Approximation in an Open Network	558
14.7.2	Diffusion Approximation in a Closed Network	561
14.8	Bounds and Approximations for Bandwidth Allocation	564
14.8.1	Minimum Envelope Process and Rate	564
14.8.2	Effective Bandwidth of Fluid Source	565
14.8.3	Effective Bandwidth of Markov Modulated Fluid Source	571
14.9	Discussion and Further Reading	575
15	Time-dependent Solutions of Queues	576
15.1	Time-dependent Solution for M/G/∞	576
15.1.1	Discrete-time M/G/∞	578

15.2	Time-dependent Solution of a Markov Process Model	578
15.2.1	Matrix Representation	578
15.2.2	Spectral Expansion	581
15.3	Eigenvectors of the BD Process	584
15.3.1	Laplace Transform Method	585
15.4	Generating Function Method	591
15.4.1	Time-dependent Solution of M/M/1	591
15.4.2	Time-dependent Solution of Multiple On-Off Source	600
15.5	Uniformization of Continuous-time Markov Chain	607
15.5.1	Time-dependent Solution of the M/M/1 Queue	609
15.6	Time-dependent Solution of a Fluid Flow Model	613
15.6.1	Laplace Transform Method	614
15.7	Discussion and Further Reading	618
 IV Simulation Modeling and Analysis		619
16	Formulation and Implementation of Simulation Models	621
16.1	Introduction	621
16.2	Self-Driven Simulation vs. Trace-Driven Simulation	622
16.3	Formulation of Simulation Models	625
16.3.1	Time Control in Simulators	625
16.3.2	Data Structures in Simulators	626
16.4	Implementation of Simulators	630
16.4.1	Simulation Languages	630
16.4.2	Simulation Platforms	632
16.4.3	OPNET Simulator	634
16.4.4	Network Simulator-2 (ns-2)	637
16.5	Techniques for Generating Random Variables	638
16.5.1	Random number Generator (RNG)	638
16.5.2	Multiplicative Congruential Method	640
16.5.3	Mixed Congruential Method	642
16.5.4	Other Methods for Random Number Generation	645
16.5.5	Generation of Variates from General Distributions	648
16.5.6	Generation of Normal or Gaussian Variates	652
16.5.7	Improved Methods to Generate General Variates	655
16.6	Case Studies	659
16.6.1	Simulation of a Single-server Queueing System	659
16.6.2	Simulation of a Queueing Network	672
16.6.3	Simulation of a Loss Network	676
16.7	Discussion and Further Reading	682
17	Simulation Experiments and Statistical Data Analysis	684
17.1	Introduction	684
17.2	Experiments and Statistical Inference	684
17.2.1	Controlled Variables and Uncontrolled Variables	685

xviii Contents

17.2.2	Significance Level and Confidence Intervals	686
17.3	Analyzing a Simulation Run	689
17.3.1	Starting Conditions, Transients, and Equilibrium	689
17.3.2	Choice of Sample Sizes	690
17.3.3	Comparison of Alternatives	700
17.3.4	Validation and Testing of Simulation Models	701
17.4	Efficient Statistical Simulation	702
17.4.1	Monte Carlo Techniques	702
17.4.2	Antithetic-variate Method	703
17.4.3	Control-Variate Method	705
17.5	Fast Simulation of Rare Events	708
17.5.1	Importance Sampling Method	710
17.5.2	Importance Sampling in Simulating a Queueing System	718
17.6	Discussion and Further Reading	722
A	Number Theory	724
	Bibliography	730
	Index	761