Performance Analysis of Complex Networks and Systems

PIET VAN MIEGHEM
Delft University of Technology

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107058606

First published 2014

Printed in the United Kingdom by CPI Group Ltd, Croydon CR0 4YY

Library of Congress Cataloging-in-Publication Data

Van Mieghem, Piet.
Performance analysis of complex networks and systems / Piet Van Mieghem, Delft University of Technology.
pages cm
ISBN 978-1-107-05860-6 (hardback)
1. System analysis–Mathematics. 2. Computational complexity. 3. Network analysis (Planning)
I. Title.
T57.V36 2014
003.01′5192–dc23 2014000241

ISBN 978-1-107-05860-6 Hardback

Waar een wil is, is een weg.
to my father

in memory of
my wife Saskia and my son Nathan

to my sons Vincent and Laurens

Cambridge University Press
978-1-107-05860-6 - Performance Analysis of Complex Networks and Systems
Piet Van Mieghem
Frontmatter
More information

Contents

Preface

This book is a development of *Performance Analysis of Communications Networks and Systems* of 2006. Its current incarnation has a broader scope, extending to complex networks, the broad term covering all types of real-world networks, that encompass communications networks. Apart from the correction of numerous errors in the earlier book, nearly all chapters have been extended. Chapter 17 on Epidemics in networks has been added, while the appendix on algebraic graph theory has been deleted, because our book *Graph Spectra for Complex Networks* amply replaces this chapter. Similar to *Graph Spectra for Complex Networks*, **art. x** has been used to refer to article x in Appendix A. The number of problems, together with their solutions in Appendix B, has been doubled at least.

Performance analysis belongs to the domain of applied mathematics. In particular, the branches of mathematics as probability theory, stochastic processes and graph theory are exploited, besides analysis (calculus) and linear algebra that are omnipresent nearly everywhere. The major aim of this book is to offer several mathematical methods to address challenges in *network science*, the rapidly growing field of complex networking. The link with technology is kept shallow, on purpose, because most technical advances in micro-electronics, in communications protocols and services, standards, etc. have a more limited lifetime and a narrower scope compared to mathematical concepts.

This book aims to present methods rigorously, hence mathematically, with minimal resorting to intuition. It is my belief that intuition is often gained after the result is known and rarely before the problem is solved, unless the problem is simple. I have tried to interpret most of the important formulas in the sense of "What does this mathematical expression teach me?" This last step justifies the word "applied", since most mathematical treatises do not interpret as it contains the risk to be imprecise and incomplete.

As prerequisites, familiarity with elementary probability and the knowledge of the theory of functions of a complex variable are assumed. In particular, the beautiful book on the *Theory of Functions* by Titchmarsh (1964) is recommended for complex function theory. Appendix A briefly summarizes concepts of linear algebra used in this book. Parts in the text in small font refer to more advanced topics or to computations that can be skipped at first reading.

Chapters 2–3 in Part I briefly review probability theory; they are included to make the remainder self-contained. Chapter 4 discusses how to compute correlation between several random variables. The central role of the Gaussian distribution is emphasized. Since the Gaussian distribution is so widely known, I have included Gauss's own derivation of his distribution, whereas the Central Limit Theorem (studied in Chapter 6 together with other limit laws) is currently considered as the main law that produces a Gaussian distribution. Chapter 5 treats powerful inequalities at an introductory level, but skips the recent (and more difficult) inequalities, such as the FKG inequality due to Fortuin, Kasteleyn and Ginibre.

The book essentially starts with Chapter 7 (Part II) on Poisson processes. The Poisson process (independent increments and discontinuous sample paths) and Brownian motion (independent increments but continuous sample paths) are considered to be the most important basic stochastic processes. We briefly touch upon renewal theory to move to Markov processes. The theory of Markov processes is regarded as a fundament for many applications, particularly in queueing theory, epidemics on networks and some instances of shortest path routing. A large part of the book is consumed by Markov processes and its applications. The last chapters of Part II dive into queueing theory. Inspired by intriguing problems in telephony at the beginning of the twentieth century, Erlang has pushed queueing theory to the scene of sciences. Since his investigations, queueing theory has grown considerably. Especially during the last decade with the advent of the Asynchronous Transfer Mode (ATM) and the worldwide Internet, many early ideas have been refined (e.g. discrete-time queueing theory, large deviation theory, scheduling control of prioritized flows of packets) and new concepts (self-similar or fractal processes) have been proposed. Queuing theory is expected to emerge again in complex networks, when structural aspects are understood.

Part III covers the parts of network science that study the structure of and processes on complex networks. While the first decade of this century has predominantly focused on the topology and structure of complex networks, the next decade aims to understand the influence of the topology on the dynamic process(es) on the network. Chapter 15 overviews the topological properties of complex networks and invokes concepts of graph theory, as well as random graph theory. Both the famous Erdős-Rényi and the Barabási-Albert random graphs are introduced. The giant component of a network, which is the largest still connected subgraph, is studied, because the giant component can be considered as the operational heart of the complex network. Finally, interdependent networks are introduced and the surprisingly different nature of cascading failures between single and interdependent networks is discussed. Chapter 16 and 17 exemplify two dynamic processes on a network: the transport along the shortest path and the spread of epidemics in networks. Chapter 18 and 19 dive deeper into routing instances as multicasting and anycasting, originally proposed for internetworking, but extendable to the newer communications types as social networking and cloud computing.

Since network science is still developing at fast pace, Part III is undoubtedly the least mature and complete. Moreover, I have predominantly relied on my own research for the exposition of topics, without the ambition to thoroughly review (and cite) contributions in the field. The book of Remco van der Hofstad (2013) treats random graphs differently inspired by Erdős' probabilistic method. Remco seriously extends the material in Chapter 15 and I recommend his book for a deeper discussion about the giant component, phase transitions, inhomogeneous random graphs, the configuration model and preferential attachment models.

To Huijuan Wang, who has used the earlier book of 2006 in her classes, I am indebted for many suggestions, corrections and a shorter proof of the degree distribution in the URT in Section 16.7.2. Numerous people have pointed me to errors in the earlier 2006 book, while others gave suggestions for this book. I am very grateful to all of them: Chandrashekhar Pataguppe Suryanarayan Bhat, Ruud van de Bovenkamp, Eric Cator, Li Cong, Edwin van Dam, Michel Dekking, David Hemsley, Remco van der Hofstad, Gerard Hooghiemstra, David Hunter, Geurt Jongbloed, Merkouris Karaliopoulos, Rob Kooij, Javier Martin Hernandez, Jil Meier, Raphi Rom, Annalisa Socievole, Bart Steyaert, Siyu Tang, Stojan Trajanovski, Matthias Waehlisch, Huijuan Wang, and to the many students at Delft University of Technology that followed my course on Performance Analysis over the last 14 years.

Although this book is intended to be of practical use, in the course of writing it, I became more and more persuaded that mathematical rigor has ample virtues of its own.

Omnia sunt incerta, cum a
mathematicae discessum est

February 2014 Piet Van Mieghem

Symbols

Only when explicitly mentioned will we deviate from the standard notation and symbols outlined here.

Random variables and matrices are written with capital letters, while complex, real, integer, etc., variables are in lower case. For example, X refers to a random variable, A to a matrix, whereas x is a real number and z is complex number. Usually, i, j, k, l, m, n are integers. Operations on random variables are denoted by [.], whereas (.) is used for real or complex variables. A set of elements is embraced by {.}.

Linear algebra

A	$n \times m$ matrix $\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \\ a_{n1} & \ldots & a_{nm} \end{bmatrix}$
$\det A$	determinant of a square matrix A; also denoted by $\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$
$\text{trace}(A)$	$= \sum_{j=1}^{n} a_{jj}$: sum of diagonal elements of A
$\text{diag}(a_k)$	$= \text{diag}(a_1, a_2, \ldots, a_n)$: a diagonal matrix with diagonal elements listed, while all off-diagonal elements are zero
A^T	transpose of a matrix, the rows of A are the columns of A^T
A^*	matrix in which each element is the complex conjugate of the corresponding element in A
A^H	$= (A^*)^T$: Hermitian of matrix A
$c_A(x)$	$= \det(A - xI)$: characteristic polynomial of A
$\text{adj}A$	$= A^{-1} \det A$: adjugate of A
$Q(\lambda)$	$= \frac{c_A(\lambda)}{\lambda I - A}$: adjoint of A
e_j	basic vector, all components are zero, except for component j that is 1
δ_{kj}	Kronecker delta, $\delta_{kj} = 1$ if $k = j$, else $\delta_{kj} = 0$

Probability theory

$\Pr[X]$	probability of the event X
$E[X]$	$= \mu$: expectation of the random variable X
$\text{Var}[X]$	$= \sigma_X^2$: variance of the random variable X
$f_X(x)$	$= \frac{dF_X(x)}{dx}$: probability density function of X
$F_X(x)$	probability distribution function of X
$\varphi_X(z)$	probability generating function of X
	$\varphi_X(z) = E\left[z^X\right]$ when X is a discrete r.v.
	$\varphi_X(z) = E\left[e^{-zX}\right]$ when X is a continuous r.v.
$\{X_k\}_{1\leq k\leq m}$	$= \{X_1, X_2, \ldots, X_m\}$
$X_{(k)}$	k-th order statistics, k-th smallest value in the set $\{X_k\}_{1\leq k\leq m}$
P	transition probability matrix (Markov process)
$1_{\{x\}}$	indicator function: $1_{\{x\}} = 1$ if the event or condition $\{x\}$ is true,
	else $1_{\{x\}} = 0$. For example, $\delta_{kj} = 1_{\{k=j\}}$
γ	$= 0.577\,215\ldots$: Euler's constant
Ω	sample space
ω	sample point

Queuing theory

t_n	arrival time of the n-th packet
r_n	departure time of the n-th packet
$\tau_n = t_n - t_{n-1}$	n-th interarrival time
x_n	service time of n-th packet
w_n	waiting time of the n-th packet
$T_n = x_n + w_n$	system time of n-th packet
$v(t)$	virtual waiting time or unfinished work at time t
$\lambda = (E[\tau])^{-1}$	average arrival rate
$\mu = (E[x])^{-1}$	average service rate
$\rho = \frac{\lambda}{\mu}$	traffic intensity
$N_A(t)$	number of arrivals at time t
$N_S(t)$	number of packets in the system (queue plus server) at time t
$N_Q(t)$	number of packets in the queue at time t