Preface

This booklet collects the abstracts of the talks presented at the *Third Madrid Conference on Queueing Theory (MCQT'10)*, Hotel San Juan de Los Reyes, Toledo, Spain, June 28 - July 1, 2010. The Madrid Conference on Queueing Theory aims to be a meeting place where scientists and technicians in the field could find a discussion forum to promote research, encourage interaction and exchange of ideas. The conference is open to all trends in Queueing Theory including development of the theory, methodology, computational advances and applications.

This third edition is gathering together approximately 90 participants from 22 countries. The Opening Lectures of the conference will be given by M. Miyazawa (Tokyo University of Science) and E.A. van Doorn (University of Twente). Eight Invited Speakers will introduce and show the border lines of a variety of queueing topics. In addition, 60 submissions have been selected as Contributed Talks.

A selection of the best papers presented at the conference will be published in special volumes of the journals *Annals of Operations Research (AOR)* and *Top*. The papers will be refereed following the usual standards of all major international journals. AOR will devote one volume to high quality papers on the topic "Computational Methods and Applications in Queueing Theory" whereas Top will devote another volume to papers which make a significant contribution to the topic "Methodological Advances in Queueing Theory".

I would like to express my sincere gratitude to all participants. Thanks are also due to the institutions which supported the conference: Universidad Complutense de Madrid, Toledo Convention Bureau and the Applied Probability Society of INFORMS.

Jesús Artalejo

Contents

Organization

– Program Committee Members –

J.R. Artalejo (Chairman)

Department of Statistics and Operations Research
Faculty of Mathematics
Complutense University of Madrid
Madrid 28040, Spain
E-mail: jesus_artalejo@mat.ucm.es

T. Altiok (Rutgers University, USA)

F. Avram (University of Pau and Pays de l'Adour, France)

A. Economou (University of Athens, Greece)

O. Kella (The Hebrew University of Jerusalem, Israel)

M. Miyazawa (Tokyo University of Science, Japan)

S. Nishimura (Tokyo University of Science, Japan)

H. Perros (North Carolina State University, USA)

J.A.C. Resing (Eindhoven University of Technology, The Netherlands)

M.S. Squillante (IBM Watson Research Center, USA)

E.A. van Doorn (University of Twente, The Netherlands)

B. van Houdt (University of Antwerp, Belgium)

H. Zhang (Academia Sinica, China)

Y.Q. Zhao (Carleton University, Canada)

– Local Committee Members –

J.R. Artalejo (Chairman)

A. Gómez-Corral (Secretary)

M.J. López-Herrero (Secretary)

M.J. Alcón

J. Amador

I. Atencia

P. Moreno

General Information

Date

June 28 (Monday) – July 1 (Thursday), 2010

Venue

Hotel San Juan de Los Reyes
Toledo

Organized by

Complutense University of Madrid

Supported by

Complutense University of Madrid (UCM)
Applied Probability Society of INFORMS

Internet site

http://www.mat.ucm.es/~mcqt/confe10/conf10.html

Publication

A selection of the papers presented at the Conference will appear in special volumes of the journals *Annals of Operations Research* and *Top*.

Conference Information

Social Program

The social programme is open both to the participants and their accompanying registered persons. A personal card of invitation for each registered person is included among the conference documentation.

Toledo Trip

The "Toledo Trip" is scheduled for the afternoon of Tuesday, June 29, 2010. The tour will include the guided visit to some of the most emblematic monuments in Toledo: Cathedral, Church of Saint John of the Kings, Synagogue of Saint Mary the White.

- Date & Time: June 29 (Tuesday) 2010, 14:45-18:15
- Place: Main entrance of Hotel San Juan de Los Reyes. The departure time is 14:45. Your strict punctuality is kindly requested.

Conference Dinner

There will be a conference dinner on Wenesday June 30.

- Date & Time: June 30 (Wednesday) 2010, 20:30-22:30
- Place: Restaurant "*El Casón de Los López de Toledo*", Sillería 3, Toledo. http://www.casontoledo.com

A Map of the City

Plano: Gira, Estudio de Diseño

Program Schedule

June 28, 9:00 hours to 18:00 hours
Main Hall: Registration
Main Room: Welcome and Opening Session
Room 1: Sessions Mo-R1a, Mo-R1b
Room 2: Sessions Mo-R2a, Mo-R2b

9:00 – 10:00 Registration

Welcome – J.R. Artalejo

Opening Session – Chair: Y.Q. Zhao

10:00 – 11:00 Plenary Talk: **Light tail asymptotics in multidimensional reflected random walks and queueing networks**
Author: *M. Miyazawa* (Tokyo University of Science, Japan)

11:00 – 11:30 Coffee Break

Opening Session

11:30 – 12:30 Plenary Talk: **Rate of convergence to stationarity of the system** $M/M/N/N+R$
Author: *E.A. van Doorn* (University of Twente, The Netherlands)

12:30 – 15:00 Lunch Time

Session Mo-R1a – Chair: U. Yechiali

15:00 – 15:25 Contributed Talk: Approximations for the waiting-time distribution in polling systems with renewal arrivals
Authors: *J.L. Dorsman*, *R.D. van der Mei* and *E.M.M. Winands*

15:25 – 15:50 Contributed Talk: A polling model with smart customers
Authors: *M.A.A. Boon*, *A.C.C. van Wijk*, *I.J.B.F. Adan* and *O.J. Boxma*

15:50 – 16:15 Contributed Talk: Two-queue systems where customers of each queue are the servers of the other queue
Authors: *E. Perel* and *U. Yechiali*

Session Mo-R2a – Chair: I. Atencia

15:00 – 15:25 Contributed Talk: FCFS in dual-class discrete-time queueing systems
Authors: *S. De Clercq*, *K. Laevens*, *B. Steyaert* and *H. Bruneel*

15:25 – 15:50 Contributed Talk: Boundary probabilities in the discrete-time selective-repeat ARQ queueing problem
Authors: *S. De Vuyst*, *S. Wittevrongel* and *H. Bruneel*

15:50 – 16:15 Contributed Talk: A discrete-time queueing system with optional LCFS discipline
Authors: *I. Atencia* and *A.V. Pechinkin*

16:15 – 16:45 Coffee Break

Session Mo-R1b – Chair: G. Latouche

16:45 – 17:10 Contributed Talk: Asymptotic behavior for $MArP/PH/2$ queue with join the shortest queue discipline
Author: *Y. Sakuma*

17:10 – 17:35 Contributed Talk: A multi-server queue with Markovian arrivals and priority services
Authors: *S.R. Chakravarthy* and *H. Qing*

17:35 – 18:00 Contributed Talk: Queues with boundary assistance and the many effects of truncation
Authors: *G. Latouche*, *G.T. Nguyen* and *P.G. Taylor*

Session Mo-R2b – Chair: K. Laevens

16:45 – 17:10 Contributed Talk: Recursive equations for multiclass multi-server queueing systems
Author: *B. Rabta*

17:10 – 17:35 Contributed Talk: A multi-class M/PH/1 queue with deterministic impatience times
Authors: *J. Kim*, *B. Kim* and *J. Kim*

17:35 – 18:00 Contributed Talk: Queues with random order of service and multiple classes of customers
Authors: *K. Laevens*, *W. Rogiest*, *J. Walraevens* and *H. Bruneel*

June 29, 9:00 hours to 12:50 hours
Room 1: Sessions Tu-R1a, Tu-R1b
Room 2: Sessions Tu-R2a, Tu-R2b

Session Tu-R1a – Chair: E.A. Feinberg

9:00 – 9:50 Invited Talk: **Bandwidth allocation under delay percentiles**
Authors: *X. Mountrouidou*, *K. Kontovassilis*, *B. Anjum* and *H. Perros* (NC State University, USA)

9:50 – 10:15 Contributed Talk: Dynamic control of a single-server two-class retrial queueing system
Author: *A. Winkler*

10:15 – 10:40 Contributed Talk: Optimality of trunk reservation for an $M/M/k/N$ queue with several types of customers and holding costs
Authors: *E.A. Feinberg* and *F. Yang*

Session Tu-R2a – Chair: A.N. Dudin

9:00 – 9:50 Invited Talk: **Priority queuing and tree-structured Markov chains**
Authors: *B. van Houdt* (University of Antwerp, Belgium), *Q.-M. He* and *J. van Velthoven*

9:50 – 10:15 Contributed Talk: Drift conditions for fluid queues
Authors: *M. Govorun*, *G. Latouche* and *M.A. Remiche*

10:15 – 10:40 Contributed Talk: Multi-server queueing systems with broadcasting service disciplines, servers breakdowns and heating
Authors: *A. Dudin*, *S. Bin* and *C.S. Kim*

10:40 – 11:10 Coffee Break

Session Tu-R1b – Chair: I. Adan

11:10 – 11:35 Contributed Talk: Queueing system with state-dependent controlled batch arrivals and server under maintenance
Authors: *B. Krishna Kumar*, *S. Pavai Madheswari* and *S.R. Anantha Lakshmi*

11:35 – 12:00 Contributed Talk: A continuous-time queueing system with renovation
Authors: *I. Atencia*, *I. Fortes* and *S. Sánchez*

12:00 – 12:25 Contributed Talk: Analysis of an $M/M/k$ system with exponential setup times under staggered boot up
Authors: *A. Gandhi*, *M. Harchol-Balter* and *I. Adan*

12:25 – 12:50 Contributed Talk: Aggregate modelling of multi-processing workstations
Author: *I. Adan*

Session Tu-R2b – Chair: I. Mitrani

11:10 – 11:35 Contributed Talk: Approximation of $M/G/c/K$ retrial queue
Author: *Y.W. Shin*

11:35 – 12:00 Contributed Talk: Queueing model of the FSO/RF hybrid channel with heterogeneous links subject to failures
Authors: *D. Efrosinin* and *V. Rykov*

12:00 – 12:25 Contributed Talk: Numerical investigation of a $Geo/PH/1$ retrial queue with non-persistent customers
Author: *S.I. Rabia*

12:25 – 12:50 Contributed Talk: Managing performance and power consumption in a server farm
Author: *I. Mitrani*

12:50 – 14:45 Lunch Time

14:45 – 18:15 Toledo Trip

June 30, 9:00 hours to 18:00 hours
Room 1: Sessions We-R1a, We-R1b, We-R1c, We-R1d
Room 2: Sessions We-R2a, We-R2b, We-R2c, We-R2d

Session We-R1a – Chair: M. Telek

9:00 – 9:50 Invited Talk: **Exact tail asymptotics for random walks in the quarter plane**
Author: *Y.Q. Zhao* (Carleton University, Canada)

9:50 – 10:15 Contributed Talk: Tail behaviour of a finite-/infinite-capacity priority queue
Authors: *T. Demoor*, *J. Walraevens*, *F. Fiems* and *H. Bruneel*

10:15 – 10:40 Contributed Talk: Performance analysis of binary exponential backoff MAC protocol for cognitive radio in the IEEE 802.16e/m network
Authors: *S. Jin*, *J.S. Park* and *B.D. Choi*

10:40 – 11:05 Contributed Talk: On minimal representation of rational arrival processes
Authors: *P. Buchholz* and *M. Telek*

Session We-R2a – Chair: A. Krishnamoorthy

9:00 – 9:50 Invited Talk: **Waiting time approximation in multi-class queueing systems with multiple types of class-dependent interruptions**
Authors: *Ö.S. Ulusçu* and *T. Altıok* (Rutgers University, USA)

9:50 – 10:15 Contributed Talk: Equilibrium balking behavior in an unreliable queue with complete removals at failure epochs
Authors: *O. Boudali* and *A. Economou*

10:15 – 10:40 Contributed Talk: Machine repair problem with variable servers considering balking concept
Authors: *K.-H. Wang* and *Y.-C. Liou*

10:40 – 11:05 Contributed Talk: Queues with service interruption: a survey and some new results
Author: *A. Krishnamoorthy*

11:05 – 11:30 Coffee Break

Session We-R1b – Chair: B. D'Auria

11:30 – 11:55 Contributed Talk: Fluid limits and maximal load in multi-wavelength optical buffer systems
Authors: *K. De Turck*, *W. Rogiest*, *D. Fiems*, *H. Bruneel* and *S. Wittevrongel*

11:55 – 12:20 Contributed Talk: A fluid model analysis of streaming video in the presence of time-varying service capacity
Authors: *J.W.Bosman*, *R. Nunez Queija* and *R.D. van der Mei*

12:20 – 12:45 Contributed Talk: Markov modulated fluid queues with server vacations
Authors: *J.W. Baek*, *H.W. Lee*, *S.W. Lee*, *S. Ahn* and *L. Feng*

12:45 – 13:10 Contributed Talk: Brownian queues with modulated buffer
Authors: *B. D'Auria* and *O. Kella*

Session We-R2b – Chair: A. Pacheco

11:30 – 11:55 Contributed Talk: Influence of correlation in the arrival process on batch-service queueing systems
Authors: *D. Claeys*, *B. Steyaert*, *J. Walraevens*, *K. Laevens* and *H. Bruneel*

11:55 – 12:20 Contributed Talk: Computational analysis of queue-length distribution arising in the continuous-time queue: $GI/G/1$
Authors: *M.L. Chaudhry* and *X. Yang*

12:20 – 12:45 Contributed Talk: A simple and complete computational analysis of $M/G_j^{(m,M)}/1/\infty$ and $M/G_j^{(m,M)}/1/B$ queues using roots
Authors: *M.L. Chaudhry* and *J. Gai*

12:45 – 13:10 Contributed Talk: Customer loss probabilities in busy-periods of batch arrival oscillating $M/G/1/n$ queueing systems
Authors: *F. Ferreira*, *A. Pacheco* and *H. Ribeiro*

13:10 – 15:00 Lunch Time

Session We-R1c – Chair: J. Hasenbein

15:00 – 15:50 Invited Talk: **Stochastic analysis and optimization of queueing networks for cloud computing**
Authors: *S. Ghosh* and *M.S. Squillante* (IBM Thomas J. Watson Research Center, USA)

15:50 – 16:15 Contributed Talk: Fair connection admission control with adaptive thresholds for wireless multimedia networks
Authors: *T.O. Kim*, *A.S. Alfa* and *B.D. Choi*

16:15 – 16:40 Contributed Talk: Scheduling in tandem flexible server queueing systems
Authors: *J. Hasenbein* and *B. Kim*

Session We-R2c – Chair: M. Baykal-Gursoy

15:00 – 15:50 Invited Talk: **On first passage problems for piecewise deterministic processes with phase-type jumps**
Authors: *F. Avram* (Universite de Pau, France) and *X. de Lucas Araujo*

15:50 – 16:15 Contributed Talk: $M/D/1$, $M/D/\infty$ and stick breaking
Authors: *J.S.H. van Leeuwaarden*, *A.H. Lopker* and *A.J.E.M. Janssen*

16:15 – 16:40 Contributed Talk: Completion time analysis for infinite server queues with two service speeds
Authors: *M. Baykal-Gursoy* and *Z. Duan*

16:40 – 17:10 Coffee Break

Session We-R1d – Chair: P. Brill

17:10 – 17:35 Contributed Talk: Inference statistic for retrial queues with negative arrivals
Authors: *L. Berdjoudj* and *D. Aissani*

17:35 – 18:00 Contributed Talk: $M/G/c/c$ queues with workload-dependent arrival rate
Authors: *P. Brill* and *M.L. Huang*

Session We-R2d – Chair: A. Ridder

17:10 – 17:35 Contributed Talk: Universal generation of fractal statistics: data-traffic, queues, and physics
Authors: *I. Eliazar* and *J. Klafter*

17:35 – 18:00 Contributed Talk: Importance sampling for rare events in random walks and queueing models
Authors: *A. Ridder* and *T. Taimre*

20:30 – 22:30 Conference Dinner

July 1, 9:00 hours to 13:10 hours
Room 1: Sessions Th-R1a, Th-R1b
Room 2: Sessions Th-R2a, Th-R2b

Session Th-R1a – Chair: R. Nobel

9:00 – 9:50 Invited Talk: **The single server queue with synchronized repeated services**
Authors: *A. Economou* (University of Athens, Greece), *S. Kapodistria* and *J. Resing*

9:50 – 10:15 Contributed Talk: A single server queue with synchronized abandonments
Author: *S. Kapodistria*

10:15 – 10:40 Contributed Talk: Matrix continued fraction approach to level-dependent QBD process and its application to multiserver retrial queues
Authors: *T. Phung-Duc*, *H. Masuyama*, *S. Kasahara* and *Y. Takahashi*

10:40 – 11:05 Contributed Talk: A discrete-time retrial queueing model with abandonments
Author: *R. Nobel*

Session Th-R2a – Chair: Y. Sakuma

9:00 – 9:50 Invited Talk: **Polling tree networks**
Authors: *P. Beekhuizen*, *D. Denteneer* and *J. Resing* (Eindhoven University of Technology, The Netherlands)

9:50 – 10:15 Contributed Talk: Distribution of the successful and blocked events in a discrete-time retrial queue
Authors: *J. Amador* and *P. Moreno*

10:15 – 10:40 Contributed Talk: Analysis of $MAP/G^{(a,b)}/1/N$ queue with multiple vacations and closedown times
Authors: *A. Senthil Vadivu* and *R. Arumuganathan*

10:40 – 11:05 Contributed Talk: Join the shortest queue among similar c parallel service stations
Authors: *M. Kobayashi*, *Y. Sakuma* and *M. Miyazawa*

11:05 – 11:30 Coffee Break

Session Th-R1b – Chair: D. Fiems

11:30 – 11:55 Contributed Talk: Scaling limits of cyclically varying birth-death processes and applications
Authors: *M. Jonckheere*, *Y. Nazarathy* and *R. Nunez Queija*

11:55 – 12:20 Contributed Talk: Performance analysis of sleep mode operation in IEEE 802.16m advanced mobile WiMAX
Authors: *S. Baek*, *J.J. Son* and *B.D. Choi*

12:20 – 12:45 Contributed Talk: Stability analysis of a two-queue cascade network
Authors: *E. Morozov* and *B. Steyaert*

12:45 – 13:10 Contributed Talk: Light-correlation analysis of queues with autoregressive service times
Authors: *D. Fiems* and *B. Prabhu*

Session Th-R2b – Chair: R.D. van der Mei

11:30 – 11:55 Contributed Talk: Flexible server allocation and customer routing policies for two parallel queues when service rates are not additive
Authors: *H.-S. Ahn* and *M.E. Lewis*

11:55 – 12:20 Contributed Talk: Periodic and continuous review analytical models for airline check-in queues
Authors: *S. Moosa*, *M. Parlar* and *B. Rodrigues*

12:20 – 12:45 Contributed Talk: Optimal routing in parallel, non-observable queue and the price of anarchy revisited
Authors: *J. Anselmi* and *B. Gaujal*

12:45 – 13:10 Contributed Talk: Optimal resource allocation of time-reservation systems
Authors: *R. Yang*, *S. Bhulai* and *R.D. van der Mei*

13:10 – 14:15 Farewell Cocktail

Abstracts

In this section, abstracts are listed in alphabetical order of the first author.

Aggregate modelling of multi-processing workstations

I. Adan

Abstract.– We propose an aggregate model for manufacturing systems consisting of flow lines with finite buffers and parallel machines. The model is a multi-server station with process times depending on the work in process (WIP). An algorithm is developed to measure the WIP-dependent process times directly from industrial data such as arrival times at and departure times from the manufacturing system. Simulation results show that the aggregate model accurately predicts the mean flow time.

Flexible server allocation and customer routing policies for two parallel queues when service rates are not additive

H.-S. Ahn and *M.E. Lewis*

Abstract.– In this paper we consider the simultaneous control of routing and capacity allocation of two servers with two parallel stations. Customers arrive to each station in accordance with independent Poisson processes. Station dependent holding costs are accrued per customer per unit time. When a customer arrives it may join the queue at the station where it arrived or be routed at cost r to the other station. All service requirements are assumed to be exponential with rate 1. When the servers work at separate stations the service rates are μ, while when the work at same station it is μ_c. The sub-additive ($2\mu > \mu_c$) and super-additive ($\mu_c > 2\mu$) cases are considered.

Distribution of the successful and blocked events in a discrete-time retrial queue

J. Amador and *P. Moreno*

Abstract.– We consider a discrete-time multi-server retrial queue with geometrical arrival and service process and classical retrial policy. The probability distributions, the first and second moments and the cross moments of the successful and blocked events made by the external and repeated customers are analyzed. Several numerical examples and a cost function illustrate the study.

Optimal routing in parallel, non-observable queue and the price of anarchy revisited

J. Anselmi and *B. Gaujal*

Abstract.– We consider a network of parallel, non-observable queues and analyze the Price of Anarchy (PoA), an index measuring the worst-case performance loss of a decentralized system with respect to its centralized counterpart. Our analysis is undertaken from the new point of view where the router has the memory of previous dispatching decisions, which significantly complicates the nature of the problem.

In the limiting regime where the demands proportionally grow with the network capacity, we provide a tight lower bound on the socially-optimal response time and a tight upper bound on the PoA by means of convex programming. Then, we exploit this result to show, numerically, that the billiard routing yields a response time which is remarkably close to our lower bound, implying that it minimizes response time.

To study the added-value of non-Bernoulli routers, we introduce the Price of Forgetting (PoF) and prove that it is bounded from above by two, which is tight in heavy-traffic. When homogeneous queues are considered, both the PoF and the PoA boil down to a simple expression of the Lambert W function. This is increasing in the network utilization and contrasts with the case of memoryless routers, which makes the PoA equal to one.

Finally, structural properties are derived numerically for the PoF. These claim that the benefit of having memory in the router is independent of the network size and heterogeneity, while monotonically depending on the network load only. These properties yield simple product-forms well-approximating both the socially-optimal response time and the PoA.

A continuous-time queueing system with renovation

I. Atencia, *I. Fortes* and *S. Sánchez*

Abstract.– We study in this paper a single-line queueing system with poisson arrivals and renovation, that is, once a service is completed all the customers in the queue get service simultaneously where service times are exponentially distributed.

We carry out an extensive analysis of the system, including the joint generating function of the number of customers in the service and in the queue, the mean number of the customers in the system and in the queue, and the mean queueing time and sojourn time.

Finally, we give numerical examples to illustrate the effect of the parameters on several performance characteristics.

A discrete-time queueing system with optimal LCFS discipline

I. Atencia and *A.V. Pechinkin*

Abstract.– We consider a discrete-time queueing system in which the arriving customers decide with a certain probability to be served under a LCFS-PR discipline and with complementary probability to join the queue. The service times of the expelled customers are independent of their previous ones.

The arrivals are assumed to be geometrical and the service times are arbitrarily distributed. We carry out an extensive analysis of the system developing recursive formulae and generating functions for the steady-state distribution of the number of customers in the system and obtaining also recursive formulae and generating functions for the stationary distribution of the busy period and sojourn time as well as some performance measures.

On first passage problems for piecewise deterministic processes with phase-type jumps

F. Avram and *X. de Lucas Araujo*

Abstract.– We provide some examples of "solvable" spectrally negative piecewise deterministic processes with phase-type and mixtures of exponentials jumps, in the sense that the computation of survival, "two-sided" ruin probabilities and "expected dividends" (the Dirichlet and mixed Neumann problems in analysis terminology) are "quasi-explicit". Interpretations as busy period problems for some non standard queuing models is possible.

Markov modulated fluid queues with server vacations

J.W. Baek, *H.W. Lee*, *S.W. Lee*, *S. Ahn* and *L. Feng*

Abstract.– In this talk, we consider the Markov modulated fluid queues with server vacations. As soon as the fluid level becomes 0, the server leaves for vacations of random lengths. During the vacation period, the fluid level increases vertically or linearly depending upon the model. We consider two vacation cases: multiple and single.

We derive the LSTs and the mean performance measures of the fluid content and present computational examples. We also discuss the applications of these vacation fluid models to communication systems and inventory systems. Lastly, the possibility of decompositions of the fluid distributions will be discussed.

Performance analysis of sleep mode operation in IEEE 802.16m advanced mobile WiMAX

S. Baek, *J.J. Son* and *B.D. Choi*

Abstract.– The IEEE 802.16m is a currently being processed standard for IMT-Advanced next generation mobile networks. Due to the mobility, power saving is one of the significant issues for the battery-powered mobile stations. The IEEE 802.16m has a sleep mode operation as a power saving mechanism. Under the sleep mode, traffic indication messages are periodically sent by the base station at the beginning of every constant interval called the sleep cycle. We mathematically analyze the sleep mode operation in IEEE 802.16m. We construct a 2-dimensional discrete time embedded Markov chain in which the embedded points are the beginning of the sleep cycle and the end of the frame that a packet completes its transmission. The Markov chain is of $M/G/1$ type whose steady-state probability is obtained by the matrix analytic method. We obtain the average power consumption of a mobile station and the average delay of a packet. We show that there is a trade-off between the power consumption and the average delay. Using the performance analysis, we find the optimal lengths of the sleep cycle and the close-down time which minimize the power consumption while satisfying the required quality of service (QoS) constraint on the average packet delay. The numerical results show that the power consumption of sleep mode in IEEE 802.16m is better than that in legacy IEEE 802.16e.

Completion time analysis for infinite server queues with two service speeds

M. Baykal-Gursoy and *Z. Duan*

Abstract.– An infinite server queue with two-speed modulated generally distributed service time is considered. Service completion time for customers in such a queue with generally distributed down and exponentially distributed up periods is studied and Laplace transform of the completion time distribution is obtained. Special cases with exponential service times and down times are also investigated.

Polling tree networks

P. Beekhuizen, *D. Denteneer* and *J. Resing*

Abstract.– In this talk we consider concentrating tree networks of polling stations. The model is primarily motivated by the performance analysis of networks on chips for applications where all traffic has the same destination. Assuming that the service discipline in the final node of the network is HoL-based we construct a reduced system, consisting of only one node, for which there exists a distributional relation between the queue contents of the original network and the queue contents of the reduced system. Here, HoL-based essentially means that the decision which queue to serve only depends on whether or not queues are occupied (now and in the past) but not on the number of packets present in the queues. The reduction result can be used to approximate mean end-to-end delays in polling tree networks by using exact or approximate expressions for mean waiting times in single-station polling systems.

In the second part of the talk we focus on polling tree networks with flow control. To study how the total throughput of the network is divided over packets from the different sources (which depends on flow control limits, buffer sizes and service disciplines), we consider a closed queueing network which resembles a network of polling stations with flow control operating under heavy load.

Inference statistic for retrial queues with negative arrivals

L. Berdjoudj and *D. Aissani*

Abstract.– This paper studies maximum likelihood as well as confidence intervals of an $M/M/1$ retrial queue with negative arrivals. Under study state condition. We derive the maximum likelihood estimates of the mean primary arrivals, mean negative arrivals, mean retrial rates and the mean service rate. We also develop the confidence interval formula for the parameter "traffic of the system", the probability of empty system and the expected number of customers in this system.

A polling model with smart customers

M.A.A. Boon, *A.C.C. van Wijk*, *I.J.B.F. Adan* and *O.J. Boxma*

Abstract.– We consider a single-server, cyclic polling system with switch-over times. A distinguishing feature of the model is that the rates of the Poisson arrival processes at the various queues depend on the server location. We refer to this as 'smart customers', as an example of this is the situation where arriving customers choose which queue they join, based on

the current position of the server. For this system, we derive the waiting time distribution for each customer type, by applying a generalized version of the distributional form of Little's law to the joint queue length distribution at departure epochs. We also provide an alternative, more efficient way to determine the mean queue lengths and mean waiting times, using Mean Value Analysis for polling systems. Under certain conditions, a Pseudo-Conservation Law for the total amount of work in the system holds. Typical features of the model under consideration are demonstrated in several numerical examples, for instance the optimal queue for customers to join.

A fluid model analysis of streaming video in the presence of time-varying service capacity

J.W. Bosman, *R. Nunez Queija* and *R.D. van der Mei*

Abstract.– In this paper a mathematical model is proposed for the analyses of constant bit rate streaming video over a network. At playback side video has to be buffered in order to prevent video playback freeze. The speed for transferring video depends on the cross traffic behaviour on the network. The resulting transfer speed over the network is modelled by a continuous time Markov Chain (CTMC) where the states correspond to the network speed available for transferring the streaming video. This model consists of two fluid queue parts. The first part models the network used for video transfer (constant rate in, variable rate out). The second part models the video playback buffer (variable rate in, constant rate out). The network and playback part are connected to each other by fitting a CTMC for describing the output process from the network queue. The resulting fit is validated by comparison to the analyses of the network and playback part at once. This model will significantly reduce the complexity for analysing more complex settings as it reduces multi-dimensional analyses to smaller one-dimensional parts.

Equilibrium balking behavior in an unreliable queue with complete removals at failure epochs

O. Boudali and *A. Economou*

Abstract.– We consider a single-server Markovian queue with server failures, complete removals of customers and repair times. Failures of the server occur according to a Poisson Process. Whenever a failure occurs, all the customers are forced to abandon the system. The system is rendered inoperative and an exponential repair time process is set on. We consider two different models according to whether the system admits customers during the repair time or not. We assume that arriving customers decide whether

to join the system or balk based on a natural reward-cost structure which incorporates their desire for service, their unwillingness both to wait and to be removed due to a failure as well as failure compensation.

We examine customer behavior under various levels of information regarding the system state. More specifically, a customer, before making his decision, may be fully informed about the exact state of the system (state of server and number of present customers), partially informed or not informed at all. We derive equilibrium strategies for the customers under the various levels of information. We also illustrate several qualitative aspects of the model and the behavior of the customers by presenting some numerical scenarios.

$M/G/c/c$ queues with workload-dependent arrival rate

P. Brill and *M.L. Huang*

Abstract.– We derive the steady-state probability density function (pdf) of the total workload in an $M/G/c/c$ queue such that the arrival rate depends on the current total workload in the occupied servers. The talk first reviews the steady-state pdf of the workload in the standard $M/G/c/c$ queue. Then we give a set of model equations for the pdf of the workload, in the state-dependent model in when the arrival rate depends on the total current workload. The method of analysis utilizes order statistics and level crossing theory. Solutions are formulated for several $M/M/c/c$ variants.

On minimal representation of rational arrival processes

P. Buchholz and *M. Telek*

Abstract.– Rational Arrival Processes (RAPs) are a general class of stochastic processes which include Markovian Arrival Processes (MAPs) as a subclass. We study in this paper RAPs and their representations of different size.

We show transformation methods between different representations and present conditions to evaluate the size of the minimal representation. By showing some analogies to linear system theory, a minimization approach is defined which allows one to transform a RAP into one of its minimal representation (a representation of minimal size). An algorithm for computing a minimal representation is also given.

A multi-server queue with Markovian arrivals and priority services

S.R. Chakravarthy and *H. Qing*

Abstract.– We consider a multi-server queueing system in which two types of customers arrive according to a Markovian arrival process. Type 1 customers have preemptive priority over Type 2 customers. A Type 2 arrival finding all servers busy will be lost. However, a Type 1 customer finding all c servers busy with at least one Type 2 in service will get into service by pre-empting one of the Type 2 customers in service. The pre-empted Type 2 customers enter into a buffer of finite capacity. These (preempted) customers eventually leave the system after completing a service. This model is analyzed in steady state by exploiting the special nature of the queueing model. A number of useful performance measures along with some illustrative examples are reported.

A simple and complete computational analysis of $M/G_j^{(m,M)}/1/\infty$ and $M/G_j^{(m,M)}/1/B$ queues using roots

M.L. Chaudhry and *J. Gai*

Abstract.– Bar-Lev et al. [1] developed the generating function for the steady-state probabilities of the embedded Markov chain for the $M/G_j^{(m,M)}/1$ system. The general solution requires the calculation of the roots of the denominator of the underlying generating function. They claim that the use of roots may result in numerical inaccuracies especially when the decision variable M assumes a large value. In view of this, they obtain the solution by solving a finite space (B) model. This involves solving B linear equations. The problem becomes computationally challenging if B is large since $(M+2)$ symbolic differentiations are required to get those B linear equations.

In this paper, we first present a computationally simple solution to $M/G_j^{(m,M)}/1$ queue using roots. We also give an alternative simple method to solve the finite-space queue $M/G_j^{(m,M)}/1/B$.

Computational analysis of queue-length distribution arising in the continuous-time queue: $GI/G/1$

M.L. Chaudhry and *X. Yang*

Abstract.– The $GI/G/1$ queue is a generalized queueing model useful for a large number of practical applications. In this paper, we propose to numerically discuss the waiting-time distribution and from it the queue-length distribution using the distributional Little's law. We discuss three cases: (1) when the distributions of inter-arrival and service times have rational Laplace transform, (2) when such distributions do not have rational Laplace transforms, and (3) when the distributions involved do not have Laplace

transforms. Throughout our computations, we show that the root-finding method first successfully developed and implemented by Chaudhry in various papers can be fruitfully deployed to get all the desired results.

Influence of correlation in the arrival process on batch-service queueing systems

D. Claeys, *B. Steyaert*, *J. Walraevens*, *K. Laevens* and *H. Bruneel*

Abstract.– Numerous studies demonstrate that correlation in the arrival process might have a significant influence on the behaviour of queueing systems. The impact of correlation on batch-service queueing systems, nevertheless, has attracted nearly no attention, and therefore constitutes the subject of this contribution. More specifically, we deduce various performance measures related to the buffer content in a system whereby the batch server only initiates service if the buffer content reaches or exceeds some server threshold. Such a threshold leads to a better utilization of the server capacity. We further include a general dependency between the service time of a batch and the number of customers within it. In addition, we model the correlation by a discrete batch Markovian arrival process (D-BMAP), i.e. the distribution of the number of customer arrivals per slot depends on a background state which is determined by a first-order Markov chain. This process is the subject of many papers, as it has the advantage that it can capture most arrival patterns up to any desired precision.

By means of the obtained performance measures, we examine the influence of correlation on the behaviour of the system. For instance, we demonstrate that disregarding correlation, although it might lead to a severe underestimation of the buffer content, has only a small influence on the value of the optimal server threshold.

Brownian queues with modulated buffer

B. D'Auria and *O. Kella*

Abstract.– In this talk we present the analysis of a two-sided regulated Brownian motion, generally known as Brownian queue, whose reflecting barriers depend on an independent random environment. The environment is an irreducible finite state space Markov chain and depending on its state we assume that the Brownian Motion's parameters change together with the level value of the upper reflecting barrier.

What makes this model interesting and different form the classical ones is the intrinsic presence of discontinuities in its dynamics due to the abrupt changes of the barrier level that imply impulsive losses in the buffer content of the queue.

For this model we study the stationary distribution, and in the special case of a two-state environment we analyze in more detail the time duration from a fixed instant up to the next discontinuity of the process.

FCFS in dual-class discrete-time queueing systems

S. De Clercq, *K. Laevens*, *B. Steyaert* and *H. Bruneel*

Abstract.– The problem with the FCFS server discipline in discrete-time queueing systems is that it doesn't actually say what happens if multiple customers enter the system at the same time, which in the discrete-time paradigm translates into 'during the same time-slot'. In other words, it doesn't specify in which order such customers are served? When we consider multiple types of customers, each requiring different service time distributions, the precise order of service even starts to affect things like queue content and delays of arbitrary customers, so specifying this order will be prime. In this paper we study a dual-class discrete-time queueing system with a general independent arrival process and generally distributed service times. The service discipline is FCFS and customers entering during the same time-slot are served in an arbitrary order. It will be our goal to search for queue content and delays of certain types of customers. If one thinks of the time-slot as a continuous but bounded time period, the arbitrary service order is equivalent to FCFS if different customers have different arrival epochs and if the arrival epochs are independent of class. For this reason we propose two distinct ways of analysing; one utilizing permutations, the other considering a bounded continuous time frame. The equivalent problem in continuous time has thus far only been solved for a Poissonian arrival process, and some specific cases of MAP.

Fluid limits and maximal load in multi-wavelength optical buffer systems

K. De Turck, *W. Rogiest*, *D. Fiems*, *H. Bruneel* and *S. Wittevrongel*

Abstract.– The conversion of the internet backbone to all-optical technology is currently an important topic in research. Although major technical obstacles remain, the promised bandwidth gains for such conversions are spectacular.

Buffers are an important component of the backbone, and they are currently implemented in the electrical domain by means of random access memory (RAM). However, it has proven to be very hard to make true optical RAM, and this has brought so-called fiber-delay lines (FDLs) to the attention: in an FDL, data is not truly stored, but merely delayed for a fixed amount of time. A set of fiber delay lines can thus mimic to some extent

optical RAM. They have different performance characteristics however, in particular the occurrence of so-called voids causes performance losses which remain to be fully quantified.

From a queueing-theoretical point of view, the resulting systems get substantially harder to analyse, especially for the realistic case of multi-wavelength systems, where closed-form expressions are hard to obtain. In this contribution, we therefore look at the fluid limit of the process. The rewards of this approach are substantial: we are able to answer not only questions on the stability, but also on optimal wavelength allocation strategies. Moreover we show how fluid limits can help to find good approximative models.

Boundary probabilities in the discrete-time selective-repeat ARQ queueing problem

S. De Vuyst, *S. Wittevrongel* and *H. Bruneel*

Abstract.– We consider the scenario of reliably sending packets from transmitter to receiver over an error-prone channel, using the Selective-Repeat ARQ protocol. At the transmitter side, we assume a discrete-time queue with a single server and infinite capacity which is fed by an uncorrelated stream of packet arrivals. For each transmitted packet, the receiver returns a negative or positive acknowledgement (NACK or ACK) notifying the transmitter that either an error occurred and the packet has to be sent again or that the packet was received correctly and can thus be removed from the queue. The transmission times of the packets are fixed, as well as the feedback delay of the channel.

Assuming a static probability for a transmission error to occur, this queueing problem was solved more or less exactly by Konheim in 1980. Nevertheless, his method suffers badly from state-space explosion since the number of unknown boundary probabilities increases exponentially with the feedback delay. Determining these unknowns amounts to solving a huge, possibly ill-conditioned, set of linear equations.

Our contribution is to review the analysis of Konheim and to propose several approximation methods for obtaining the unknown boundary probabilities, some of which are methods adapted from statistical physics. Our aim is to come up with a solution that is both fast to compute and asymptotically correct for large feedback delays. Extensions to Markov-modulated arrival processes are considered as well.

Tail behaviour of a finite-/infinite-capacity priority queue

T. Demoor, *J. Walraevens*, *F. Fiems* and *H. Bruneel*

Abstract.– Priority queues have been studied extensively throughout the last few decades. One of the main performance measures of interest is the loss probability, caused by the evident finiteness of queues. From analytical point of view, finite-capacity queues are more complex to study than queues with infinite capacity. Therefore, the most common approach for studying the loss probability is to analyze the tail behaviour of the corresponding infinite-capacity system. From this behaviour, the loss probability of the finite-capacity system can be approximated. However, this approximation can be rather inaccurate.

Consider a system with two priority classes where the low-priority queue has infinite capacity but the high-priority queue can only accommodate N packets. In the system where both queues have infinite capacity, the tail behaviour of the queue content of the low-priority class can be geometric or non-geometric, dependent on whether the dominant singularity of the respective generating function is a pole or a branch point. However, in the system with a finite high-priority queue the tail is always geometric, even when N is large. Investigating this discrepancy is the topic of our contribution.

In the finite system, the low-priority queue content can be expressed by a matrix equation. These matrices contain (partial) generating functions. Finding the poles corresponds to solving a non-linear eigenvalue problem. However, finding all poles is infeasible for larger N. Therefore, a numerical method is used to obtain the poles in a certain region (around the dominant pole).

Approximations for the waiting-time distribution in polling systems with renewal arrivals

J.L. Dorsman, *R.D. van der Mei* and *E.M.M. Winands*

Abstract.– A polling system is a queueing system that consists of a number of queues, attended by a single server, typically in a cyclic order. Polling models find many applications in areas such as computer-communication systems, manufacturing systems, maintenance systems and traffic systems.

The vast majority of papers in the literature assume Poisson arrivals and are focused on the expected delay at each of the queues. However, in many applications the Poisson assumption is not realistic and moreover, remarkably little attention has been paid to the evaluation of the complete distribution of the delay at the queues.

Motivated by this, in this paper we focus on the waiting-time distributions for cyclic polling models with renewal arrivals, and exhaustive service at each of the queues. In the absence of exact results, we propose a closed-form approximation for the complete waiting time distribution by combining known light- and heavy-traffic results. Extensive simulation results show

that the approximations are highly accurate over a wide range of parameter combinations.

Multi-server queueing systems with broadcasting service disciplines, servers breakdowns and heating

A. Dudin, *S. Bin* and *C.S. Kim*

Abstract.– Multi-server queueing systems of the $MAP/PN/N$ type with broadcasting service disciplines, servers breakdowns and heating are under study. Broadcasting service discipline assumes engagement of all free servers to the service of an arriving customer. Full broadcasting and partial broadcasting are considered. The latter one assumes that some set of the idle servers are reserved in situation if the number of available servers is large or, oppositedly, small. Essential advantages of the broadcasting service discipline in the cases when the service of a customer can be provided with an error, when the servers are subject to breakdowns and repairs and when the preliminary heating of an idle server is required are numerically shown. The broadcasting service discipline can provide higher probability of the successful service of a customer and lesser response time in comparison to the classical service discipline. Effect of correlation and variation in the arrival and breakdown processes and variation in the service, recovering and heating processes is clarified. Effect of the partial broadcasting is discussed.

The single server queue with synchronized repeated services

A. Economou, *S. Kapodistria* and *J. Resing*

Abstract.– We consider a single server queueing system with generally distributed synchronized services. More specifically, customers arrive according to a Poisson process and there is a single server that provides service, if there is at least one customer present in the system. Upon the initialization of a service, all present customers start to get service simultaneously. We consider the gated version of the model, that is, customers who arrive during a service time do not receive service immediately but wait for the beginning of the next service time. At service completion epochs, all served customers decide simultaneously and independently whether they will leave the system or stay for another service. The probability that a served customer gets another service is the same for all customers. Therefore, the number of customers in the system is reduced according to a binomial distribution.

We study the model and derive its main performance measures that include the equilibrium distribution of the number of customers at service completion epochs and in continuous time, the busy period and the sojourn time distributions. Moreover, we prove some limiting results regarding the

behavior of the system in the extreme cases of the synchronization level. Several variants and extensions of the model are also discussed.

Queueing model of the FSO/RF hybrid channel with heterogeneous links subject to failures

D. Efrosinin and *V. Rykov*

Abstract.– With the recent developments in semiconductor technology, free space optical (FSO) or optical wireless communication channel has become more attractive comparing to optical fiber communications or radio frequency (RF) systems. An FSO system offers much higher data rates, easier to deploy and less expensive. However, despite the advantages of an FSO system, there are some limitations for optical wireless, e.g. the adverse weather conditions. The practical and obvious solution to this problem would be to combine the less reliable but higher rate FSO link with more reliable but lower rate RF link. In this case the natural question is the formulation of the link allocation policy.

In the paper we consider the experimental hybrid FSO/RF channel which is analyzed by a multi-server queue with two types of heterogeneous servers subject to the failures. The FSO link can partially or complete fail due to the atmospheric attenuation caused by haze, heavy fog and snowfalls. The RF link can fail due to the heavy rain. Under normal operating conditions data is transmitted over the FSO link and the RF link is waiting as a hot stand-by. To control the allocation between the FSO and RF link we set-up a threshold-based policy using the information about the signal level of the FSO link and the number of waiting packets. We provide the performance and availability link analysis under the given threshold policy and formulate some optimization problems to improve the links switching scheme. The performance results are compared to other heuristic control policies.

Universal generation of fractal statistics: data-traffic, queues, and physics

I. Eliazar and *J. Klafter*

Abstract.– We present a stochastic superposition model which is capable of generating - in a universal fashion - various "fractal statistics": anomalous diffusion, Levy laws, 1/f noises, and self-similarity. The model considers the superposition of independent stochastic processes: all processes sharing a common - yet arbitrary - stochastic process-pattern; each process having its own random parameters - initiation epoch, amplitude, and frequency. The stochastic superposition model is general and robust, and arises naturally in diverse fields of science and engineering: transmission channels

and routers, and Internet servers in Communication; background noises in Physics and Electrical Engineering; probe dynamics in "stochastic baths", and shot noise processes in Physics; river flows in Hydrology; tax revenues of states in Economics. In the context of Queueing Theory, the stochastic superposition model can be regarded as a generalized $M/G/\infty$ model, in which the superimposed processes represent the random workload processes of incoming jobs. Considering a specific process-statistic of the superposition model's output, we focus on the following universality question: Is there a randomization of processes' parameters which renders the output's process-statistic invariant with respect to the processes' common stochastic pattern? The answer - for various process-statistics - turns out to be affirmative, and yield the aforementioned "fractal statistics". This research thus establishes a unified framework that: (i) explains the universal emergence of "fractal statistics" in diverse fields of science and engineering, and (ii) provides an explicit model and randomization-algorithms that universally yield desired "fractal statistics". Observed from the "$M/G/\infty$ perspective", this framework serves as natural and robust model for data-traffic processes displaying "fractal statistics" - which are prevalent in contemporary communication systems.

Optimality of trunk reservation for an $M/M/k/N$ queue with several types of customers and holding costs

E.A. Feinberg and *F. Yang*

Abstract.– We study optimal admission to an $M/M/k/N$ queue with several customer types. The cost structure consists of revenues collected from admitted customers and holding costs, both of which depend on customer types. The goal is to find an admission policy that minimizes average rewards per unit time. Under natural assumptions we show that an optimal policy has a trunk reservation form. Previously problems with the same holding costs for all customers have been studied in the literature.

Customer loss probabilities in busy-periods of batch arrival oscillating $M/G/1/n$ queueing systems

F. Ferreira, *A. Pacheco* and *H. Ribeiro*

Abstract.– In classical queueing systems it is assumed that customers arrive single at a service facility and find an infinite waiting room. However, these assumptions are often violated in practice. In addition, it is a well-known fact that standard queueing systems are not able to achieve higher rates of service utilization concurrently with small loss probabilities. In order to improve the performance of queueing systems there have been proposed

in the literature systems whose characteristics depend on the evolution of the state of the system. In this line, threshold systems with hysteresis have service rates or service characteristics that change over time as forward and backward barriers are up-crossed or down-crossed, respectively.

We propose to analyze particular threshold systems with a single backward and forward barrier which are usually called oscillating queueing systems. Among these, we will address systems with (compound) Poisson input. Taking profit of the Markov regenerative structure of these oscillating systems, we derive the probability mass function of the number of customer losses in a busy period. The derived computational procedure is easy to implement and leads to a fast numerical computation of these loss probabilities. We illustrate the effectiveness of the computational procedure considering a wide variety of queues with different capacities, batch size distributions, and arrival and service parameters.

Light-correlation analysis of queues with autoregressive service times

D. Fiems and *B. Prabhu*

Abstract.– We consider a single-server queueing system. The arrival process is modelled as a Poisson process while the service times of the consecutive customers constitute a sequence of autoregressive random variables. Our interest into autoregressive service times comes from the need to capture bandwidth fluctuations on wireless network links. If these fluctuations are slow in comparison with the transmission times of the packets, transmission times of consecutive packets are correlated. Such correlation needs to be taken into account for an accurate performance assessment. By means of a transform approach, we obtain a functional equation for the joint transform of the queue content and the current service time at departure epochs in steady state. To the best of our knowledge, this functional equation cannot be solved by exact mathematical techniques, despite its simplicity. However, by means of a Taylor series expansion in the parameter of the autoregressive process, a "light-correlation" approximation is obtained for performance measures such as moments of the queue content and packet delay. We illustrate our approach by means of some numerical examples, thereby assessing the accuracy of our approximations by means of simulation.

Analysis of an $M/M/k$ system with exponential setup times under staggered boot up

A. Gandhi, *M. Harchol-Balter* and *I. Adan*

Abstract.– In this talk, we consider the $M/M/k$ queueing system with

setup times. This particular queueing model is common in manufacturing systems, where idle machines are turned off to save on operating costs, as well as in server farms, where idle servers are turned off to conserve power. However, turning servers back on requires a significant setup time, and possibly some setup cost (financial cost or power consumption). In the particular model we consider, at most one server can be in setup at any point of time, to limit the incurred setup cost. This model is referred to in data centers as the staggered boot up model.

While single servers with setup are well studied in the literature, the $M/M/k$ with setup is far less tractable and not well understood. The only case that has been studied via analysis is the case of staggered boot up, which was studied by Artalejo and Economou in 2005, but did not lead to closed form solutions.

In this talk, we provide the first analytical closed form expressions for the mean response time, limiting distribution of the system states, as well as the z-transform for the number of jobs in system for the staggered boot up model.

Our analysis reveals a very interesting property of the staggered boot up model: The distribution of response time can be decomposed into the sum of response time for an $M/M/k$ system without setup times and the exponential setup time.

Stochastic analysis and optimization of queueing networks for cloud computing

S. Ghosh and *M.S. Squillante*

Abstract.– Cloud computing represents a fundamental shift for future computing environments, which in turn creates fundamental differences in the corresponding stochastic capacity planning and resource allocation models and their optimization. In particular, the stochastic models for determining the optimal number of system resources to deploy and the optimal allocation of these system resources in the cloud need to take into account much burstier resource request patterns and much more complex resource requirements (including strong dependence structures) than in traditional models. Furthermore, the general utility function to be optimized needs to reflect additional complexities of cloud computing environments.

Motivated by these fundamental differences, we present a methodology for the stochastic analysis and optimization of capacity planning and resource allocation in queueing networks inspired by cloud computing environments. Our objective function is based on rewards that are generated when performance guarantees are satisfied and on penalties that are incurred otherwise. The corresponding performance criteria are obtained from Service-Level-Agreements (SLAs) between cloud service providers and each

customer class based on throughput and sojourn time metrics involving tail distributions. We derive a stochastic analysis of the underlying queueing network which includes bursty arrivals over time and service times with strong dependence structures. A stochastic optimization of resource capacity and resource allocation is then derived within the context of this class of queueing networks, where SLAs can be taken into account in both the constraints and the objective function. Numerical experiments illustrate and quantify the benefits of our approach.

Drift conditions for fluid queues

M. Govorun, *G. Latouche* and *M.A. Remiche*

Abstract.– For the purposes of obtaining stability conditions for fluid queues we may consider their close relationship to Quasi Birth-and-Death (QBD) processes, as is shown by V. Ramaswami in 1999. Ramaswami shows that the computation of the steady state distribution of a fluid queue can be directly obtained from the analysis of a QBD model.

In 1981 Neuts gives one drift criterion to determine whether a QBD process is positive recurrent. This criterion is based on the stationary distribution of the phase process, and its physical meaning is identical to the well-known "average drift" criterion for fluid queues.

In 2003 Latouche and Taylor introduce additional drift conditions for QBDs and prove them in a general context. We transform these drift conditions for QBDs to a fluid queues framework using the QBD – fluid queue connection.

As a result we have three new stability conditions. These new conditions are based on some eigenvectors of different matrices built from the phase generator matrix and the fluid input rates matrix. However, their probabilistic interpretation is not as clear as for the original drift condition for a fluid queue; in particular, the explanation from the fluid buffer behaviour in a long-term perspective no longer applies.

Scheduling in tandem flexible server queueing systems

J. Hasenbein and *B. Kim*

Abstract.– We prove a conjecture posed in Andradottir and Ayhan (OR, 2005), which studied a queueing system with flexible server stations. The system has two multiple server stations in tandem, with flexible servers which can be assigned to either station. The objective is to assign the servers so as to maximize throughput when there is a finite buffer between the two servers. We prove that a service ratio rule is optimal, for an arbitrary number of servers.

Performance analysis of binary exponential backoff MAC protocol for cognitive radio in the IEEE 802.16e/m network

S. Jin, *J.S. Park* and *B.D. Choi*

Abstract.– Radio spectrum is one of the most scarce and valuable resource for wireless communication. The traditional spectrum allocation policy which is allowed to use spectrum only for licensed users (or primary users) faces scarcity of spectrum. The concept of cognitive radio has been proposed for unlicensed users (or secondary users) to opportunistically access the unused spectrum by primary users.

In this paper, we propose a distributed MAC protocol for cognitive radio in the IEEE 802.16e/m WiMAX. Our proposed MAC for cognitive radio adopts truncated binary exponential backoff scheme which reflects the frame structure of IEEE 802.16e/m and can be very well adapted to dynamic unused spectrum and number of secondary users.

For performance of our proposed MAC, we construct 3-dimensional Markov chain with states (i,j,k) where i is backoff stage, j is backoff counter and k is the number of idle channels in a frame. By applying censored Markov chain method, we obtain the steady state probability of Markov chain and then throughput and packet delay of secondary users. We give numerical examples showing how much throughput of secondary users can be achieved depending on the dynamic unused channels and the number of secondary users.

Scaling limits of cyclically varying birth-death processes and applications

M. Jonckheere, *Y. Nazarathy* and *R. Nunez Queija*

Abstract.– Fluid limits of stochastic queueing systems have received considerable attention in recent years. The general idea is to scale space, time and/or system parameters as to obtain a simpler, yet accurate description of the system. A basic example is the single server queue with time speeded up and space scaled down at the same rate. A second well known example is the Markovian infinite server queue with the arrival rate speeded up and space scaled down at the same rate. Such scalings and their network generalizations are often useful for obtaining stability conditions and approximating optimal control policies.

We consider birth-death processes with general transition rates. We obtain an asymptotic scaling result, generalizing the Markovian single server and infinite server cases. We apply our results to the steady-state analysis of queueing systems with cyclic or time varying behaviour. Examples are

systems governed by deterministic cycles, queues with hysteresis control and queues with Markov-modulated arrival or service rates. The unifying property of such systems, is that if they are properly scaled, the resulting trajectories follow a cyclic or piece-wise deterministic behaviour which is determined by the asymptotic scaling.

Our results include approximations of the stationary distributions which are shown to be asymptotically exact. Such approximations prove to be useful in optimization and control problems which appear in power management and speed scaling of computer systems. We also outline and show preliminary experiments of an additional application: change point detection problems in internet traffic.

A single server queue with synchronized abandonments

S. Kapodistria

Abstract.– We study a single server queue with synchronized abandonments, i.e. all customers at given opportunities decide simultaneously, but independently, to abandon the system with the same probability p. More explicitly, we assume that customers arrive according to a Poisson process, there is a single server and the service times are generally distributed. Furthermore, we assume that the abandonment opportunities occur according to a Poisson process (so we can think that the transportation facility's arrivals occur according to this process). Then, at an abandonment opportunity epoch, every customer decides to abandon the system with probability p or remains in the system waiting for service with probability q=1-p, independently of the others.

First, we will present the Markovian case, that is the $M/M/1$ queue with synchronized abandonments, and distinguish two abandonment scenarios, in the first scenario all customers including the one in service can abandon the system, while in the second one we exclude the customer in service from the abandonment process. Subsequently, we will extend the talk to the case of the $M/G/1$ queue with synchronized abandonments.

A multi-class $M/PH/1$ queue with deterministic impatience times

J. Kim, *B. Kim* and *J. Kim*

Abstract.– We consider a multi-class $M/PH/1$ queue with deterministic impatience times in which there are several classes of customers. Customers of each class arrive according to a Poisson process and have phase-type service requirements. A customer in the queue departs the system giving up his/her service if he/she has not started service within a deterministic time depending on his/her class. We find a related Markov process by us-

ing Markovian structure of the phase-type distributions for services, and then obtain the stationary distribution of the Markov process. By using the results of the stationary distribution of the Markov process, we obtain performance measures such as the loss probability, the waiting time distribution and the queue size distribution. Numerical examples are also given. This paper generalizes the results of Xiong et al. (2008), which give analytic results for an $M/G/1$ queue with deterministic impatience time only when there is a single class of customers and the service time distribution has a hyper-exponential distribution of order 2.

Fair connection admission control with adaptive thresholds for wireless multimedia networks

T.O. Kim, *A.S. Alfa* and *B.D. Choi*

Abstract.– Recently, there has been in explosive demand for multimedia services that require high-level quality-of-service (QoS). Connection admission control (CAC) is crucial to determine connection-level and packet-level QoS in wireless multimedia networks. Most CAC protocols are designed to enhance network utilization, while causing a bias against wideband connections. Motivated by this, we propose a fair CAC protocol for VBR multimedia services.

We consider voice and video connections as representative narrowband and wideband traffic, respectively. To model voice and video traffic sources, we use discrete-time Markov modulated deterministic processes. In order to achieve high network utilization by statistical multiplexing gain, our protocol determines the total required bandwidth of ongoing connections by constructing a finite $M/G/1$ type discrete-time Markov chain describing the volumes of aggregated packets from connections of each type, considering requirements on packet-level QoS. CAC is based on the remaining bandwidth (which equals to the difference between total bandwidth and required bandwidth of ongoing connections) and the dynamic threshold (which equals to the amount of additionally required bandwidth to admit one more video connection and varies adaptively with the number of ongoing video connections). If remaining bandwidth is more than or equal to the dynamic threshold, any arriving connection is admitted; otherwise, rejected. In this manner, our protocol achieves the same blocking probability for voice and video connections. We obtain the connection blocking probabilities by constructing a continuous time Markov chain describing the number of connections. Numerical results show that our protocol enhances network utilization compared to previous fair CAC protocols.

Join the shortest queue among similar c parallel service stations

M. Kobayashi, *Y. Sakuma* and *M. Miyazawa*

Abstract.– We are concerned with an $M/M/c$ queue with join the shortest queue discipline for an arbitrary positive integer c. We show that the tail decay rate of the stationary distribution for the shortest queue length is bounded by the c-th power of the traffic intensity of the corresponding $M/M/c$ queue with a single waiting line. To this end, we consider the convergence domain of the joint moment generating function of the stationary distribution for the whole system. We conjecture this upper bound is identical with the decay rate, and discuss about its reasons.

Queueing system with state-dependent controlled batch arrivals and server under maintenance

B. Krishna Kumar, *S. Pavai Madheswari* and *S.R. Anantha Lakshmi*

Abstract.– In this paper, a single server Markovian queue with the server under maintenance is studied. A batch of customers is allowed whenever the server is idle such that each individual customer in the batch is subjected to a control admission policy upon arrival. For this queueing system, the transient probabilities of the system size and mean of the system size are obtained explicitly. The steady-state analysis and associated key performance measures of the system are also carried out. Finally, extensive numerical illustrations are presented.

Queues with service interruption: a survey and some new results

A. Krishnamoorthy

Abstract.– Queues with interruption have been extensively investigated by several researchers. Vacation queues form a class of Queues with interruption. Most of the vacation models assume that only on a service completion, and not while a service is in progress, does the server go for a vacation. However, interruption can occur even when a service is in progress. For example it can be in the form of server break down or the server pre-empts the service of the customer to attend a higher priority work and so on. It is precisely this type of interruption that we discuss here. We assume that only when a service is in progress does interruption takes place. We discuss interruptions of both Type I and II (of the type described by Type I and Type II counters, respectively). In Type I interruption, the interrupted service does not get further extended due to the incidence of another interruption (ie., it gets locked the moment an interruption strikes), whereas in Type II interruption, the duration of interruption gets extended).

On recovering from an interruption, whether of Type I or II, it is essen-

tial to decide whether the service which got interrupted has to be resumed from where it got disrupted or to start from scratch. We give a rule for determining this.

First we give a survey of related work starting from White and Christie (1958; Journal of Operations research), followed by Gaver (1962-Jl. Royal Statistical Society), Keilson (1968-Annals of Math. Stat)), Takine & Sen Gupta (1993, QUESTA)), Ibe and Trivedi (1985, QUESTA), Fiems et al. (2008, EJOR) and so on along with a few papers of Krishnamoorthy et al dealing with queues with interruption. Then a few new results are discussed. Stability of such queueing systems, performance measures and some numerical results are provided.

Queues with random order of service and multiple classes of customers

K. Laevens, *W. Rogiest*, *J. Walraevens* and *H. Bruneel*

Abstract.– We derive results for a discrete-time queueing system operating under the random-order-of-service policy that serves customers that each belong to one of multiple classes. Customer service times are independent random variables with class-dependent distribution. The numbers of new arrivals in each slot are independent and identically distributed over time, but can show correlation across classes within a single slot, i.e., we consider a batch-Bernoulli input process with general multi-class batch size distribution.

Using a transform-based approach, results are obtained for the first few moments of the system content at various epochs. Analyzing the embedded Markov chain at service starts leads to an equation the joint generating function of the per-class customers in system satisfies. From it, moments can be extracted without need for an explicit solution for that generating function. Renewal arguments then lead to the moments of the system content at random slot boundaries.

A somewhat unexpected result shows that, under certain conditions on the batch-size and service-time distributions, the random-order-of-service policy can outperform the first-come-first-served policy in terms of mean system content and thus, by virtue of Little's law, also in terms of mean customer waiting time.

Results for the equivalent continuous-time system can be obtained along similar lines or through a limit procedure. This is briefly touched upon in the paper.

Queues with boundary assistance and the many effects of truncation

G. Latouche, *G.T. Nguyen* and *P.G. Taylor*

Abstract.– Consider a simple system with two queues, each with its own Poisson stream of customers and its own server, of which the service time is exponentially distributed. If a server is free and its associated queue is empty, then it serves a neighboring customer, provided there is one. The service rate of a server does not change regardless of whether it is serving its own or a neighbouring customer.

This system is an example of queues with boundary assistance and may be modeled by a Quasi-Birth-and-Death process, with the level being the minimum queue length and the phase the difference between the two queue lengths. We investigate the asymptotic behavior of the level in the original infinitely-many-phase system and in its finite approximations where the queue length difference is bounded. In the infinite system, we determine the convergence norm of the rate operator of the QBD, and consequently the interval in which the decay rate of the minimum queue length lies.

We consider four finite approximations: one is the infinitely-many-phase system truncated without augmentation, and three are obtained from different augmenting schemes such that stochasticity is preserved. We show that the first has a monotonically increasing decay rate that approaches the convergence norm of the rate operator in the infinite system, as the truncation size tends to infinity. All three truncated and augmented systems have decay rates that are independent on the truncation size. Finally, we observe that the stability regions for the original system and the aforementioned finite approximations are nonequivalent.

Managing performance and power consumption in a server farm

I. Mitrani

Abstract.– We examine the problem of managing a server farm in a way that attempts to satisfy the conflicting objectives of high performance and low power consumption. A subset of servers is designated as 'reserve'. Those reserves are powered up when the number of jobs in the system is sufficiently high, and are powered down when that number is sufficiently low. The question of how to choose the number of reserves, the up threshold and the down threshold is answered by analyzing a suitable queueing model and minimizing an appropriate cost function. The solution is obtained in essentially a closed form, in the sense that there is no need to solve simultaneous sets of equations. Nevertheless, the search for the optimal policy may be computationally expensive when the number of servers in the farm is large. For that reason, heuristic policies are proposed that are easily computable and are close to optimal. Numerical results are presented.

Light tail asymptotics in multidimensional reflected random walks and queueing networks

M. Miyazawa

Abstract.– We are concerned with tail asymptotics of the stationary distribution of a reflecting random walk on the multidimensional nonnegative integer quadrant, provided it exists. This random walk is assumed to have homogeneous jump transitions on each boundary faces in addition to the inside of the quadrant. We are particularly interested to compute the tail decay rates of the stationary distribution in the directions of the coordinate axes under the light tail regime.

This decay rate problem is generally very hard to solve in the sense of computing them from modelling primitives except for some two dimensional cases. Its major difficulty comes from the presence of the boundary faces on which jump transitions may be very different. This is a problem of so called large deviations. Its theory can be applied, but may not be convenient to compute the decay rates. There are other approaches to attack this problem, which have been developed independently.

The purpose of this talk is two holds. We first overview the current approaches to attack the problem from a unified viewpoint. We then take up a moment generating approach, recently developed by the author and his colleagues, and consider its possible extensions, including for the more than two dimensional case. We will also remark on the reflecting random walk modulated by a background Markov chain.

From the modelling viewpoint, the reflecting multidimensional random walk has a wide rage of applications. A typical application is a discrete time queueing network. We exemplify our results for it.

Periodic and continuous review analytical models for airline check-in queues

S. Moosa, *M. Parlar* and *B. Rodrigues*

Abstract.– A well managed check-in counter system improves not only the overall image of the airport as perceived by the passengers but also impacts on the downstream operations. Hence, this research addresses the important question of how best to manage the allocation of check-in counters in an exclusive use check-in counter system dedicated to a specific flight. Almost all prior works on this important topic are only simulations studies. In this study, we propose tractable analytical models guided by real observations at an award winning airport in Asia. More specifically, we consider how the airport operator should review congestion status and then decide whether to open more counters or not. One striking feature of the problem

is that the population of customers is finite and unlike in the mainstream finite population queues, the customers in this queue do not go back to the population after service. Hence, the system can never reach steady state. We propose both periodic and continuous review analytical models. Our approach facilitates real-time decision-making by balancing passenger wait costs, counter staffing costs and aircraft delay costs. We use a dynamic programming approach for decision making. Results of model testing for the continuous review model with real life data collected at an international airport will also be presented. The model and the methodology used in this research can help airport service providers, airport regulatory authorities and airlines in the management of the valuable resource of check-in counters and the associated check-in staff.

Stability analysis of a two-queue cascade network

E. Morozov and *B. Steyaert*

Abstract.– We consider a cascade of two coupled infinite-capacity single-server queues, both with generally distributed and i.i.d. interarrival and service times. The configuration is such that, whenever the second queue becomes empty while customers are awaiting service at the first queue, one customer can jump to the second queue to be served there. The opposite event does not occur however: customers that arrive in the second queue remain there to await service. Such a queueing network is related to the so-called cross-trained server paradigm with two servers, where one server is fit to handle a limited set of customers, whereas the second server has been trained to handle all types of customers that enter.

For this system we will investigate the stability condition of both queues, by examining the regeneration cycles of system's workload process. More specifically, we determine the range of system parameters for which the mean regeneration time is finite, and hence, this process is positive recurrent.

Moreover, we will indicate how to extend these results to more general scenarios, including 1) multiserver queues; 2) a different service time distribution for customers jumping from the first server to the second one; 3) jumps occur only if the queue size in the first station exceeds a level $d \geq 1$. Finally, we will discuss the possibility to extend this analysis to any number of queues that are configured as a cascade, i.e., if a queue becomes idle, it can handle a customer that is waiting in one of the non-empty upstream queues.

Bandwidth allocation under delay percentiles

X. Mountrouidou, *K. Kontovassilis*, *B. Anjum* and *H. Perros*

Abstract.– We consider the issue of how much bandwidth should be allocated on each router along the path of an MPLS connection so that the end-to-end delay of packets switched through this connection is bounded statistically. That is, γ percent of the time it is less than a given value. For this, we model a connection by a tandem queueing network of infinite capacity queues, and we assume that the packet transmission time at each router is exponentially distributed with the same rate μ. The arrival of packets to the connection is depicted by an MMPP2. We obtain a closed-form solution for the waiting time distribution in an $MMPP2/M/1$ queue, which is in the form of a hyper-exponential distribution with two stages. Using a simple nodal decomposition, we calculate approximately the end-to-end delay and subsequently the γ percentile. The source of the approximation is the characterization of the departure process from each node as an MMPP2, which is carried out using existing results from the literature. Subsequently, we use the γ percentile in a simple search algorithm to obtain the bandwidth necessary, depicted by the service rate μ at each queue. Finally, we also propose some alternative methods that permit us to calculate the bandwidth μ by analyzing the first node of the tandem network only.

A discrete-time retrial queueing model with abandonments

R. Nobel

Abstract.– This paper presents a one-server queueing model with retrials and abandonments in discrete time. In every time slot a generally distributed number (batch) of customers arrives. The different numbers of arrivals in consecutive slots are mutually independent. Each customer requires a geometrically distributed service time, counted in slots. Customers arriving in a slot can start their service only at the beginning of the next slot [i.e. delayed access]. When upon arrival customers find the server busy, they enter a virtual waiting space, the so-called orbit, and try to approach the server some random (geometric) time later, individually and independently from the other customers in the orbit. When upon arrival customers find the server idle, then one of the incoming customers (randomly chosen) starts his service at the beginning of the next slot, whereas the other incoming customers, if any, (re)enter the orbit. Every time slot each customer in the orbit decides to abandon the orbit forever with a fixed (abandon) probability, i.e. independently from the other customers in the orbit and his elapsed time in the orbit. Arrivals have precedence over departures [i.e. late arrivals], and abandonments have precedence over arrivals. The steady-state behaviour of this system is studied by the generating function method.

Two-queue systems where customers of each queue are the servers of the other queue

E. Perel and *U. Yechiali*

Abstract.– We consider systems comprised of two interlacing queues where customers of each queue are the servers of the other queue. Denoting by L_i the number of customers in queue $i = 1, 2$, we study four models: (i) Queue 1 (Q_1) operates as a multi-server limited-buffer $M(\lambda_1)/M(\mu_1)/L_2/max\{0, N-L_2\}$ system, while queue 2 (Q_2) operates as a single-server $M(\lambda_2)/M(\mu_2 L_1)/1$ queue. That is, at any moment, the L_2 customers present in Q_2 act as the servers of the limited-buffer Q_1, where service time of each individual customer is exponentially distributed with parameter μ_1. The customers of Q_2 are served by the L_1 customers in Q_1, who join hands together to form a single server with exponentially distributed service time of rate $\mu_2 L_1$. Arrivals to Q_i follow a Poisson process with rate λ_i. (ii) Q_1 operates as in model (i), but Q_2 operates as a multi-server $M(\lambda_2)/M(\mu_2)/L_1$ system. (iii) Q_1 is a limited buffer single-server $M(\lambda_1)/M(\mu_1 L_2)/1/N-1$ queue, while Q_2 is an $M(\lambda_2)/M(\mu_2 L_1)/1$ queue. (iv) Q_1 is as in model (iii), but Q_2 is an $M(\lambda_2)/M(\mu_2)/L_1$ system. For each model we derive the (conditional) probability generating function of L_1 (given L_2) and determine the condition for stability, where in models (i) and (ii) we apply Matrix Geometric analysis. We further calculate the means of L_i, along with their correlation coefficient, and compare between the models.

Matrix continued fraction approach to level-dependent QBD process and its application to multiserver retrial queues

T. Phung-Duc, *H. Masuyama*, *S. Kasahara* and *Y. Takahashi*

Abstract.– We consider the computation of the stationary distribution of a level-dependent quasi-birth-and-death (QBD) process. Based on a matrix continued fraction approach, we develop a simple algorithm to compute the rate matrices and the stationary distribution. The algorithm is easy to implement and is less memory-consuming than the conventional algorithm developed by Bright and Taylor (Stochastic Models vol. 11, pp. 497-525, 1995). In order to demonstrate the efficiency of our algorithm, we apply it to some multiserver retrial queues. Finally, we discuss an extension of our algorithm to a level-dependent Markov chain of $GI/M/1$ type.

Numerical investigation of a $Geo/PH/1$ retrial queue with non-persistent customers

S.I. Rabia

Abstract.– We consider a discrete-time retrial queue with geometric ar-

rival process and phase-type service time distribution. An arriving customer who finds the server busy may choose to depart immediately and conduct no further retrials. Orbiting customers are also non-persistent. A returning customer who cannot join the server has a positive probability of leaving the system without being served. Using the supplementary variables technique, the system is modeled as a Markov chain. By applying a sample path argument, the necessary and sufficient condition for the stability of this process is derived. Both direct truncation and generalized truncation approaches are applied to approximate the steady-state distribution of the system state and hence obtain different performance measures. A set of numerical results is presented to compare the two techniques in terms of time and accuracy. It is shown that although the idea of Neuts of Rao does not exist in the present model, applying their generalized truncation technique gives a good approximation. Emphasis is also put on the effect of non-persistence on the system performance.

Recursive equations for multiclass multiserver queueing systems

B. Rabta

Abstract.– Queues with multiple classes of jobs and multiple servers arise in many real life applications. We propose a set of equations that describe the behaviour of such systems and show the application to derive fast simulation algorithms. Networks of multiserver queues with multiple job classes are also investigated.

Importance sampling for rare events in random walks and queueing models

A. Ridder and *T. Taimre*

Abstract.– We present a method to obtain state- and time-dependent importance sampling estimators by repeatedly solving a minimum cross-entropy (MCE) program as the simulation progresses. We use this method to obtain a state- and time-dependent estimator for tail probabilities of random walks, and show logarithmic efficiency in general and strong efficiency when the jumps are Gaussian. We go on to construct an importance sampling estimator for buffer overflow problems in simple exponential queueing models. We present some numerical comparisons between our algorithms and others from the literature.

Asymptotic behavior for $MArP/PH/2$ queue with join the shortest queue discipline

Y. Sakuma

Abstract.– We are concerned with a parallel queueing model with two servers, where an arriving customer joins the shortest queue. In Sakuma, Miyazawa and Zhao (2006), we studied a $PH/M/2$ queue with join the shortest queue discipline, and obtained the tail decay rate of the stationary distribution by using the matrix analytic approach. The main objectives of this study are to generalize the preceding result, and to clarify difficulty when we apply a similar technique used in Sakuma et al. (2006). We extend the former queueing model to a $MArP/PH/2$ queue with join the shortest queue discipline, and show that the geometric tail asymptotics of the stationary distribution is obtained under a certain condition on the service time distribution.

Analysis of $MAP/G^{(a,b)}/1/N$ queue with multiple vacations and closedown times

A. Senthil Vadivu and *R. Arumuganathan*

Abstract.– In this paper, a batch service finite queue with multiple vacations and closedown times is considered. Customers arrive according to Markovian Arrival Process (MAP). On completion of a service, if the queue length is less than 'a', then the server performs a closedown work and then leaves for a vacation of random length. When the server returns from vacation and if the queue length is still less than 'a' he avails another vacation so on until the server finds 'a' customers waiting in the queue. After the completion of a service, if the number of customers in the queue is greater than 'a' then the server will continue the batch service with general bulk service rule. Customers are served in batches of maximum size 'b' with the minimum threshold value 'a'. Using supplementary variable imbedded Markov chain technique, we obtain queue length distributions at arbitrary epochs. Some key performance measures are obtained. Numerical illustration is also developed.

Approximation of $M/G/c/K$ retrial queue

Y.W. Shin

Abstract.– An approximation for the number of customers at service facility in $M/G/c/K$ retrial queue is provided with the help of the approximations of ordinary $M/G/c/K$ loss system and ordinary $M/G/c$ queue. The interpolation between two ordinary systems is used for the approximation.

Waiting time approximation in multi-class queueing systems with

multiple types of class-dependent interruptions

Ö.S. Ulusçu and *T. Altıok*

Abstract.– We consider a single-server queue subject to multiple types of operation-independent and class-dependent interruptions motivated by operations and vessel queueing at entrances of waterways. A case in point is the Istanbul Strait. We are using waiting-time arguments and service completion time analysis to obtain the expected waiting time of a customer (vessel) in the aforementioned queue with two classes of customers and k possibly simultaneous and class-dependent service interruptions. In this paper, we generalize the queueing model with multiple types of class-independent interruptions presented in the previous work and consider class-dependent service interruptions.

Rate of convergence to stationarity of the system $M/M/N/N+R$

E.A. van Doorn

Abstract.– The talk concerns the $M/M/N/N+R$ service system, characterized by N servers, R waiting positions, Poisson arrivals and exponential service times. Representations and bounds for the rate of convergence to stationarity of the number of customers in the system will be discussed, and its behaviour as a function of N and R will be described. Attention will also be paid to the setting in which the arrival rate is a function of N rather than a constant.

Priority queuing and tree-structured Markov chains

B. van Houdt, *Q.-M. He* and *J. van Velthoven*

Abstract.– In this talk we demonstrate how tree-like processes can be used to analyze a general class of priority queues with $K > 2$ service classes and possibly correlated Markovian arrivals. The key result is that the operation of a K-class priority queue can be captured by means of an alternate system that operates in a stack-like manner. The evolution of this alternate system can be reduced to a highly structured tree-like Quasi-Birth-Death (QBD) Markov process, the solution of which is realized through matrix analytic methods. The advantage of this approach, compared to solving this priority queueing system via a classical QBD, resides in the fact that the block matrices involved are of a considerably smaller size. Our approach and its efficiency is exemplified by means of a series of numerical examples. This work extends a prior paper with $K = 3$ (by B. Van Houdt and C. Blondia) and as such we will consider this special case first.

$M/D/1$, $M/D/\infty$ and stick breaking

J.S.H. van Leeuwaarden, *A.H. Lopker* and *A.J.E.M. Janssen*

Abstract.– We consider the length of a busy period in the $M/D/\infty$ queue and show that it coincides with the sojourn time of the first customer in an $M/D/1$ processor-sharing queue. We further show that the busy period is intimately related with the stationary waiting time in the $M/D/1$ first-come-first-served queue. Starting point of our analysis are three related stochastic relations that belong to the hitting time of a simple renewal age process. The main part of the talk is devoted to the discussion of some properties of the associated distribution function. We present three characterizations for the distribution function of the busy period and an asymptotic expression for its tail distribution.

Machine repair problem with variable servers considering balking concept

K.-H. Wang and *Y.-C. Liou*

Abstract.– This paper studies the machine repair problem with variable servers in which failed machines balk (do not enter) with a probability $1-b_n$. Failure and repair times of the failed machines are assumed to be exponentially distributed. A recursive method is used to derive analytic steady-state solutions through which several system performance measures can be obtained. A cost model is developed to determine the optimal values of the number of busy servers and balking rate, and maintain the balking rate at a certain level. We use the direct search method and Newton's method to find the global minimum value until the balking rate constraint is satisfied. Numerical results are provided in which various system performance measures are evaluated under optimal operating conditions. Two numerical examples are provided to illustrate Newton's method.

Dynamic control of a single-server two-class retrial queueing system

A. Winkler

Abstract.– Wireless local area networks have become popular due to ease of installation, and location freedom with the gaining popularity of laptops. Carrier sense multiple access (CSMA), in computer networking, is a network control protocol in which all participants of the network watch the status of the electrical bus and frames (messages) are only transmitted, if the bus

is detected free. There are a lot of different types of CSMA protocols in order to improve the performance. E.g. CSMA/CD (collision detection) stops transmitting a frame, when it detects another signal, and then waits a random time interval before trying to send that frame again. CSMA/CA (collision avoidance) tries to avoid collisions on the bus by waiting a random time after detecting the bus as free. Inspired by these protocols we present our model.

We consider the non-preemptive assignment of a single server to two retrial queues. Customers arrive at both queues according to Poisson processes. They are served on FIFO basis in accordance with a cost- or time-optimal routing policy. The customer at the head of the queue generates a Poisson stream of repeated requests for service. All service times are exponential with rates depending on the queues. The costs to be minimized are costs per unit time a customer spends in the system. In case of the scheduling problem a complete solution to the allocation problem is given. In case of new arrivals to the system explicit routing policy under given cost structure is presented.

Optimal resource allocation of time-reservation systems

R. Yang, *S. Bhulai* and *R. van der Mei*

Abstract.– We study the optimal resource allocation in time-reservation systems. Customers arrive at a service facility and receive service in two steps. In the first step information is gathered from the customer, which is then sent to a pool of computing resources; in the second step the information is processed after which the customer leaves the system. A central decision maker has to decide when to reserve computing power from the pool of resources, such that the customer does not have to wait for the start of the second service step and that the processing capacity is not wasted due to the customer still being serviced at the first step. The decision maker simultaneously has to decide on how many processors to allocate for the second processing step such that reservation and holding costs are minimized. We show via dynamic programming that the optimal number of processors follows a step function with as extreme policy the bang-bang control. Moreover, we provide new fundamental insight in the dependence of the optimal policy on the distribution of the information gathering times.

Exact tail asymptotics for random walks in the quarter plane

Y.Q. Zhao

Abstract.– Many two-dimensional stochastic systems, including queueing models, can be treated as random walks in the quarter plane. For such a sys-

tem satisfying an ergodic condition, tail asymptotic properties in the joint stationary distribution are very important and often lead to approximations, since an explicit expression for the distribution is usually unavailable (except a few special cases). In this talk, we introduce two newly developed methods, one based on the non-linear optimization presentation (Miyazawa) and the other based on the algebra function determined by the kernel equation and asymptotic analysis. These methods can lead to the characterization of the exact tail asymptotics in the joint (and also marginal) distribution(s), which have been found to be one of the three cases (under a certain condition or for a typical random walk).

List of Participants

Adan, I.J.B.F. (Tu-R1b – page 21)
Eindhoven University of Technology, The Netherlands
E-mail: iadan@win.tue.nl

Alcón, M.J.
Complutense University of Madrid, Spain
E-mail: mjalcon@estad.ucm.es

Altıok, T. (We-R2a – page 50)
Rutgers University, USA
E-mail: altiok@rci.rutgers.edu

Amador, J.
Complutense University of Madrid, Spain
E-mail: jamador@estad.ucm.es

Anselmi, J. (Th-R2b – page 22)
INRIA & LIG Laboratory, France
E-mail: jonatha.anselmi@gmail.com

Artalejo, J.R.
Complutense University of Madrid, Spain
Email: jesus_artalejo@mat.ucm.es

Arumuganathan, R. (Th-R2a – page 50)
PSG College of Technology, India
E-mail: ran_psgtech@yahoo.co.in

Atencia, I. (Mo-R2a – page 23)
University of Málaga, Spain
E-mail: iatencia@ctima.uma.es

Avram, F. (We-R2c – page 23)
Université de Pau, France
E-mail: florin.avram@univ-pau.fr

Baek, J.W. (We-R1b – page 23)
Sungkyunkwan University, Korea
E-mail: rainbeak@skku.edu

Baek, S. (Th-R1b – page 24)
Korea University, Korea
E-mail: sangkyubaek@korea.ac.kr

Baykal-Gursoy, M. (We-R2c – page 24)
Rutgers University, USA
E-mail: gursoy@rci.rutgers.edu

Berdjoudj, L. (We-R1d – page 25)
University of Bejaia, Algeria
E-mail: l_berdjoudj@yahoo.fr

Bosman, J.W. (We-R1b – page 26)
CWI, The Netherlands
E-mail: j.w.bosman@cwi.nl

Boudali, O. (We-R2a – page 26)
University of Athens, Greece
E-mail: olboudali@math.uoa.gr

Brill, P. (We-R1d – page 27)
University of Windsor, Canada
E-mail: brill@uwindsor.ca

Bruneel, H.
Ghent University, Belgium
E-mail: hb@telin.ugent.be

Chakravarthy, S.R. (Mo-R1b – page 27)
Kettering University, USA
E-mail: schakrav@kettering.edu

Chaudhry, M.
Royal Military College of Canada, Canada
E-mail: chaudhry-ml@rmc.ca

Choi, B.D.
Korea University, Korea
E-mail: queue@korea.ac.kr

Claeys, D. (We-R2b – page 29)
Ghent University, Belgium
E-mail: dieter.claeys@telin.ugent.be

D'Auria, B. (We-R1b – page 29)
Carlos III University of Madrid, Spain
E-mail: bernardo.dauria@uc3m.es

De Clercq, S. (Mo-R2a – page 30)
Ghent University, Belgium
E-mail: sdclercq@telin.ugent.be

De Vuyst, S. (Mo-R2a – page 31)
Ghent University, Belgium
E-mail: sdv@telin.ugent.be

Demoor, T. (We-R1a – page 31)
Ghent University, Belgium
E-mail: thdemoor@telin.ugent.be

Dendievel, S.
Université Libre de Bruxelles, Belgium
E-mail: sdendiev@ulb.ac.be

Dorsman, J.L. (Mo-R1a – page 32)
CWI & Vrije Universiteit Amsterdam, The Netherlands
E-mail: j.l.dorsman@cwi.nl

Dudin, A.N. (Tu-R2a – page 33)
Belarusian State University, Belarus
E-mail: dudin@bsu.by

Economou, A. (Th-R1a – page 33)
University of Athens, Greece
E-mail: aeconom@math.uoa.gr

Efrosinin, D. (Tu-R2b – page 34)
Johannes Kepler University, Austria
E-mail: dmitry.efrosinin@jku.at

Eliazar, I. (We-R2d – page 34)
Holon Institute of Technology, Israel
E-mail: eliazar@post.tau.ac.il

Feinberg, E.A. (Tu-R1a – page 35)
Stony Brook University, USA
E-mail: eugene.feinberg@sunysb.edu

Feng, L.
Sungkyunkwan University, Korea
E-mail: fengliyan@skku.edu

Fiems, D. (Th-R1b – page 36)
Ghent University, Belgium
E-mail: dieter.fiems@ugent.be

Fortes, I. (Tu-R1b – page 22)
University of Málaga, Spain
E-mail: ifortes@uma.es

Gai, J. (We-R2b – page 28)
Royal Military College of Canada, Canada
E-mail: jing.gai@rmc.ca

Gandhi, A. (Tu-R1b – page 36)
Carnegie Mellon University, USA
E-mail: anshulg@cs.cmu.edu

Gómez-Corral, A.
Complutense University of Madrid, Spain
E-mail: antonio_gomez@mat.ucm.es

Govorun, M. (Tu-R2a – page 38)
Université Libre de Bruxelles, Belgium
E-mail: mgovorun@ulb.ac.be

Harchol-Balter, M.
Carnegie Mellon University, USA
E-mail: nstenger@cs.cmu.edu

Hasenbein, J. (We-R1c – page 38)
University of Texas at Austin, USA
E-mail: jhas@mail.utexas.edu

Huang, M.L.
Brock University, Canada
E-mail: mhuang@brocku.ca

Jin, S. (We-R1a – page 39)
Korea University, Korea
E-mail: irarykim@gmail.com

Kapodistria, S. (Th-R1a – page 40)
University of the Aegean, Greece
E-mail: stellakap@aegean.gr

Kasahara, S.
Kyoto University, Japan
E-mail: kasahara@ieee.org

Kim, J.
Chungbuk National University, Korea
E-mail: jeongsimkim@chungbuk.ac.kr

Kim, J. (Mo-R2b – page 40)
Korea University, Korea
E-mail: b1155@hotmail.net

Kim, T.O. (We-R1c – page 41)
Korea University, Korea
E-mail: violetgl@korea.ac.kr

Kobayashi, M. (Th-R2a – page 41)
Tokyo University of Science, Japan
E-mail: j6308702@ed.noda.tus.ac.jp

Krishna Kumar, B. (Tu-R1b – page 42)
Anna University, India
E-mail: drbkkumar@hotmail.com

Krishnamoorthy, A. (We-R2a – page 42)
Cochin University of Science and Technology, India
E-mail: achyuthacusat@gmail.com

Laevens, K. (Mo-R2b – page 43)
Ghent University, Belgium
E-mail: kl@telin.ugent.be

Latouche, G. (Mo-R1b – page 43)
Université Libre de Bruxelles, Belgium
E-mail: guy.latouche@ulb.ac.be

Lee, S.W.
Sungkyunkwan University, Korea
E-mail: swlee94@skku.edu

Lewis, M. (Th-R2b – page 21)
Cornell University, USA
E-mail: mark.lewis@cornell.edu

López García, M.
Complutense University of Madrid, Spain
E-mail: martin_lopez@mat.ucm.es

López-Herrero, M.J.
Complutense University of Madrid, Spain
E-mail: lherrero@estad.ucm.es

Lopker, A.H. (We-R2c – page 52)
Eindhoven University of Technology, The Netherlands
E-mail: lopker@eurandom.tue.nl

Masuyama, H.
Kyoto University, Japan
E-mail: masuyama@sys.i.kyoto-u.ac.jp

Mitrani, I. (Tu-R2b – page 44)
Newcastle University, U.K.
E-mail: isi.mitrani@ncl.ac.uk

Miyazawa, M. (Opening Session – page 45)
Tokyo University of Science, Japan
E-mail: miyazawa@is.noda.tus.ac.jp

Moosa, S. (Th-R2b – page 45)
Singapore Management University, Singapore
E-mail: sharafalim@smu.edu.sg

Moreno, P. (Th-R2a – page 21)
Pablo de Olavide University, Spain
E-mail: mpmornav@upo.es

Nazarathy, Y. (Th-R1b – page 39)
Eurandom, The Netherlands
E-mail: nazarathy@eurandom.tue.nl

Nobel, R. (Th-R1a – page 47)
Vrije Universiteit Amsterdam, The Netherlands
E-mail: rnobel@feweb.vu.nl

Pacheco, A. (We-R2b – page 35)
IST – Technical University of Lisbon, Portugal
E-mail: apacheco@math.ist.utl.pt

Perros, H. (Tu-R1a – page 46)
North Carolina State Uiversity, USA
E-mail: hgperros@gmail.com

Phung-Duc, T. (Th-R1a – page 48)
Kyoto University, Japan
E-mail: tuan@sys.i.kyoto-u.ac.jp

Rabia, S.I. (Tu-R2b – page 48)
Alexandria University, Egypt
E-mail: shrfrabia@hotmail.com

Rabta, B. (Mo-R2b – page 49)
University of Neuchâtel, Switzerland
E-mail: boualem.rabta@unine.ch

Resing, J.A.C. (Th-R2a – page 24)
Eindhoven University of Technology, The Netherlands
E-mail: resing@win.tue.nl

Ridder, A. (We-R2d – page 49)
Vrije Universiteit Amsterdam, The Netherlands
E-mail: aridder@feweb.vu.nl

Rogiest, W. (We-R1b – page 30)
Ghent University, Belgium
E-mail: wrogiest@telin.ugent.be

Sakuma, Y. (Mo-R1b – page 49)
Tokyo University of Science, Japan
E-mail: sakuma-y@rs.noda.tus.ac.jp

Sánchez, S.
University of Málaga, Spain
E-mail: sixto@uma.es

Shin, Y.W. (Tu-R2b – page 50)
Changwon National University, Korea
E-mail: ywshin@changwon.ac.kr

Squillante, M.S. (We-R1c – page 37)
IBM Thomas J. Watson Research Center, USA
E-mail: mss@watson.ibm.com

Stadje, W.
University of Osnabrueck, Germany
E-mail: wolfgang@mathematik.uni-osnabrueck.de

Steyaert, B. (Th-R1b – page 46)
Ghent University, Belgium
E-mail: bs@telin.ugent.be

Telek, M. (We-R1a – page 27)
Technical University of Budapest, Hungary
E-mail: telek@hit.bme.hu

Van der Mei, R. (Th-R2b – page 53)
CWI & Vrije Universiteit Amsterdam, The Netherlands
E-mail: mei@cwi.nl

Van Doorn, E.A. (Opening Session – page 51)
University of Twente, The Netherlands
E-mail: e.a.vandoorn@utwente.nl

Van Houdt, B. (Tu-R2a – page 51)
University of Antwerp, Belgium
E-mail: benny.vanhoudt@ua.ac.be

Van Wijk, A.C.C. (Mo-R1a – page 25)
Eindhoven University of Technology, The Netherlands
E-mail: a.c.c.v.wijk@tue.nl

Wang, K.-H. (We-R2a – page 52)
National Chung-Hsing University, Taiwan
E-mail: khwang@amath.nchu.edu.tw

Winkler, A. (Tu-R1a – page 52)
Johannes Kepler University, Austria
E-mail: anastasia.winkler@jku.at

Wittevrongel, S.
Ghent University, Belgium
E-mail: sw@telin.ugent.be

Yang, X. (We-R2b – page 28)
Royal Military College of Canada, Canada
E-mail: xiaofeng.yang@rmc.ca

Yechiali, U. (Mo-R1a – page 47)
Tel-Aviv University & Afeka College of Engineering, Israel
E-mail: uriy@post.tau.ac.il

Zhao, Y.Q. (We-R1a – page 53)
Carleton University, Canada
E-mail: zhao@math.carleton.ca