Resource Management and Performance Analysis of Wireless Communication Networks

Shunfu Jin • Wuyi Yue

Resource Management and Performance Analysis of Wireless Communication Networks

Springer

Shunfu Jin
School of Information Science and Engineering
Yanshan University
Qinhuangdao
Hebei, China

Wuyi Yue
Department of Intelligence and Informatics
Konan University
Kobe, Japan

ISBN 978-981-15-7755-0 ISBN 978-981-15-7756-7 (eBook)
https://doi.org/10.1007/978-981-15-7756-7

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

This book provides analytical methods, approaches, algorithms, and simulation to evaluate numerically the quality of service and optimize the resource allocation in various wireless communication networks including broadband wireless access networks, cognitive radio networks, and cloud computing systems. The methods, techniques, and algorithms provided in this book include traffic analysis, applications of queueing theory and Markov chain theory, game theory, intelligent optimization, and operations research. In order to understand these methods, algorithms and techniques, basic concepts of computer communication networks and knowledge of queueing theory are needed. A familiarity with stochastic processes would also be useful.

Preface

Resource management techniques, primarily power, energy consumption minimization, and network design have become increasingly important in wireless communication networks (WCNs), since the explosion of demand for mobile devices. Specifically, the typical optimization problem in WCNs is to efficiently reduce the necessary energy consumption while maintaining the quality of service (QoS) of network users.

In the operation of WCNs, ensuring green wireless communication, managing resources, maintaining efficient energy production and conservation, while at the same time guaranteeing the best possible QoS are all key factors being implemented. Other important issues in resource management and energy conservation of WCNs are how to share the limited wireless resources and how to apply sleep mode technology.

Queueing theory and Markov chains are commonly used as powerful methods for analyzing performance and evaluating communication networks. There are several books on the topic of WCNs using queueing theory and Markov chains. However, although queueing theory and optimization techniques play vital roles in the deployment and operation of almost every type of network, none of the existing books covers the topics of resource management and energy conservation in WCNs. Resource management and energy conservation are the keys to producing successful WCNs in the future.

This book provides the fundamental concepts and principles underlying the study of queueing systems as they apply to resource management and energy conservation in modern WCNs. This book gives analytical methods, approaches, algorithms, and simulation to evaluate numerically the QoS and optimize the resource allocation in various WCNs including broadband wireless access (BWA) networks, cognitive radio networks (CRNs), and cloud computing systems.

This book explains the constructed stochastic models that are at the core of evaluating system performance and presents intelligent searching algorithms to optimize the strategy under consideration. This book also provides sufficient analytical methods, approaches, and numerical simulation for students, analysts,

managers, and industry people who are interested in using queueing theory to model congestion problems.

This book will be the first to provide an overview of the numerical analysis that can be gleaned by applying queueing theory, traffic theory, and other analytical methods. It will provide readers with information on recent advances in the resource management of various WCNs, such as BWA networks, CRNs, and cloud computing systems.

The authors of this book have been engaged in researching the performance evaluation of communication networks for nearly twenty years. This book has grown out of the authors' collaborative research on the performance evaluation for the resource management and energy conservation in WCNs, including BWA networks, CRNs, and cloud computing systems. The subject area discussed in the book is timely, gave the recent remarkable growth in wireless networking and the convergence of personal wireless communications, Internet technologies, and real-time multimedia. Each chapter of this book will give a detailed introduction of key topics in resource management of WCNs. The technical depth of the knowledge imparted in each chapter aims to satisfy experts in the field.

The methods, algorithms and techniques provided in this book include traffic analysis, applications of queueing theory and Markov chain theory, game theory, intelligent optimization, and operations research. In order to understand these methods, algorithms and techniques, basic concepts of computer communication networks and knowledge of queueing theory are needed. A familiarity with stochastic processes would also be useful.

We organize the book into three parts with a chapter on Introduction. In the Introduction, we briefly explain the background of our topic and give an overview of WCNs as they relate to the networks covered in this book. Then we introduce the queueing systems in general terms, as well as basic concepts and analysis methods that relate to the general theory of the stochastic processes used to capture the important properties and the stochastic behaviors of the network systems under research. We also show how to obtain the rate matrices, to solve the matrices and vectors numerically, and to optimize the system performances in queueing models and Markov chain models, effectively. These analytic methods, approximation methods, and techniques are used to analyze the system performance for resource management and energy conservation strategies on WCNs. Moreover, we define some of the important performance measures and common definitions to use in this book to present the performance analysis and optimization of the system models on WCN systems.

As an overview, we also outline the organization of the book. In particular, we explain the analysis methods for single- and multiple-vacation models, priority queueing systems, evaluation measures, and analytical methods and processes relating to performance optimization of the network systems.

In Parts I, II, and III, we present the performance evaluation and optimization for different WCNs, such as BWA networks, CRNs, and cloud computing systems, by applying queueing theory and the Markov chain models. Common performance measures like the energy conservation level, the energy-saving rate, the average

response time, the system throughput, the spectrum utilization, and the switching rate are used in the evaluation. The optimization methods used are the steepest descent method, Newton's method, and the Intelligent optimization algorithm.

Specifically, Part I discusses the sleep mode in BWA networks, including Worldwide Interoperability for Microwave Access (WiMAX), and Long-Term Evolution (LTE). Part I includes 7 chapters beginning with Chap. 2, looking at how under the sleep mode operation, an MS operates two modes: the awake mode and the sleep mode. Revolving around this standard, we investigate some vacation queueing models with two types of busy periods: busy period in the listening state and busy period in the awake state, with a sleep-delay period, with a wake-up procedure, and with batch arrivals.

Part II discusses the dynamic spectrum allocation and energy-saving strategy in CRNs. There are 7 chapters in Part II, beginning with Chap. 9. In these chapters, we present an analytic framework to evaluate the system performance by constructing priority queueing models with possible service interruptions, using multiple channels, with several types of vacation mechanisms, and possible transmission interruptions.

Part III discusses the virtual machine (VM) allocation and sleep mode in cloud computing systems aiming to realize green cloud computing. Part III includes 6 chapters. Beginning with Chap. 16, we illustrate how from the perspective of multiple servers, we have an insight into queueing models with task migrations, wake-up thresholds, variable service rates, partial vacations, and second optional services.

The system model and the performance analysis offered in each chapter of Parts I–III are independent of others offered in other chapters, although depending on the class of queueing system involved, there is some common ground between the techniques employed. Each chapter contains its own system and offers important analysis methods and numerical results. The readers or students will find it helpful to refer to Chap. 1 initially, but after that the remaining chapters are stand-alone units which can be read in any order.

We would like to thank all the publishers to grant permission for our original publications.

The authors also would like to thank the editorial and publishing staff of Springer Nature, in particular, Celine Lanlan Chang, executive editor, computer science, and Beracah John Martyn and Nick Zhu, production editors, for their support and cooperation.

Finally, the authors would like to thank National Natural Science Foundation (No. 61872311 and No. 61973261) and Hebei Province Natural Science Foundation (No. F2017203141), China, for supporting this publication. The authors are also grateful for GRANT-IN-AID FOR SCIENCE RESEARCH, MEXT, Japan which led to this book.

Qinhuangdao, China Shunfu Jin
Kobe, Japan Wuyi Yue
May 2020

Contents

Abbreviations

3GPP	3rd Generation Partnership Project
5G	5th Generation
BS	Base station
BWA	Broadband wireless access
CDC	Cloud data center
CRN	Cognitive radio network
CTMC	Continuous-time Markov chain
CDMA	Code-division multiple access
D-MAP	Discrete-time Markovian arrival process
DPM	Dynamic power management
DRX	Discontinuous reception
DTMC	Discrete-time Markov chain
DVFS	Dynamic voltage and frequency scaling
EAS	Early arrival system
eNodeB	Evolved node B
FCFS	First-come first-served
FDM	Frequency division multiplexing
IaaS	Infrastructure as a service
i.i.d.	Independent and identically distributed
LAS	Late arrival system
LTE	Long-Term Evolution
LTE-A	LTE-Advanced
MAC	Medium access control
M-ADRX	Multi-threshold Automated configuration DRX
MS	Mobile station
NRT-VR	Non-real-time variable rate
PaaS	Platform as a service
PGF	Probability generating function
PH	PHase type
PM	Physical Machine
PSO	Particle swarm optimization

PU	Primary user
QBD	Quasi-birth–death
QoE	Quality of experience
QoS	Quality of service
RT-VR	Real-time variable rate
SaaS	Software as a service
S-ADRX	Single-threshold Automated configuration DRX
SLA	Service level agreement
SNR	Signal-to-noise ratio
SOR	Successive over relaxation
SU	Secondary user
TLBO	Teaching-learning-based optimization
UE	User equipment
UGS	Unsolicited grant service
UMC	Underlying Markov chain
VIP	Very important person
VM	Virtual machine
WCN	Wireless communication network
WiMAX	Worldwide Interoperability for Microwave Access

List of Figures

List of Tables