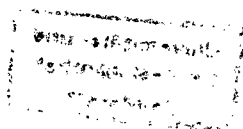


Richard J. Boucherie • Nico M. van Dijk
Editors

Queueing Networks

A Fundamental Approach



 Springer

Contents

1	On Practical Product Form Characterizations	1
	Nico M. van Dijk	
	A: Product Forms: Single Station Hierarchy	2
1.1	Introduction	2
1.2	Product Forms: Three Balances	4
1.2.1	Station Balance: B-D or Erlang-Engset systems	4
1.2.2	Class balance: Coordinate convex property (CCP)	6
1.2.3	Job Local Balance: Necessity	15
1.2.4	LCFS-pre case: Nonexponential	19
1.2.5	Symmetric Disciplines and Job-Local-balance (JLB)	22
1.3	Invariant Disciplines and JLB	25
1.3.1	Invariance Condition	25
1.3.2	Service invariant examples	28
1.3.3	A generalized symmetric insensitivity result	32
1.4	An application, literature discussion and hierarchy review	35
1.4.1	An $M G c c+m$ application	35
1.4.2	Literature discussion	37
1.4.3	A hierarchy review	39
	B: Product Forms: Tandem and Cluster Structures	41
1.5	Tandem Queues	41
1.5.1	Introduction	41
1.5.2	Product Form Tandem Queues	45
1.5.3	Service examples	49
1.5.4	Blocking examples	51
1.5.5	Mixed examples	54
1.6	Jacksonian clusters	56
1.6.1	A Jackson cluster	56
1.6.2	A restricted Jackson cluster	58
1.6.3	A conservative product form protocol	60
1.7	Product form bounds for networks of restricted clusters	62
1.7.1	Instructive tandem extension	63

1.7.2	A Jackson Tandem	66
1.7.3	A nested case	68
1.7.4	Further illustrative examples	69
1.7.5	An Optimal Design Application	73
1.8	A hospital application	73
1.8.1	Motivation	73
1.8.2	Model formulation	74
1.8.3	Bounds and application	76
1.9	Evaluation	77
1.9.1	Literature	77
1.9.2	Review Part B	79
1.9.3	Some remaining questions	80
	References	81
2	Order Independent Queues	85
	A.E. Krzesinski	
2.1	Introduction	85
2.2	The OI Queue	87
2.2.1	The Definition of an OI Queue	88
2.2.2	The Implications of the OI Conditions	88
2.2.3	The Stationary Distribution	89
2.2.4	Models Covered by the OI Class	92
2.3	Numerical Techniques for the OI Queue	95
2.3.1	Aggregating the State Space	95
2.3.2	The Performance Measures: the MSCCC Queue	96
2.4	The OI Loss Queue	102
2.4.1	The Stationary Distribution	103
2.4.2	The Performance Measures: the MSCCC Loss Queue ...	105
2.4.3	OI Loss Networks	109
2.5	OI Applications	110
2.5.1	Multiported Memory	110
2.5.2	A Messaging Card	110
2.5.3	Multilayer Window Flow Control	111
2.5.4	Machine Scheduling Model	111
2.5.5	Blocked Calls Cleared	111
2.5.6	Blocked Calls Queued	112
2.5.7	Blocked Calls Queued with Source Rejection	113
2.5.8	Local and Long Distance Calls	113
2.5.9	Local and Transit Calls	114
2.5.10	Hierarchical Tree Networks	115
2.5.11	Local and External Networks	115
2.5.12	Transit Calls among Networks	116
2.6	An Algorithm to Compute the Performance Measures of the MSCCC	118
	References	119

3	Insensitivity in Stochastic Models	121
	P.G. Taylor	
3.1	Introduction	121
3.2	The Erlang Loss System as a Symmetric Queue	126
3.3	The Erlang Loss System as a GSMP	129
3.4	Insensitive Queueing Networks	132
3.5	Non-Standard Insensitive Models	136
3.6	Conclusion	137
	References	138
4	Palm Calculus, Reallocatable GSMP and Insensitivity Structure	141
	Masakiyo Miyazawa	
4.1	Introduction	141
4.2	Shift operator group	142
4.3	Point processes	146
4.4	Palm distribution	149
4.5	Inversion formula	153
4.6	Detailed Palm distribution	157
4.7	Time and event averages	159
4.8	Rate conservation law	163
4.9	PASTA: a proof by the rate conservation law	167
4.10	Relationship among the queueing length processes observed at different points in time	170
4.11	An extension of the rate conservation law	173
4.12	Piece-wise deterministic Markov process (PDMP)	176
4.13	Exponentially distributed lifetime	182
4.14	GSMP and RGSMP	185
4.15	Exponential and non-exponential clocks in RGSMP	189
4.16	Product form decomposability	193
4.17	Applications to queues and their networks	201
4.18	Further insensitivity structure in RGSMP	205
4.19	Bibliographic notes	213
	References	214
5	Networks with Customers, Signals, and Product Form Solutions	217
	Xiuli Chao	
5.1	Introduction	217
5.2	Quasi-Reversibility of Queues	219
5.3	Quasi-Reversibility of Queues with Triggered Departures	224
5.4	Networks of Quasi-Reversible Nodes	226
5.5	Networks with Signals and Triggered Movements	235
5.6	Networks with Positive and Negative Signals	241
5.6.1	Single Class of Positive and Negative Signals	242
5.6.2	Multiple Classes of Positive and Negative Signals	246
5.7	Necessary and Sufficient Conditions for Product Form	250

5.8	Quasi-Reversibility Revisited	256
5.9	Networks with Random Customer Shuffling	260
5.10	Conclusion	264
	References	265
6	Discrete Time Networks with Product Form Steady States	269
	Hans Daduna	
6.1	Introduction	269
6.2	Bernoulli Servers with Different Customer Types	271
6.3	Closed Cycles of Bernoulli Servers	275
6.3.1	Steady State Behaviour and Arrival Theorem	275
6.3.2	Delay Times in a Closed Cycle	279
6.3.3	Computational Algorithms for Closed Cycles	281
6.3.4	Large Cycles	282
6.4	Open Tandems of Bernoulli Servers	284
6.4.1	Steady State and Arrival Theorem	284
6.4.2	Delay Times in Tandems	287
6.4.3	Open Tandems of Unreliable Bernoulli Servers	288
6.5	Doubly stochastic and geometrical servers	289
6.5.1	Common Properties	290
6.5.2	The Doubly Stochastic Server	290
6.5.3	The Geometrical Server	295
6.5.4	Networks of Doubly Stochastic and Geometrical Nodes	296
6.6	Batch Movement Networks	300
6.6.1	The General Network Model	300
6.6.2	Walrand's S-Queues and Networks	304
6.6.3	Closed Networks of Unreliable S-Queues	305
6.6.4	Networks with Triggered Batch Movements	307
	References	308
7	Decomposition and Aggregation in Queueing Networks	313
	Tijs Huisman and Richard J. Boucherie	
7.1	Introduction	313
7.2	Model	317
7.2.1	The nodes	317
7.2.2	Interaction between the nodes	320
7.2.3	The network	322
7.3	Decomposition	324
7.4	Aggregation	330
7.5	Examples	335
7.5.1	Quasi-reversible nodes linked via state-dependent routing	335
7.5.2	Biased local balance	337
7.5.3	A pull network	339
7.5.4	An assembly network	340
	References	343

8	Stochastic Comparison of Queueing Networks	345
	Ryszard Szekli	
8.1	Introduction	345
8.1.1	Jackson networks	349
8.1.2	Gordon-Newell networks	350
8.1.3	Ergodicity of classical networks	350
8.2	Stochastic monotonicity and related properties for classical networks	352
8.2.1	Stochastic orders and monotonicity	352
8.2.2	Stochastic monotonicity and networks	356
8.2.3	Bounds in transient state	357
8.2.4	Bounds in stationary state	357
8.2.5	Sojourn times in networks	359
8.3	Properties of throughput in classical networks	361
8.3.1	Uniform conditional variability ordering, a relation between closed and open networks	361
8.3.2	Effect of enlarging service rates in closed networks	362
8.3.3	Majorization, arrangement and proportional service rates	363
8.3.4	Throughput and number of jobs	365
8.4	Routing and correlations	365
8.4.1	Correlation inequalities via generators	367
8.4.2	Doubly stochastic routing	373
8.4.3	Robin-Hood transforms	374
8.4.4	Dependence orderings and monotonicity	376
8.5	Jackson networks with breakdowns	381
8.5.1	Product formula	381
8.5.2	Bounds via dependence ordering for networks with breakdowns	383
8.6	General networks	384
8.6.1	Dependence and variability in input	387
8.6.2	Comparison of workloads	387
8.6.3	Throughput in general networks	390
	References	391
9	Error Bounds and Comparison Results: The Markov Reward Approach For Queueing Networks	397
	Nico M. Van Dijk	
A:	General results	398
9.1	Motivation	398
9.1.1	A first example	398
9.1.2	Two more examples	402
9.1.3	Objectives	405
9.1.4	Approach	405
9.1.5	Outline	406
9.2	Stochastic Comparison	407

9.2.1	Preliminaries	407
9.2.2	Stochastic comparison	408
9.2.3	Stochastic comparison failure	411
9.3	Markov reward approach	413
9.3.1	Preliminaries	413
9.3.2	Comparison Result	414
9.3.3	Error bound Result	416
9.3.4	Truncation Error Bound	420
9.3.5	Comparison of MRA and SC	421
B: Applications		425
9.4	Application 1: Instructive Breakdown Example	425
9.4.1	Analytic bounds for the bias-terms	425
9.4.2	Comparison Result	429
9.4.3	Error Bounds	431
9.5	Application 2: Finite Tandem Queue	433
9.5.1	Problem Motivation	433
9.5.2	Comparison Result (Bounds)	435
9.5.3	Technical verification of Bias-Terms	436
9.6	Application 3: Truncation of Finite Jackson Network	440
9.6.1	Description and motivation.	441
9.6.2	Truncation	442
9.6.3	Analytic Error bound	447
9.6.4	Application: Cellular Mobile Network Application	450
9.7	Evaluation	451
9.7.1	Extensions	452
9.7.2	Further Research	453
9.7.3	Other applications.	454
References		456
10	Stability of Join-the-Shortest-Queue networks: Analysis by Fluid Limits	461
	J. G. Dai, John J. Hasenbein and Bara Kim	
10.1	Join-the-shortest-queue networks	461
10.2	The network process and the fluid model	464
10.3	JSQ networks with two stations	469
10.4	Two examples with three stations	474
10.5	JSQ networks with homogeneous feedback	482
10.6	Further study	486
References		486
11	Methods in Diffusion Approximation for Multi-Server Systems: Sandwich, Uniform Attraction and State-Space Collapse	489
	Hong Chen and Heng-Qing Ye	
11.1	Introduction	489
11.2	Preliminaries	492
11.3	Multi-Server Queue: Sandwich Method	494

11.4	A Multi-Class Queue under FIFO Service Discipline: Uniform Attraction and State-Space Collapse	498
11.4.1	Fluid Approximation and Uniform Attraction	500
11.4.2	Diffusion Approximation	503
11.5	Multi-Channel Queues under JSQ Routing Control	509
11.5.1	Fluid Approximation and Uniform Attraction	511
11.5.2	Diffusion Approximation	514
11.6	Notes	518
11.7	Appendix	519
11.7.1	Proof of Lemma 11.5.2	519
11.7.2	Proof of Proposition 11.5.3	521
11.7.3	Proof of Lemma 11.5.6	523
	References	529
12	Queueing Networks with Gaussian Inputs	531
	Michel Mandjes	
12.1	Introduction	531
12.2	Preliminaries on Gaussian processes	533
12.2.1	Gaussian sources	533
12.2.2	Classifications	534
12.2.3	Schilder's theorem	536
12.3	Single queues	539
12.3.1	Steady-state queue length	539
12.3.2	Logarithmic asymptotics	540
12.3.3	The shape of the loss curve	541
12.3.4	The buffer-bandwidth curve is convex	545
12.4	Tandem networks	547
12.4.1	Alternative formulation	547
12.4.2	Lower bound	550
12.4.3	Tightness; two regimes	553
12.4.4	Approximation	554
12.5	Priority queues	555
12.5.1	Lower bound	556
12.5.2	Tightness; two regimes	557
12.5.3	Approximation	558
12.6	Concluding remarks	559
	References	559
13	Mean Values Techniques	561
	Ivo Adan, Jan van der Wal	
13.1	Introduction	561
13.2	PASTA property and Little's law	562
13.3	MVA for single-station systems	563
13.3.1	$M M 1$	563
13.3.2	$M G 1$	564
13.3.3	$M G 1$ with priorities	564

13.3.4	$M G 1$ with least attained service	565
13.3.5	Server vacations	567
13.3.6	$M M c$	567
13.3.7	$M M c$ with priorities	568
13.3.8	Retrials	569
13.3.9	Polling	569
13.4	AMVA for single-station systems	571
13.4.1	$M G c$	571
13.4.2	$M G c$ with priorities	572
13.5	ASTA property in PF networks	572
13.5.1	Open single-class PF networks	572
13.5.2	Open multi-class PF networks	573
13.5.3	Closed multi-class PF networks	574
13.6	MVA for open PF networks	576
13.7	MVA for closed single-class PF networks	576
13.8	MVA for closed multi-class PF networks	579
13.9	AMVA for open networks	580
13.10	AMVA for closed single-server networks	580
13.10.1	Class-dependent general service times	581
13.10.2	Priorities	582
13.10.3	Multiple visits to a station	582
13.11	AMVA for closed multi-server networks	583
13.12	The Schweitzer-Bard approximation	583
	References	585
14	Response Time Distributions in Networks of Queues	587
	Michael Grottke, Varsha Apte, Kishor S. Trivedi, and Steve Woollet	
14.1	Introduction	587
14.2	Closed Markovian networks	589
14.2.1	Tagged customer approach	590
14.2.2	Example: Central server model	590
14.3	Open Markovian networks of $M/M/c/b \leq \infty$ queues	594
14.3.1	Response time blocks	595
14.3.2	Building the Markov chain from a queuing network	598
14.3.3	Examples	601
14.4	Open Markovian networks of queues with general PH service time distributions	611
14.4.1	Building blocks with general PH service time distributions	611
14.4.2	Building the Markov chain	612
14.4.3	Example: CPU and disk queuing system	614
14.5	Non-Markovian networks	615
14.5.1	Approximating non-PH distributions	616
14.5.2	Modeling response time distributions using semi-Markov processes	624
14.6	Conclusions	637

References	639
15 Decomposition-Based Queueing Network Analysis with FiFiQueues .	643
Ramin Sadre, Boudewijn R. Haverkort	
15.1 Introduction	643
15.2 The decomposition approach	645
15.3 Whitt's Queueing Network Analyzer	648
15.3.1 Model class	649
15.3.2 Traffic descriptors	649
15.3.3 Superposition of traffic streams	649
15.3.4 Splitting traffic streams	650
15.3.5 Servicing jobs	651
15.3.6 Node performance	651
15.3.7 Network-wide performance	652
15.3.8 Complexity	652
15.4 FiFiQueues	653
15.4.1 Model class	653
15.4.2 Traffic descriptor	654
15.4.3 Superposition of traffic streams	654
15.4.4 Splitting traffic streams	655
15.4.5 Servicing jobs	655
15.4.6 Node performance	659
15.4.7 Network-wide performance	660
15.4.8 Complexity	661
15.4.9 The FiFiQueues network designer	662
15.5 Performance of FiFiQueues	664
15.5.1 Evaluation of FiFiQueues	665
15.5.2 Performance evaluation of a web server	668
15.5.3 Summary	676
15.6 Summary and conclusions	676
15.7 Appendix: Jackson queueing networks	677
15.7.1 Model class	677
15.7.2 Traffic descriptor	677
15.7.3 Superposition of traffic streams	678
15.7.4 Splitting traffic streams	678
15.7.5 Servicing jobs	678
15.7.6 Node performance	678
15.7.7 Network-wide performance	679
15.7.8 Complexity	679
15.8 Appendix: MAPs, PH-distributions and QBDs	680
15.8.1 Markovian Arrival Processes (MAPs)	681
15.8.2 Phase-type (PH) renewal processes	683
15.8.3 Infinite QBDs	684
15.8.4 Finite QBDs	686
15.9 Appendix: Existence of the fixed point	688

15.9.1	Notation and Brouwer's theorem	688
15.9.2	Properties of D	688
15.9.3	Continuity of H	689
15.9.4	Continuity for $c_a^2 = 1$	690
15.9.5	Continuity for $c_a^2 = \frac{1}{m}, m \in \{2, \dots, 10\}$	693
	References	697
16	Loss Networks	701
	Stan Zachary and Ilze Ziedins	
16.1	Introduction	701
16.2	Uncontrolled loss networks: stationary behaviour	705
16.2.1	The stationary distribution	705
16.2.2	The single resource case	706
16.2.3	The Kaufman-Dziong-Roberts (KDR) recursion	707
16.2.4	Approximations for large networks	708
16.3	Controlled loss networks: stationary behaviour	711
16.3.1	Single resource networks	711
16.3.2	Multiple resource models	713
16.4	Dynamical behaviour and stability	715
16.4.1	Fluid limits for large capacity networks	715
16.4.2	Single resource networks	717
16.4.3	Multi-resource networks: the uncontrolled case	718
16.4.4	Multi-resource networks: the general case	719
16.4.5	The diverse routing limit	723
16.5	Further developments and open questions	724
	References	726
17	A Queueing Analysis of Data Networks	729
	Thomas Bonald and Alexandre Proutière	
17.1	Introduction	729
17.2	Capacity region	730
17.3	Traffic characteristics	737
17.4	Stability issues	738
17.5	Flow throughput	739
17.6	Queueing analysis	740
17.7	Resource allocation	742
17.8	Insensitivity results	745
17.9	A single link	747
17.10	Performance bounds	751
17.11	Examples	756
17.12	Open issues	762
17.13	Bibliographical notes	762
	Appendix	763
	References	764

18 Modeling a Hospital Queueing Network 767
Stefan Creemers and Marc Lambrecht

18.1 Introduction 767

18.2 Problem Description 769

18.3 A hospital queueing system 771

 18.3.1 Modeling arrival rate and natural service times 773

 18.3.2 Variability from preemptive and nonpreemptive outages . 777

18.4 Applications 786

 18.4.1 Numerical example 786

 18.4.2 The impact of interrupts 788

 18.4.3 The impact of pooling 789

 18.4.4 Finding the optimal number of patients in a service
 session 791

18.5 Conclusion 796

References 796