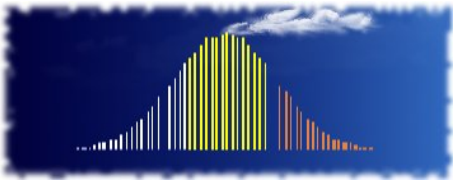
17th

informs



Applied Probability Conference

July 15-17, 2013

Program

Monday July 15, 2013

	Track 1	Track 2	Track 3	Track 4	Track 5	Track 6	Track 7	Track 8	Track 9
Room	JVdC A	JVdC B	JVdC C	JVdC D	JVdC E	JVdC F	JVdC G	JVdC H	Cabildo I
9:00am - 9:15am	**Opening Remarks** (Room: Juan Vasquez de Coronado D-E)								
9:15am - 10:15am	**Plenary - Onesimo Hernandez-Lerma** (Room: Juan Vasquez de Coronado D-E)								
MA 10:45am - 12:15pm	*Random Graph Models*	*Green Scheduling 1*	*Advanced Applications of Simulation*	*AP in Finance, Economics and Insurance II*	*Limit results for queueing and related models*	*Heavy tails and related topics*	*Markov Chains and MDPs*	*Statistical modelling and analysis*	*Queues and Health*
	Moyal	**Righter / Down**	**Fu / Henderson / Zhou**	**Cadenillas**	**Atar**	**Roy / Maulik**	**Arlotto / Steele**	**Yi**	**Örmeci**
	Moyal Chen Lelarge	Gandhi Harchol-Balter Adan Down	Dieker Lim Ramanan Glasserman	Blanchet Melnikov Choulli Seco	Bramson Gurvich Shwartz Atar	Das Grabchak Hazra	Fill Lyzinski Arlotto Feinberg	Wong Li Liu Yi	Niyirora Pender Örmeci
12:15pm - 2.00pm	**Lunch**								
MB 2:00pm - 3:30pm	*Random Graphs: Epidemics*	*Green Scheduling 2*	*Rare-Event Simulation*	*AP in Finance, Economics and Insurance I*	*Asymptotic Analysis of Stochastic Networks*	*Stationary analysis of Markov chains*	*Markov Decision Processes*	*Network Routing*	*Queues I*
	Hasenbein	**Righter / Down**	**Fu / Henderson / Zhou**	**Cadenillas**	**Ramanan**	**Adan**	**Feinberg**	**Honnappa**	**Ding**
	Hasenbein Lelarge Van Leeuwaarden Britton	Gopalakrishnan Squillante Hyytiä	Liu Botev Rojas-Nandayapa Blanchet	Cadenillas Frei Pérez-Hernández Choi	Ward Kruk Dieker Ramanan	Kulkarni Kapodistria Smit Fralix	Canbolat Spieksma Feinberg Katehakis	Ziedins Kardeş Hampshire Honnappa	Xu He Büke Ding
MC 4:00pm - 5:30pm	*Random Graphs: Algorithms*	*Green Scheduling 3*	*Simulation: Estimation*	*Optimal Stopping and American Options*	*Systems with parallel skilled-based service*	*Matrix Analytic Methods*	*Markov Decision Processes II*	*Networks*	*Queues II*
	Litvak	**Righter / Down**	**Fu / Henderson / Zhou**	**Ludkovski**	**Adan**	**Van Houdt**	**Van de Ven**	**Lewis**	**Shin**
	Litvak Reittu Spieksma	Wang Kulkarni Gautam	Nakayama Vázquez-Abad Pasupathy Glynn	Hu Ludkovski Yao	Harchol-Balter Mairesse Ward Weiss	Telek Nielsen He Block	Blok Jacobson Fernandez Van de Ven	Kunnumkal Chen Lewis	Yom-Tov Xu Ertiningsih Shin
6:30pm - 8:00pm	**Reception, Small Town Costa Rican "Fiesta"**								

Contents

1

Program Committee and Organizing Committee

General Conference Chair:
Emmanuel Fernandez
University of Cincinnati
emmanuel.fernandez@uc.edu

Program Committee Chair:
John Hasenbein
University of Texas at Austin
jhas@mail.utexas.edu

Program Committee Co-Chairs:
Ivo Adan
Eindhoven University of Technology, The Netherlands
i.adan@tue.nl

Emmanuel Fernandez
University of Cincinnati
emmanuel.fernandez@uc.edu

Marko Boon
Eindhoven University of Technology, The Netherlands
marko@win.tue.nl

Local Arrangements Committee:
Chair:
Jorge Romero
Universidad de Costa Rica
jromero@eie.ucr.ac.cr

Co-chair:
Edgardo Vargas
Instituto Tecnologico Costa Rica - San Carlos
edvargas@itcr.ac.cr

Program Committee

Name	Affiliation
Ivo Adan	Eindhoven University of Technology
Allesandro Arlotto	Duke University
Rami Atar	Technion
Kostia Avrachenkov	INRIA
Moshen Bayati	Stanford University
Tom Britton	Stockholm University
Abel Cadenillas	University of Alberta
Pavel Chigansky	The Hebrew University of Jerusalem
Bernardo d'Auria	University Carlos III de Madrid
Doug Down	McMaster University
Eugene Feinberg	SUNY Stony Brook
Emmanuel Fernandez	University of Cincinnati
Michael Fu	University of Maryland
Natarajan Gautam	Texas A&M
Samim Ghamami	UC Berkeley
David Goldberg	Georgia Tech
Varun Gupta	University of Chicago
Itai Gurvich	Northwestern
Robert Hampshire	Carnegie Mellon University
John Hasenbein	University of Texas at Austin
Shane Henderson	Cornell University
Vladimir Kaishev City	University London
Stella Kapodistria	Eindhoven University of Technology
N. Bora Keskin	The University of Chicago
Jeff Kharoufeh	University of Pittsburgh
Henry Lam	Boston University
Nelly Litvak	University of Twente
Yingdong Lu	IBM Watson
Mike Ludkovski	UC Santa Barbara
Avi Mandelbaum	Technion
Krishanu Maulik	Indian Statistical Institute
Pascal Moyal	Université de Technologie de Compiègne
Yoni Nazarathy	The University of Queensland
Giang Nguyen	The University of Adelaide
Mariana Olvera	Columbia University
Zbyszek Palmowski	University of Wroclaw
Kavita Ramanan	Brown University
Rhonda Righter	UC Berkeley
Parthanil Roy	Michigan State University
Mark Squillante	IBM Watson
J. Michael Steele	The University of Pennsylvania
Peter Taylor	University of Melbourne
Benny van Houdt	University of Antwerpen
Johan van Leeuwaarden	Eindhoven University of Technology
Bo Zhang	IBM Watson
Hanqin Zhang	University of Singapore
Enlu Zhou	University of Illinois

2

Sponsors

The organizers gratefully acknowledge the support of the following organizations:

Universidad de Costa Rica

Co-sponsored by the INFORMS Applied Probability Society

3

General Information

Location and Conference Buildings

The conference will take place at the Costa Rica Marriott Hotel, San Jose. A ploor plan can be found on the next page.

(Pre-)Registration and Conference Office

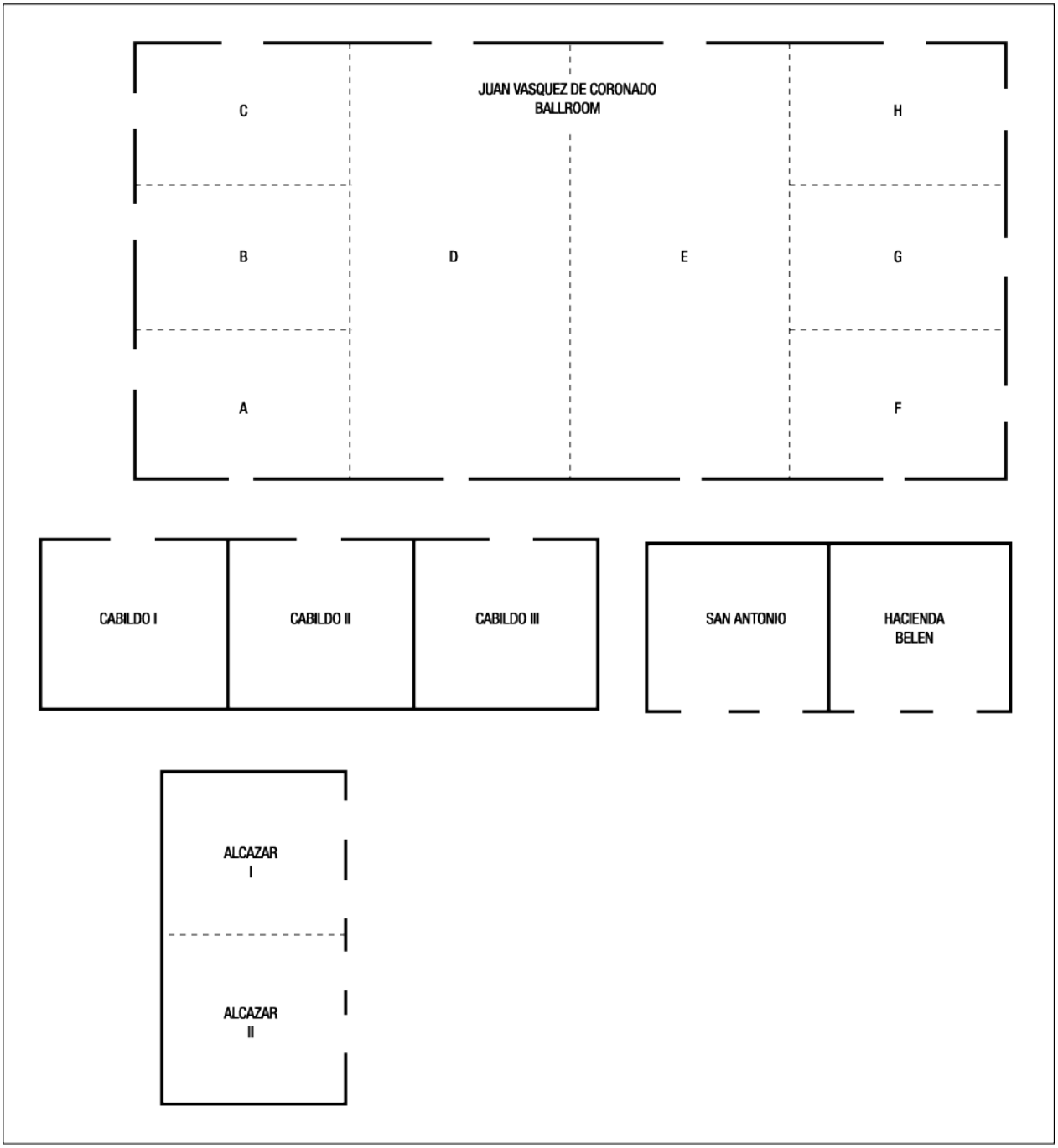
The conference registration and hospitality desk will open on Sunday July 14, from 3:00-8:00 pm, and Monday-Wednesday from 8:00 am - 6:00 pm.

Refreshments and Meals

Coffee, tea and refreshments will be served during the breaks and before the first session of the day.

Applied Probability Society Business Meeting

Time and date: Monday July 16, 6:00pm – 6:45pm.
Location: Cabildo I.



Floor plan Costa Rica Marriott Hotel

4

Program at a Glance

Opening, plenaries, and tutorials take place in Juan Vasquez de Coronado (JVdC) D-E.

Welcome reception takes place in the Pool Area.

APS Business Meeting takes place in Cabildo I.

The banquet on Tuesday takes place in Juan Vasquez de Coronado (JVdC) D-E.

Monday July 15, 2013

9:00am	–	09:15am	Opening remarks
9:15am	–	10:15am	Plenary - Onesimo Hernandez-Lerma
10:45am	–	12:15pm	Parallel Sessions MA
12:15pm	–	2:00pm	Lunch
2:00pm	–	3:30pm	Parallel Sessions MB
4:00pm	–	5:30pm	Parallel Sessions MC
6:30pm	–	8:30pm	Reception, Small Town Costa Rican “Fiesta”

Tuesday July 16, 2013

8:30am	–	10:00am	Parallel Sessions TA
10:15am	–	11:15am	Tutorial - Bill Massey
11:30am	–	1:00pm	Parallel Sessions TB
1:00pm	–	2:15pm	Lunch
2:15pm	–	3:45pm	Parallel Sessions TC
4:00pm	–	5:00pm	Tutorial - Tsachy Weissman
6:00pm	–	6:45pm	APS Business Meeting
6:30pm	–	7:30pm	Refreshments
7:30pm	–	9:30pm	Banquet (with musical entertainment)

Wednesday July 17, 2013

8:30am	–	9:30pm	Plenary - Paul Glasserman
10:00am	–	11:30am	Parallel Sessions WA
11:45am	–	1:15pm	Lunch
1:30pm	–	7:00pm	Social event: Visit to the Gold Museum

Rooms

Track 1
JVdC A

Track 2
JVdC B

Track 3
JVdC C

Track 4
JVdC D

Track 5
JVdC E

Track 6
JVdC F

Track 7
JVdC G

Track 8
JVdC H

Track 9
Cabildo I

5

Plenaries and Tutorials

All plenaries and tutorials take place in room Juan Vasquez de Coronado D-E (see the floor plan on page 10).

Monday 9:15am – 10:15am Plenary

Stochastic Dynamic Games

Onésimo Hernández-Lerma, Mathematics Department, CINVESTAV–IPN, México City

This talk is an introduction to stochastic dynamic games and some of their applications. It includes cooperative and noncooperative games, and some important special cases such as compromise solutions, zero–sum games, and games against nature (also known as minimax or worst–case control problems). It also includes recent results on the existence and characterizations of dynamic potential games. One of these characterizations is particularly interesting because it identifies a class of dynamic potential games in which Nash (or noncooperative) equilibria coincide with Pareto (or cooperative) equilibria. This latter fact is not very common.

Tuesday 10:15am – 11:15am Tutorial

Variational calculus for dynamical system optimization with dynamic rate queueing applications

W.A. Massey, Princeton University, United States of America, wmassey@princeton.edu

The dynamic optimization of constrained dynamical systems is based on the calculus of variations and its various incarnations over the past three centuries. Various fluid and diffusion limit theorems give us many dynamic rate queueing systems that are well approximated by related low dimensional dynamical systems.

This is a tutorial that discusses these optimization methods by introducing the concepts of Lagrangians, Hamiltonians, and Bellman value functions. The performance and management of services derived from the leasing of shared resources modeled by queueing systems provide us with many examples that apply and illustrate these optimizing techniques.

Joint work with R.C. Hampshire, Carnegie Mellon University, United States of America, hamp@cmu.edu.

Tuesday 4:15pm – 5:15pm Tutorial

Some relations between Information and Estimation

Tsachy Weissman, Stanford University, tsachy@stanford.edu

I will give a tour through a sparse sample of the information theory literature - both classical and recent - on relations between information and estimation. Beyond aesthetic value, these relations underlie some of the main tools in Shannon theory, such as the Entropy Power Inequality. They also give considerable insight into and a quantitative understanding of several estimation theoretic objects, such as the costs of causality and of mismatch, as well as the performance and structure of minimax estimators. Further, they enable the transfer of analytic tools and algorithmic know-how from one domain to another. Examples will be given to illustrate these points.

Wednesday 8:30am – 9:30pm

Financial Networks and Systemic Risk

Paul Glasserman, Columbia University, New York, pg20@columbia.edu

What is the nature and magnitude of network effects in destabilizing the financial system? Versions of these questions have generated a large and growing literature on network models of systemic risk. A significant challenge in applying these models is the lack of public data on interconnections between financial institutions. In this talk, after reviewing some background, I will present results from joint work with Peyton Young (Oxford) in which we bound the magnitude of network effects using information about individual nodes but no information about network topology. We illustrate the results using European bank data. Time permitting, I will discuss other aspects of systemic risk as well.

6

Sessions

Monday 10:45am - 12:15pm

Session MA1 - Random Graph Models, Chair: Moyal in JVdC A, see page 31

1. CSMA algorithms for large random graphs
   Moyal, P.; Bermolen, P.; Jonckheere, M.

2. Directed random graphs with given degree distributions
   Chen, N.; Olvera-Cravoito, M.

3. Community Detection in the Labelled Stochastic Block Model
   Lelarge, M.; Heimlicher, S.; Massoulié, L.; Xu, J.

Session MA2 - Green Scheduling 1, Chair: Righter / Down in JVdC B, see page 32

1. How queueing-theoretic models for data center power management differ from reality
   Gandhi, A.; Harchol-Balter, M.; Kozuch, M.A.

2. The RRR Method for the Exact Analysis of Repeating Markov Chains with Uni-Directional Transitions
   Gandhi, A.; Doroudi, S.; Harchol-Balter, M.; Scheller-Wolf, A.

3. Optimization of Server Farms
   Adan, I.J.B.F.; Kulkarni, V.G.; Van Wijk, A.C.C.

4. On Optimal Policies for Energy Aware Servers
   Down, D.G.; Maccio, V.

Session MA3 - Advanced Applications of Simulation, Chair: Fu / Henderson / Zhou in JVdC C, see page 33

1. On simulation of constants from the theory of Gaussian processes
   Dieker, A.B.; Yakir, B.

2. On a Least Absolute Deviations Estimator of a Convex Function
   Lim, E.; Luo, Y.

3. Computing volumes of convex bodies
   Gamarnik, D.; Ramanan, K.

4. Bounding Wrong-Way Risk by Simulation
   Glasserman, P.; Yang, L.

Session MA4 - Applied Probability in Finance, Economics and Insurance II, Chair: Cadenillas in JVdC D, see page 34

1. Heavy Traffic Limits and Applications to Insurance and Pension Fund Dynamics
   Blanchet, J.H.; Chen, X.; Lam, H.
2. On Polynomial Technique Applied To Financial And Actuarial Modeling
   Melnikov, A.
3. Non-Arbitrage for a Class of Informational Markets
   Choulli, T.; Aksamit, A.; Deng, J.; Jeanblanc, M.
4. Stochastic correlation in asset management
   Seco, L.A.; Reuss, A.; Olivares, P.; Zagst, R.

Session MA5 - Limit results for queueing and related models, Chair: Atar in JVdC E, see page 35

1. Join the Shortest Queue Networks: Their Limiting Behavior and Open Problems
   Bramson, M.; Lu, Yi; Prabhakar, B.
2. Diffusion models and steady-state approximations for exponentially ergodic Markovian queues
   Gurvich, I.
3. Asymptotics of epidemic-like models with applications to peer-to-peer networks
   Shwartz, A.; Altman, E.; Nain, P.; Xu, Y.
4. Three new state space collapse results for the multiclass G/G/1 queue
   Atar, R.; Biswas, A.; Lev-Ari, A.; Shifrin, M.

Session MA6 - Heavy tails and related topics, Chair: Roy / Maulik in JVdC F, see page 36

1. Fine-tuning risk assessment through hidden regular variation
   Das, B.; Mitra, A.; Resnick, S.
2. Properties Of Tempered Stable Distributions
   Grabchak, M.
3. Applications of Abelian and Tauberian theorems to free probability
   Hazra, R.S.; Maulik, K.

Session MA7 - Markov Chains and Markov Decision Problems, Chair: Arlotto / Steele in JVdC G, see page 37

1. Comparison Inequalities and Fastest-mixing Markov Chains
   Fill, J.A.; Kahn, J.
2. Strong Stationary Duality for Diffusions
   Lyzinski, V.; Fill, J.A.
3. Optimal On-Line Selection of an Alternating Subsequence: A Central Limit Theorem
   Arlotto, A.; Steele, J.M.
4. Optimality Conditions for Total-Cost Partially Observable Markov Decision Processes with General State and Action Spaces
   Feinberg, E.A.; Kasyanov, P.O.; Zgurovsky, M.Z.

Session MA8 - Statistical modelling and analysis for complex data, Chair: Yi in JVdC H, see page 39

1. Nature-Inspired Metaheuristic Algorithms for Finding Efficient Experimental Designs
   Wong, W.K.
2. Response adaptive randomization in the presence of mismeasurement
   Li, X.
3. Bias analysis for misclassification in a multicategorical exposure in a logistic regression model
   Liu, Y.; Liu, J.
4. The most powerful test and the optimal design of response adaptive clinical trials
   Yi, Y.

Session MA9 - Queues and Health, Chair: Örmeci in Cabildo I, see page 40

1. Nursing Home Bed Planning with Medicaid Bed-Hold Payment Policies
   Jennings, O.B.; Massey, W.A.; Niyirora, J..
2. Dynamic Rate Coupled Processors with Abandonment: Approximation and Computation
   Pender, J.
3. Length-of-Stay and Optimal Portfolio of Surgical Procedures
   Örmeci, L.; Bavafa, H.; Savin, S.

Monday 2:00pm - 3:30pm

Session MB1 - Random Graphs and Complex Networks: Epidemics, Chair: Hasenbein in JVdC A, see page 41

1. Rapid Detection of Viruses in Large Contact Networks
   Hasenbein, J.; Lee, J.; Morton, D.
2. Diffusions and cascades in random networks
   Lelarge, M.; Coupechoux, E.
3. Critical random graphs and epidemics
   Van Leeuwaarden, J.S.H.; Aidekon, E.; Van der Hofstad, R.W.; Janssen, A.J.E.M.; Kliem, S.; Simatos, F.
4. Dynamic networks in growing populations and epidemics thereon
   Britton, T.; Turova, T.; Trapman, P.

Session MB2 - Green Scheduling 2, Chair: Righter / Down in JVdC B, see page 42

1. Scheduling and staffing strategic servers
   Doroudi, S.; Gopalakrishnan, R.; Ward, A.R.; Wierman, A.
2. Stochastic Optimal Control for a Class of Dynamic Resource Allocation Problems
   Squillante, M.S.; Gao, X.; Lu, Y.; Sharma, M.; Bosman, J.W.
3. Task Assignment in a Server Farm with Switching Delays and General Energy-Aware Cost Structure
   Hyytiä, E.; Righter, R.; Aalto, S.

Session MB3 - Rare-Event Simulation, Chair: Fu / Henderson / Zhou in JVdC C, see page 43

1. Rare-event Simulations for Random Differential Equations
   Liu, J.
2. Rare-event probability Estimation via Empirical Likelihood
   Botev, Z.I.
3. Rare event simulation for the tail probability of the maximum of dependent random variables
   Rojas-Nandayapa, L.
4. On Stochastic Approximations, Quasi-stationary Distributions, and Detection of Moving Targets
   Blanchet, J.H.; Glynn, P.W.; Zheng, S.

Session MB4 - Applied Probability in Finance, Economics and Insurance I, Chair: Cadenillas in JVdC D, see page 44

1. Optimal Government Debt Control: Explicit Formula for the Optimal Debt Ceiling
   Cadenillas, A.; Huaman, R.
2. Optimal Execution of a VWAP Order: a Stochastic Control Approach
   Frei, C.
3. Characterization of the minimal penalty of a convex risk measure with applications to robust utility maximization problem for a market model based on Lévy processes.
   Pérez-Hernández, L.
4. On Sannikov's principal-agent problem
   Choi, K.

Session MB5 - Asymptotic Analysis of Stochastic Networks, Chair: Ramanan in JVdC E, see page 45

1. On the dynamic control of matching queues
   Gurvich, I.; Ward, A.R.
2. A free boundary problem arising in order book dynamics
   Kruk, Ł.; Ramanan, K.
3. From local to global stability in stochastic processing networks through quadratic Lyapunov functions
   Dieker, A.B.; Shin, J.
4. Talk 4: Large deviations of mean-field interacting particle systems
   Dupuis, P.; Ramanan, K.; Wu, W.

Session MB6 - Advances in stationary analysis of Markov chains, Chair: Adan in JVdC F, see page 46

1. Optimal routing of customers in polling systems
   Kulkarni, V.G.; Lee, N.; Adan, I.J.B.F.; Lefeber, A.A.J.
2. Stationary analysis of the shortest queue polling model
   Adan, I.J.B.F.; Boxma, O.J.; Kapodistria, S.; Kulkarni, V.G.
3. Explicit Solutions and Other Properties of Successively Lumpable Quasi Skip Free Processes
   Smit, L.C.; Spieksma, F.M.; Katehakis, M.N.

4. A New Random-Product Representation For The Stationary Distribution Of A Markov Chain
   Buckingham, P.; Fralix, B.

Session MB7 - Markov Decision Processes, Chair: Feinberg in JVdC G, see page 47

1. Markov Population Decision Chains with Constant Risk Posture
   Canbolat, P.G.
2. Convergence properties of approximating Markov Decision Processes with unbounded jump rates
   Spieksma, F.M.
3. Sufficiency of Markov Policies for Continuous-Time Markov Decision Processes
   Feinberg, E.A.; Mandava, M.; Shiryaev, A.N.
4. Cash-Flow Based Dynamic Inventory Management
   Katehakis, M.N.

Session MB8 - Network Routing, Chair: Honnappa in JVdC H, see page 48

1. Self-Optimising State-Dependent Routing In Parallel Queues With Batch Servers
   Ziedins, I.
2. Wait or share service?
   Customers' choice at the equilibrium
   Kardeş, E.
3. Is the Curb 80%Full or 20%Empty? Analysis of San Francisco's Parking Experiment
   Hampshire, R.C.
4. Timing and Routing Games in Transitory Generalized Jackson Networks
   Honnappa, H.; Jain, R.

Session MB9 - Queues I, Chair: Ding in Cabildo I, see page 50

1. Queueing System Topologies with Limited Flexibility
   Tsitsiklis, J.N.; Xu, K.
2. A one-dimensional diffusion model for overloaded queues with customer abandonment
   He, S.
3. Stabilizing policies for probabilistic matching systems
   Büke, B.; Chen, H.; Rasónyi, M.
4. A Fluid Model for Overloaded Queues with Scoring-Based Priority Rules
   Ding, Y.; Glynn, P.

Monday 4:00pm - 5:30pm

Session MC1 - Random Graphs and Complex Networks: Algorithms, Chair: Litvak in JVdC A, see page 51

1. Ranking algorithms on directed random networks
   Chen, N.; Litvak, N.; Olvera-Cravoito, M.

2. Stochastic network models inspired by Szemerédi's Regularity Lemma, with applications to big data analysis and compression
   Reittu, H.; Bazsó, F.; Norros, I.

4. Taboos in the computation of the Effective Graph Resistance
   Spieksma, F.M.

Session MC2 - Green Scheduling 3, Chair: Righter / Down in JVdC B, see page 52

1. Green Computing by Putting Idle Servers in Sleep
   Wang, S.; Chen, J.; Liu, X.

2. Energy Efficient Virtual Appliance Deployment Framework
   Lee, N.; Kulkarni, V.

3. Efficiently Operating Energy-limited Wireless Nodes
   Mohapatra, A.; Gautam, N.

Session MC3 - Simulation: Estimation, Chair: Fu / Henderson / Zhou in JVdC C, see page 53

1. Confidence Intervals For Quantiles
   Nakayama, M.K.; Calvin, J.M.

2. Statistical bias correction for stochastic optimization
   Vázquez-Abad, F.

3. Sampling within Algorithmic Recursions
   Pasupathy, R.; Hashemi, F.; Ghosh, S.

4. Exact Estimation vs Exact Simulation
   Glynn, P.W.; Rhee, C.H.

Session MC4 - Optimal Stopping and American Options, Chair: Ludkovski in JVdC D, see page 54

1. Snell envelope with small probability criteria
   Hu, P.

2. Sequential Regression Methods for Optimal Stopping
   Ludkovski, M.

3. Robust Optimal Stopping under Volatility Uncertainty
   Yao, S.; Bayraktar, E.

Session MC5 - Systems with parallel skilled-based service, Chair: Adan in JVdC E, see page 55

1. Matching Queues with Flexible Servers
   Harchol-Balter, M.; Gardner, K.; Doroudi, S.; Scheller-Wolf, A.

2. Dynamic bipartite matching models
   Bušić, A.; Gupta, V.; Mairesse, J.

3. Routing to Minimize Waiting and Callbacks in Large Call Centers
   Ward, A.R.; Zhan, D.

4. Structure of FCFS infinite matching
   Weiss, G.; Adan, I.J.B.F.

Session MC6 - Matrix Analytic Methods, Chair: Van Houdt in JVdC F, see page 56

1. A constructive proof of {PH
   Horvath, I.; Telek, M.

2. Queueing models with matrix-exponential distributions and rational arrival processes
   Nielsen, B.F.; Bean, N.

3. Queues with Customer Interjections
   He, Q.M.; Chavoushi, A.A.

4. Delay and Energy Efficiency of Tree Algorithms with Free Access
   Block, R.; Van Houdt, B.

Session MC7 - Markov Decision Processes II, Chair: Van de Ven in JVdC G, see page 57

1. Monotonic successive approximations in queueing systems
   Blok, H.; Bhulai, S.; Spieksma, F.M.

2. Stochastic Sequential Assignment Problem with Threshold Criteria
   Jacobson, S.H.; Baharian, G.

3. Simulation-Based Approximate Dynamic Programming Approaches for Semiconductor Manufacturing Operations
   Chen, X.; Fernandez, E.; Kelton, W.D.; Ramirez-Hernandez, J.A.

4. Modeling battery energy storage
   Harsha, P.; Petrik, M.; Van de Ven, P.M.

Session MC8 - Networks, Chair: Lewis in JVdC H, see page 59

1. Randomization Approaches for Network RM with Choice Behavior
   Kunnumkal, S.

2. Necessary conditions for the invariant measure of a random walk to be a sum of geometric terms
   Chen, Y.; Boucherie, R.J.; Goseling, J.

3. The Independence Number of a Maximal Outerplanar Graph
   Lewis, T.M.

Session MC9 - Queues II, Chair: Shin in Cabildo I, see page 59

1. Staffing Service Systems with Load Dependent Service Rate
   Dong, J.; Feldman, P.; Yom-Tov, G.B.

2. Queuing with Future Information
   Spencer, J.; Sudan, M.; Xu, K.

3. A Quasi-Skip-Free processes
   with "quasi" product form stationary distribution
   Ertiningsih, D.; Spieksma, F.M.; Smit, L.

4. Approximation of serial lines with multiple servers and finite buffer
   Shin, Y.W.; Moon, D.H.; Lee, H.S.

Tuesday 8:30am - 10:00am

Session TA1 - Random Graphs and Complex Networks: Connectivity, Chair: Britton in JVdC A, see page 61

1. Giant Component in Random Multipartite Graphs with Given Degree Sequences
   Misra, S.; Gamarnik, D.

2. Retransmission Delays over Correlated Channels
   Jelenković, P.R.; Skiani, E.

3. The largest component of a hyperbolic model of complex networks
   Müller, T.

4. Degree-degree dependencies in random graphs with heavy-tailed degrees
   Litvak, N.

Session TA2 - Applied Probability in Healthcare, Chair: Bayati in JVdC B, see page 62

1. A Dynamic Random Graph Model for Kidney Exchange
   Anderson, R.; Ashlagi, I.; Gamarnik, D.; Kanoria, Y.

2. Donor-Dependent Scoring Schemes: Shaping the Allocation of Cadaver Kidneys in a New Era
   Ata, B.; Ding, Y.; Zenios, S.

3. Queues with Delay Sensitive Service Times
   Chan, C.W.; Farias, V.F.; Escobar, G.

4. Active Postmarketing Drug Surveillance for Multiple Adverse Events
   Goh, J.; Zenios, S.

Session TA3 - Simulation Optimization 1, Chair: Fu / Henderson / Zhou in JVdC C, see page 64

1. Interactive Model-based Search for Global Optimization
   Wang, Y.; Garcia, A.

2. Sequential Monte Carlo Multi-model-based Optimization
   Chen, X.; Zhou, E.

3. Global optimization with noise corrupted function evaluations
   Calvin, J.M.

4. A Bayesian Approach to Stochastic Root Finding
   Henderson, S.G.; Frazier, P.I.; Waeber, R.

Session TA4 - Quantitative Financial Risk Management, Chair: Zhang / Ghamami in JVdC D, see page 65

1. A Systemic Risk Model for Asset Price Contagion
   Chen, C.; Iyengar, G.; Moallemi, C.C.

2. Systemic Risk with Central Counterparty Clearing
   Amini, H.

3. Efficient Monte Carlo Counterparty Credit Risk Pricing and Measurement
   Ghamami, S.; Zhang, B.

Session TA5 - Many-server queues: Approximations and control, Chair: Gurvich in JVdC E, see page 66

1. Sacrificing some optimality in an overload control
   to achieve rapid recovery
   Perry, O.; Whitt, W.

2. Sizing Step-Down Units in Hospitals
   Armony, M.; Chan, C.W.; Zhu, B.

3. Many-Server Heavy-Traffic Limits for Queueing Networks with Time-Varying Parameters and Probabilistic Routing
   Liu, Y.; Huang, C.C.

4. Scaled control in the QED regime
   Van Leeuwaarden, J.S.H.; Janssen, A.J.E.M.; Sanders, J.

Session TA6 - Queueing Networks, Chair: D'Auria / Van Leeuwaarden in JVdC F, see page 67

1. Mixing properties and delay performance of random-access networks
   Zocca, A.; Borst, S.C.; Van Leeuwaarden, J.S.H.; Nardi, F.R.

2. Online Optimization of Product-Form Networks
   Sanders, J.; Borst, S.C.; Van Leeuwaarden, J.S.H.

3. Beyond strict insensitivity
   Jonckheere, M.; Lopez, S.

4. Congestion in processor-sharing multi-class closed networks: stationary and fluid analysis
   Anselmi, J.; D'Auria, B.; Walton, N.S.

Session TA7 - Optimal Control of Stochastic Inventory Systems, Chair: Squillante in JVdC G, see page 69

1. Asymptotic Optimality of Constant-Order Policies for Lost Sales Inventory Models with Large Lead Times
   Goldberg, D.A.; Katz-Rogozhnikov, D.A.; Lu, Y.; Sharma, M.; Squillante, M.S.

2. Matching Supply and Demand in Production-Inventory Systems: Asymptotics and Insights
   Squillante, M.S.; Lu, Y.; Yao, D.D.

Session TA8 - Estimation, Chair: Pasupathy in JVdC H, see page 69

1. Estimating Waiting Times with the Time-Varying Little's Law
   Kim, S.; Whitt, W.
2. Copula-based reliability estimation for multicomponent repairable systems via time-truncated NHPP with power law intensity
   Aminzadeh, M.
3. "Online" Quantile and Density Estimators
   Pasupathy, R.; Ghosh, S.

Tuesday 11:30am - 1:00pm

Session TB1 - Random Graphs and Complex Networks: Social Networks, Chair: Olvera-Cravioto in JVdC A, see page 70

1. Super Star Model: Predicting the Structure of Retweet Graphs in Twitter
   Zaman, T.; Bhamidi, S.; Steele, J.M.
2. Super Star Model: The Ghost in the Machine
   Steele, J.M.; Bhamidi, S.; Zaman, T.
3. The Logarithmic Dimension Hypothesis
   Bonato, A.
4. Co-Evolutionary Models of Community Structure in Networks
   Pralat, P.

Session TB2 - Stochastic Control and Revenue Management in Smart Grids, Chair: Zhang in JVdC B, see page 72

1. Voltage and Reactive Power Control for Power Loss Minimization Using Approximate Stochastic Annealing
   Feinberg, E.A.; Hu, J.; Yuan, E.
2. Dynamic electricity retail pricing in uncertain environments
   Jia, L.; Tong, L.
3. An inventory theory framework for the analysis of probabilistic demand response schemes
   Harsha, P.; Sharma, M.; Natarajan, R.; Ghosh, S.
4. Dynamic Pricing for Reliably Shaping of Electricity Demand
   Zhang, X.; Kalagnanam, J.; Katz-Rogozhnikov, D.A.; Squillante, M.S.; Ghosh, S.

Session TB3 - Simulation Optimization 2, Chair: Fu / Henderson / Zhou in JVdC C, see page 73

1. Maximizing Quantitative Traits in the Mating Design Problem
   Hunter, S.R.; McClosky, B.
2. Ranking and Selection with Tight Bounds on Probability of Correct Selection
   Frazier, P.I.
3. Chance-constrained selection of the best
   Nelson, B.L.; Hong, L.J.; Luo, J.

4. Stochastic Dynamic Decision Simulation: Markov Chains in Multi-stage Stochastic Programming
   Sen, S.

Session TB4 - Stochastic Calculus and Finance, Chair: Chigansky in JVdC D, see page 74

1. On the Markov property of some Brownian martingales
   Fan, J.Y.; Hamza, K.; Klebaner, F.C.

2. On a class of Stochastic Implied Volatility models
   Klebaner, F.C.

3. Mixed fractional Brownian motion: the filtering perspective
   Chigansky, P.; Kleptsyna, M.

Session TB5 - Fluid limits and its applications, Chair: Van Houdt in JVdC E, see page 75

1. Mean field approximation meets stochastic model checking
   Bortolussi, L.

2. Some fluid models in performance and dependability analysis
   Aspirot, L.; Mordecki, E.; Rubino, G.

3. A Mean Field Model for a Class of Garbage Collection Algorithms in Flash-based Solid State Drives
   Van Houdt, B.

Session TB6 - Transient Functionals of Structured Markov Chains, Chair: Taylor in JVdC F, see page 76

1. BRAVO for QED Queues
   Nazarathy, Y.; Daley, D.J.; Van Leeuwaarden, J.S.H.

2. The role of the deviation matrix in asymptotic functionals of Markov chains
   Hautphenne, S.; Latouche, G.; Taylor, P.G.

3. A time-dependent study of birth-death processes, via the knockout queue
   Fralix, B.

Session TB7 - Control of Queues, Chair: Down in JVdC G, see page 77

1. Optimal Admission Control for Tandem Loss Systems
   Silva, D.F.; Zhang, B.; Ayhan, H.

2. A few new views on queues
   Xie, J.; Zayas-Cabán, G.; Green, L.V.; Lewis, M.E.

3. Optimal rate for a queueing system in heavy traffic with superimposed On-Off arrivals.
   Ghosh, A.P.

4. Dynamic Scheduling of a $GI/GI/1+GI$ Queue with Multiple Customer Classes
   Ward, A.R.; Kim, J.

Session TB8 - Stochastic Models, Chair: Jonckheere in JVdC H, see page 78

1. Spatial stochastic models of heterogeneous cellular networks with repulsively deployed base stations
   Miyoshi, N.; Nakata, I.

2. Asymptotic Results for the First and Second Moments of Discrete-Time Bulk-Renewal Process
   Kim, J.J.; Chaudhry, M.L.

3. Queues and risk models with simultaneous arrivals
   Badila, E.S.; Boxma, O.J.; Resing, J.A.C.; Winands, E.M.M.

4. Particle systems and quasi-stationary distributions
   Groisman, P.; Jonckheere, M.

Session TB9 - Two-dimensional Markov Processes: Fluid Queues and Random Walks, Chair: Nguyen in Cabildo I, see page 79

1. Markov modulated two node fluid network: Tail asymptotics of the stationary distribution
   Miyazawa, M.

2. Tail asymptotics of the stationary distribution of a two dimensional reflecting random walk with unbounded upward jumps
   Kobayashi, M.; Miyazawa, M.

3. Asymptotic independence of (simple) two-dimensional Markov processes
   Latouche, G.

Tuesday 2:30pm - 4:00pm

Session TC1 - Optimization I, Chair: Kalathil in JVdC A, see page 80

1. An optimal dynamic funding solution to prevent default costs
   Ayesta, U.; Erausquin, M.; Ferreira, E.; Jacko, P.

2. Minimizing risk measures in bandit problems
   Yu, J.Y.

3. Decentralized Learning for Multi-player Multi-armed Bandits: An Algorithm for Near-Logarithmic Regret
   Kalathil, D.; Nayyar, N.; Jain, R.

Session TC2 - Stochastic systems in applications, Chair: Lu in JVdC B, see page 82

1. Risk-Aware Revenue Maximization in Display Advertising
   Heavlin, W.D.; Radovanović, A.

2. Performance Analysis and Scheduling in Big Data Systems
   Zhang, L.; Tan, J.; Meng, X.

3. Structural Properties and Heuristic Optimal Policies for Serials Lines with Flexible Workers
   Lu, Y.

4. Optimal Resource Capacity Management in Stochastic Networks
   Squillante, M.S.; Dieker, A.B.; Ghosh, S.

Session TC3 - Simulation and Estimation, Chair: Lam in JVdC C, see page 83

1. Iterative Methods for Robust Estimation under Bivariate Distributional Uncertainty
   Ghosh, S.; Lam, H.

Session TC4 - Markov modulation in risk and queueing processes, Chair: Palmowski / D'Auria in JVdC D, see page 83

1. The Sequential Probability Ratio Test revisited: linking statistics and ruin theory
   Albrecher, H.; Asadi, P.; Ivanovs, J.

2. Risk model with an observer in Markov environment
   Albrecher, H.; Ivanovs, J.

3. Approaching Markov-modulated Brownian motions via matrix-analytic methods
   Nguyen, G.

4. Heavy-traffic asymptotics for networks of parallel queues with Markov-modulated service speeds
   Dorsman, J.L.; Vlasiou, M.; Zwart, B.

Session TC5 - Spatial Stochastic Processes and their Applications, Chair: Ramanan in JVdC E, see page 84

1. A spatial model of cancer initiation: basic results
   Leder, K.Z.; Durrett, R.; Foo, J.

2. A spatial model of cancer initiation: approximations and applications
   Foo, J.

3. Limit theorems for Smoluchowski dynamics associated with critical continuous-state branching processes
   Iyer, G.; Leger, N.; Pego, R.

4. Conditional phase transitions: some simple examples
   Rebeschinim, P.; Van Handel, R.

Session TC6 - Funding Opportunities in OR at NSF, Chair: Jacobson in JVdC F, see page 86

1. Funding Opportunities in Operations Research at the National Science Foundation
   Jacobson, S.H.

Session TC7 - Stochastic control, Chair: Gupta / Keskin in JVdC G, see page 86

1. A study of the impact of long-range dependence on a class of threshold-type stochastic control policies
   Sharma, M.; Bosman, J.W.; Gao, X.; Lu, Y.; Squillante, M.S.

2. Approximate Dynamic Programming Approach to Stochastic Matching
   Bhat, N.P.; Moallemi, C.C.

3. Higher order Markov random fields for independent sets
   Goldberg, D.A.

4. A Brownian Model of Dynamic Pricing with Demand Model Uncertainty
   Keskin, Bora

Session TC8 - Novel directions in queueing, Chair: Kapodistria in JVdC H, see page 87

1. Service systems with skill based routing, under FCFS policies
   Adan, I.J.B.F.; Boon, M.A.A.; Weiss, G.
2. Using Estimated Patience Levels to Optimally Schedule Customers
   Bassamboo, A.; Randhawa, R.S.
3. An Aggregation Method For The Tandem Threshold Queue
   Baer, N.; Boucherie, R.J.; Al Hanbali, A.; Van Ommeren, J.C.W.
4. Factorial approach for the study of the infinite server queue with synchronized reneging
   Kapodistria, S.; Resing, J.A.C.; Phung-Duc, T.

Wednesday 10:00am - 11:30am

Session WA1 - Urn models, Chair: Maulik in JVdC A, see page 88

1. Pólya-Eggenberger-Friedman Urn Models: A New Approach
   Bandyopadhyay, A.; Thacker, D.
2. On a preferential attachment and generalized Polya's urn model
   Collevecchio, A.; Cotar, C.; LiCalzi, M.
3. Rates of convergence of color count in balanced urn models
   Dasgupta, A.; Maulik, K.; Ray, G.

Session WA2 - Finance and Risk I, Chair: Iyer in JVdC B, see page 89

1. Dynamic Pricing with Linear Price Correction
   Jasin, S.; Chen, G.; Duenyas, I.
2. Switch-when-safe cone-constrained mean-variance strategies
   Labbé, C.; Watier, F.
3. Feedback Stackelberg Solutions of Infinite-Horizon Stochastic Differential Games
   Bensoussan, A.; Chen, S.; Sethi, S.P.
4. Information and the Value of Execution Guarantees
   Iyer, K.; Johari, R.; Moallemi, C.C.

Session WA3 - Algorithmic and policy-level applications of probability, Chair: Kharoufeh / Goldberg in JVdC C, see page 91

1. Distributionally robust inventory control when demand is a martingale
   Goldberg, D.A.; Xin, L.
2. Asymptotically Optimal Online Stochastic Bin Packing
   Gupta, V.; Radovanović, A.
3. Optimal Replacement of a Component in a Partially-Observable Environment
   Flory, J.A.; Kharoufeh, J.
4. Selfish or Altruistic? The impact of customer routing in a self-service queue
   Chung, H.; Ahn, H.; Righter, R.

Session WA4 - Ruin Probability and Related Applications, Chair: Kaishev in JVdC D, see page 92

1. Expected discounted loss for a spectrally negative risk process
   Frostig, E.
2. A Risk Model with Reporting Delays
   Badescu, A.
3. Evaluating Ruin Probabilities in a Dependent Risk Model
   Dimitrova, D.S.
4. On Ruin Probability and Related Dual Models
   Kaishev, V.K.

Session WA5 - Stochastic Systems, Chair: Squillante in JVdC E, see page 93

1. Excessive delays in random-access networks
   Van Leeuwaarden, J.S.H.; Bouman, N.; Borst, S.C.
2. Diffusion scale tightness of invariant distributions of a large-scale flexible service system
   Stolyar, A.L.
3. Uniqueness of the stationary distribution of the diffusion approximation of a many-server queue
   Ramanan, K.; Aghajani, M.
4. Explicit Solutions for the Stationary Distribution of a General Class of Markov Processes
   Squillante, M.S.; Van Leeuwaarden, J.S.H.; Winands, E.M.M.

Session WA6 - Queueing Models for Computer Systems, Chair: Gautam in JVdC F, see page 94

1. The Benefit of Introducing Variability in Quality Based Service Domains
   Scheller-Wolf, A.; Xu, Y.; Sycara, K.
2. Where Queueing Theory Meets Computer System Design
   Harchol-Balter, M.
3. A Nonparametric Learning Approach to the Pricing and Provisioning Problem in Cloud Services
   Tan, Y.; Xia, C.H.
4. Stabilizing Queues with Non-Homogeneous and Multi-class Workloads in Data Centers
   Gautam, N.; Kwon, S.

Session WA7 - Estimation and Parameter Uncertainty in Queues, Chair: Nazarathy in JVdC G, see page 96

1. Does the Past Predict the Future?
   The Case of Delay Announcements in Service Systems
   Armony, M.; Bassamboo, A.; Ibrahim, R.
2. Bayesian analysis of the $MAP_2/G/1$ queueing system
   Wiper, M.P.; Cobo, P.R.; Lillo, R.E.
3. Sequential Staffing in Call Centers with Parameter Uncertainty
   Hasenbein, J.J.; Morton, D.P.; Zan, J.

Session WA8 - Approximation and Computation of Stochastic Models: Recent Advances, Chair: Lam in JVdC H, see page 96

1. Extinction Escape Times
   Leder, K.Z.; Foo, J.; Kelly, M.
2. Modeling and Analysis of Rogue Targets Avoiding a Detection
   Blanchet, J.H.; Wallwater, A.
3. Perfect sampling for infinite server and loss systems
   Blanchet, J.H.; Dong, J.
4. Stochastic Model for Limit Order Book with Asymptotic Analysis
   Blanchet, J.H.; Chen, X.

7

Abstracts

Monday 10:45am - 12:15pm

Session MA1 - Random Graph Models, Chair: Moyal in JVdC A

CSMA algorithms for large random graphs

Pascal Moyal, LMAC, Université de Technologie de Compiègne, pascal.moyal@utc.fr
Paola Bermolen, Universidad de la República, paola@ng.edu.uy
Matthieu Jonckheere, Universidad de Buenos Aires, matthieu.jonckheere@gmail.com

We aim at assessing the speed of convergence and efficiency of the CSMA decentralized algorithm, as functions of the number of nodes and the degree distribution, for homogeneous random graphs. The course of the algorithm determines the actual realization of the underlying random graph, as in a configuration model. We then examine the asymptotic behavior of the system under large graph scaling.

Directed random graphs with given degree distributions

Ningyuan Chen, IEOR, Columbia University, nc2462@columbia.edu
Mariana Olvera-Cravioto, IEOR, Columbia University, molvera@ieor.columbia.edu

We present a new directed random graph model whose in- and out-degree distributions can be chosen in advance from any type of distributions having finite mean. The model is closely related to the so-called pairing model and is based on in- and out-degree sequences of i.i.d. random variables. We explain how to circumvent the problem of how to force the two degree sequences to have the same sums without significantly destroying their distributional properties, and we present the directed versions of some classical results such as the limiting distribution for the number of self-loops and multiple edges.

Community Detection in the Labelled Stochastic Block Model

Marc Lelarge, INRIA, marc.lelarge@ens.fr
Simon Heimlicher, UMass Amherst, sh@srvr.ch
Laurent Massoulié, INRIA, laurent.massoulie@inria.fr
Jiaming Xu, University of Illinois, Urbana-Champaign, jxu18@illinois.edu

We consider the problem of community detection from observed interactions between individuals, in the context where multiple types of interaction are possible. We use labelled stochastic block models to represent the observed data, where labels correspond to interaction types. Focusing on a two-community scenario, we conjecture a threshold for the problem of reconstructing the hidden communities in a way that is correlated with the true partition. To substantiate the conjecture, we prove that the given threshold correctly identifies a transition on the behaviour of belief propagation from

insensitive to sensitive. We further prove that the same threshold corresponds to the transition in a related inference problem on a tree model from infeasible to feasible. Finally, numerical results using belief propagation for community detection give further support to the conjecture.

Session MA2 - Green Scheduling 1, Chair: Righter / Down in JVdC B

How queueing-theoretic models for data center power management differ from reality

Anshul Gandhi, Carnegie Mellon University, Pittsburgh, anshulg@cs.cmu.edu
Mor Harchol-Balter, Carnegie Mellon University, Pittsburgh, harchol@cs.cmu.edu
Michael A. Kozuch, Intel Labs, Pittsburgh, michael.a.kozuch@intel.com

Power management in data centers is an important research topic that has received considerable attention from queueing theorists recently. While queueing theory provides invaluable input into the design and analysis of power management policies for data centers, there are often many simplifying assumptions in the underlying queueing model that discourage data center operators from implementing these policies in practice.

In this work, we focus on three popular assumptions: (i) the time it takes to power up a server is negligible, (ii) the mean service rate of a server does not depend on the number of jobs, and (iii) the mean arrival rate for a given workload is constant. We first demonstrate, via actual implementation results, that the above three assumptions often do not hold in reality. We then propose simple modifications to the queueing model that allow us to relax the above three assumptions to more accurately reflect the reality.

The RRR Method for the Exact Analysis of Repeating Markov Chains with Uni-Directional Transitions

Anshul Gandhi, Carnegie Mellon University, Pittsburgh, PA, anshulg@cs.cmu.edu
Sherwin Doroudi, Carnegie Mellon University, Pittsburgh, PA, sdoroudi@andrew.cmu.edu
Mor Harchol-Balter, Carnegie Mellon University, Pittsburgh, PA, harchol@cs.cmu.edu
Alan Scheller-Wolf, Carnegie Mellon University, Pittsburgh, PA, awolf@andrew.cmu.edu

Infinite repeating Markov chains come up in many problems and are typically only solvable via numerical methods, such as matrix-analytic methods. In this talk, we introduce the Recursive Renewal Reward (RRR) method for solving a large class of repeating Markov chains, exactly, in closed form. We apply RRR towards analyzing an M/M/k with setup times for turning on servers.

Optimization of Server Farms

I.J.B.F. Adan, Eindhoven University of Technology, The Netherlands, iadan@tue.nl
V.G. Kulkarni, University of North-Carolina, USA, vkulkarn@email.unc.edu
A.C.C. van Wijk, Hamilton Institute, Ireland, a.c.c.v.wijk@gmail.com

In a server farm with infinitely many servers, each server can be busy, idle, or off. Jobs arrive according to a Poisson process and request exponential services. A new job occupies an idle server if there is one, and otherwise an off server (changing to busy). When a server becomes idle, there is the option to keep it idle or to switch it off (to save power). There are costs (per time unit) for idling and fixed costs to switch from off to on. We derive structural properties of the discounted cost (and average cost) optimal policy, analyze a simple heuristic policy and compare its performance with the optimal one.

On Optimal Policies for Energy Aware Servers

D.G. Down, McMaster University, Canada, downd@mcmaster.ca

V. Maccio, McMaster University, Canada, macciov@mcmaster.ca

As energy costs and energy used by server farms increase, so does the desire to implement energy aware policies. Although under some metrics, optimal policies for single as well as multiple server systems are known, a number of metrics remain without sufficient knowledge of corresponding optimal policies. We describe and analyze a model to determine an optimal policy for on/off single server systems under any metric based on expected response time, energy usage, and switching costs. We leverage this model in an application of random routing to show a range of non-trivial optimal routing probabilities and server configurations when energy concerns are a factor.

Session MA3 - Advanced Applications of Simulation, Chair: Fu / Henderson / Zhou in JVdC C

On simulation of constants from the theory of Gaussian processes

A.B. Dieker, Georgia Institute of Technology, USA, ton.dieker@isye.gatech.edu
B. Yakir, The Hebrew University of Jerusalem, Israel, msby@mscc.huji.ac.il

Although the understanding of Gaussian processes and fields has advanced steadily over the past decades, a variety of results related to extremes (tail asymptotics, extreme value theorems, laws of iterated logarithm) are only 'explicit' up to constants known as Pickands' constants. These constants have remained so elusive that even a reliable simulation algorithm has remained outside the scope of current methodology. In this talk, I will describe an approach that resolves this open problem. I will also stress the significance of the approach for other simulation problems, which is ongoing work with R. Birge.

On a Least Absolute Deviations Estimator of a Convex Function

E. Lim, University of Miami, USA, lim@miami.edu
Y. Luo, University of Miami, USA, ly1987510@gmail.com

When estimating a performance measure f_* of a complex system from noisy data over a domain of interest, the underlying function f_* is often known to have a certain shape characteristic such as convexity. In this case, one often uses convexity to better estimate f_* by fitting a convex function to data. However, the traditional way of fitting a convex function to data, which is done by computing a convex function minimizing the sum of least squares, takes too long to compute the fit. It also runs into an "out of memory" issue when the number of data points exceeds a few hundred. In this paper, we propose a computationally efficient way of fitting a convex function by computing the fit minimizing the sum of least absolute deviations rather than the sum of squares. We present numerical examples to illustrate the performance of the proposed estimator. We also establish the consistency of the proposed estimator and its derivative by proving that, under modest assumptions, the estimator and its derivative converge almost surely to the true values as the number of data points increases to infinity.

Computing volumes of convex bodies

D. Gamarnik, Massachusetts Institute of Technology, gamarnik@mit.edu
K. Ramanan, Brown University, Kavita_Ramanan@brown.edu

A longstanding challenging algorithmic problem has been that of computing the (approximate) volume of a convex body. In the mid-nineties, Dyer, Frieze and Kannan introduced an MCMC algorithm to approximate the volume of a convex body, which was shown to run in randomised polynomial time. In this work, we show how the volume computation problem for a certain class of polyhedral convex bodies is related to the so-called hard-core model that arises in statistical physics. We show that the hard-core model with Lebesgue measure displays no phase transitions and discuss its algorithmic implications.

Bounding Wrong-Way Risk by Simulation

P. Glasserman, Columbia University, New York, pg20@columbia.edu
L. Yang, Columbia University, New York, ly2220@columbia.edu

The dependence between stochastic inputs to a model is often difficult to characterize, even when good marginal information is available about the individual inputs. This problem arises in measuring counterparty risk, which involves dependence between the market value of transactions with a counterparty and the counterparty's probability of default. Positive dependence creates "wrong-way risk." We develop methods for bounding this effect and for interpolating between the worst case and the independent case based on the degree of uncertainty in the dependence.

Session MA4 - Applied Probability in Finance, Economics and Insurance II, Chair: Cadenillas in JVdC D

Heavy Traffic Limits and Applications to Insurance and Pension Fund Dynamics

J.H. Blanchet, Columbia University, New York NY, USA, jose.blanchet@columbia.edu
X. Chen, Columbia University, New York NY, USA, jose.blanchet@columbia.edu
H. Lam, Boston University, Boston MA, USA.

The theory behind heavy traffic approximations has attracted a great deal of attention in Operations Research. This theory has been applied in diverse areas in the context of operations management and related applications. An area that has not been traditionally studied in light of this theory is that of life insurance mathematics, including pension fund dynamics. In this talk we revisit some of the traditional life insurance models using the heavy traffic approximations and in particular many server queues. Moreover, we discuss sample path large deviations results for infinite server queues and their application to life insurance portfolios and new pension fund models obtained from a heavy traffic perspective. Part of our objective is to show that heavy traffic approximations have a great potential of applicability in this area.

On Polynomial Technique Applied To Financial And Actuarial Modeling

Alexander Melnikov, University of Alberta, Edmonton, AB, Canada, melnikov@ualberta.ca

The leading idea of our approach will be demonstrated with the help of polynomial extensions of a normal distribution. We use the product of a polynomial and the normal density to model the density function of the logarithmic returns in order to adjust the estimation for skewness, kurtosis and other moment parameters. Some other extensions will be presented in the talk. A link between the polynomial and the basic distribution density can be derived from the system of the Pearson's differential equation. Some of these distributions are useful in actuarial modeling. An empirical study is provided on selective models including the Black-Scholes model, the T-distribution model and their corresponding polynomial versions.

Non-Arbitrage for a Class of Informational Markets

T. Choulli and **J. Deng**, University of Alberta, Canada, tchoulli@ualberta.ca
A. Aksamit and **M. Jeanblanc**, Université Val d'Essonne, Evry, France

In this talk, I will address the question of how an arbitrage-free semimartingale model is affected when this model is stopped at a random horizon or when a honest time is totally incorporated. Precisely, I will focus on weak form of non-arbitrage that is crucial for the existence of optimal portfolio. This type of non-arbitrage is called the No-unbounded-Profit-with-Bounded-Risk (called NUPBR hereafter) concept. It is also known –in the literature– as the first kind of non-arbitrage. I will present many new results that can be classified into two principal types of results.

First, I will provide necessary and sufficient conditions on the random time (default time/exit time/random horizon) such that the Non-arbitrage concept remains valid for any model stopped at this random time. The second type of results consists of considering a fixed (but arbitrary) semimartingale and a random time, and derive the necessary and sufficient conditions under which this semimartingale stopped at this random time still fulfill the non-arbitrage concept. When a class of honest times will be incorporated into the model, we also provide necessary and sufficient conditions that guarantee the preservation of the non-arbitrage concept. Our class of honest times is much larger than the class of all stopping times. More importantly, our condition that defines our class of honest times sounds play important rôle in classifying honest times.

Stochastic correlation in asset management

L. A. Seco, Department of Mathematics, University of Toronto. seco@math.utoronto.ca
A. Reuss, Technische Universität München, Garching, Germany. andreas.reuss@gmx.de
P. Olivares, Department of Mathematics, Ryerson University, Toronto, Canada. pablo.olivares@ryerson.ca
R. Zagst, Technische Universität München, Garching, Germany. zagst@tum.de

This talk will review models for asset dependence, which capture fat tails and skewness, and take into account distinct market regimes. Special focus will be given to an α-stable regime-switching model, which allows for asymmetries in the dependence structure also. It will also discuss empirical implications in risk management and portfolio selection.

Session MA5 - Limit results for queueing and related models, Chair: Atar in JVdC E

Join the Shortest Queue Networks: Their Limiting Behavior and Open Problems

M. Bramson, University of Minnesota,
Yi Lu, University of Illinois,
B. Prabhakar, Stanford University

In join the shortest queue networks, incoming jobs are assigned to the shortest queue from among a randomly chosen subset of D queues in a system of N queues. After completion of service at its queue, a job leaves the system. We assume that jobs arrive into the system according to a rate - αN Poisson process, $\alpha < 1$, and that service at each queue is at rate 1.

In the fundamental paper of Vvedenskaya, Dobrushin, and Karpelevich (1996), it was shown that, if the service times at each queue are exponentially distributed, then the right tail of the equilibrium distribution at each queue decays doubly exponentially in the limit, as N goes to infinity. This is a substantial improvement over the case $D = 1$, where the queue size decays exponentially; this faster decay is important for various applications.

When the assumption of exponential service times is dropped, analysis of the problem becomes considerably more complicated, with significant aspects remaining to a large extent open. In recent work, the speaker, together with Y. Lu and B. Prabhakar, has investigated the asymptotic behavior of the tails of equilibrium distributions when service is FIFO and the service time distributions are Pareto. There, it was shown that, depending on the exact parameters that are involved, the tails can decay either polynomially, exponentially, or doubly exponentially fast. A number of other situations were also investigated.

A crucial step in the analysis of the equilibrium distributions is to show asymptotic independence of the individual queues, as N goes to infinity. While "obviously" true in a large variety of cases, this has been shown in only more limited contexts. In this talk, background for join the shortest queue networks will be given, including a summary of the above results. Asymptotic independence and certain other open problems will be discussed.

Diffusion models and steady-state approximations for exponentially ergodic Markovian queues

I. Gurvich, Kellogg School of Management, Northwestern University, i-gurvich@kellogg.northwestern.edu

Motivated by queues with many-servers, we study Brownian steady-state approximations for continuous time Markov chains (CTMCs). Our approximations are based on diffusion models (rather than a diffusion limit) whose steady-state, we prove, approximates well that of the Markov chain. Strong approximations provide such "limitless" approximations for process dynamics. Our focus here is on steady-state distributions and the diffusion model that we propose is tractable relative to strong approximations. Within an asymptotic framework, in which a scale parameter n is taken large, a uniform (in the scale parameter) Lyapunov condition is proved to guarantee that the gap between steady-state moments of the diffusion and those of the properly centered and scaled CTMCs, shrinks at a rate of $\sqrt{n}$. The uniform Lyapunov requirement is satisfied, in particular, if the scaled and centered sequence converges to a diffusion limit for which a Lyapunov condition is satisfied. Our proofs build on gradient estimates for the solutions of the Poisson equations associated with the (sequence of) diffusion models together with elementary Martingale arguments. As a by product of our analysis, we explore connections between Lyapunov functions for the Fluid Model, the Diffusion Model and the CTMC.

Asymptotics of epidemic-like models with applications to peer-to-peer networks

A. Shwartz, Technion
E. Altman, INRIA,
P. Nain, INRIA,
Y. Xu, INRIA

In a peer-to-peer network, an individual may acquire a copy of a file from another individual that has the file. Contact between individuals happens according to a (continuous time) random Markovian mechanism. The number of file owners grows through this copying mechanism, and may diminish as individuals leave the system. This is similar to an epidemic model. We study several variations of the model in the limit as population size increases. Of particular interest are the proportion of users who own a copy of the file, the probability that all copies of the file disappear from the network, etc. It turns out that the asymptotics change sharply near some critical values of the parameters. The results have implications on the ability of network owners to influence the unauthorized distribution of electronic material over the internet.

Three new state space collapse results for the multiclass G/G/1 queue

R. Atar, Technion
A. Biswas Technion,
A. Lev-Ari, Technion,
M. Shifrin, Technion

This talk is on asymptotic optimality results, in diffusion and moderate deviation scales, for the multiclass G/G/1 queue with various combinations of the following elements: scheduling/admission control, finite buffers, reneging, holding/rejection/reneging costs. In all cases the limiting dynamics is one-dimensional, and asymptotic optimality is achieved by switching from one priority discipline to another depending on the workload level. Related models we could not solve will also be mentioned.

Session MA6 - Heavy tails and related topics, Chair: Roy / Maulik in JVdC F

Fine-tuning risk assessment through hidden regular variation

B. Das, Singapore University of Technology and Design, Singapore, bikram@sutd.edu.sg

A. Mitra, WalmartLabs, mitra.abhimanyu@gmail.com
S. Resnick, Cornell University, sir1@cornell.edu

A Gaussian dependence structure implies asymptotic independence in the sense that, two correlated risk factors which are jointly normally distributed are extremely unlikely to take very high (or low) values together (unless their correlation is 1). This phenomenon can be observed in many other dependence structures which are not necessarily Gaussian in nature. We concentrate on multivariate regularly varying distributions (our paradigm for multivariate heavy-tailed distributions) which often exhibit a similar property of asymptotic independence. Under asymptotic independence, though a joint extreme event is quite unlikely, yet it is still possible. We provide a formulation of hidden regular variation to find the part that is missed under the cruder normalization of multivariate regular variation in such cases and discuss some statistical estimation procedures for assessing probabilities of such joint risk events under this definition of hidden regular variation. (This talk is based on joint work with A. Mitra and S. Resnick.)

Properties Of Tempered Stable Distributions

M. Grabchak, UNC Charlotte, Charlotte NC USA, mgrabcha@uncc.edu

Tempered stable distributions were introduced in Rosiński 2007 as models that look like infinite variance stable distributions in some central region, but they have lighter (i.e. tempered) tails. Such models have found applications in a variety of areas including mathematical finance, biostatistics, computer science, and physics. We extend this class to allow for more variety in the tails. While some cases no longer correspond to stable distributions they serve to make the class more flexible, and in certain subclasses they have been shown to provide a good fit to data. To characterize the possible tails we give detailed results about finiteness of various moments. We also give necessary and sufficient conditions for the tails to be regularly varying. This last part allows us to characterize the domain of attraction to which a particular tempered stable distribution belongs. We then characterize the weak limits of sequences of tempered stable distributions. We will conclude by discussing a mechanism by which distributions that are stable-like in some central region but with lighter tails show up in applications

Applications of Abelian and Tauberian theorems to free probability

Rajat Subhra Hazra, University of Zurich, Zurich
Krishanu Maulik, Indian Statistical Institute, Kolkatra

In this talk we first introduce the notion of free subexponentiality, which extends the notion of subexponentiality in the classical probability setup to the noncommutative probability spaces under freeness. We show that the distributions with regularly varying tails belong to the class of free subexponential distributions. In particular, we present some Abelian and Tauberian theorems for Stieltjes transform and Voiculescu transform of measures with regularly varying tails. (This talk is based a joint work with Krishanu Maulik.)

Session MA7 - Markov Chains and Markov Decision Problems, Chair: Arlotto / Steele in JVdC G

Comparison Inequalities and Fastest-mixing Markov Chains

James A. Fill, The Johns Hopkins University, USA, jimfill@jhu.edu
Jonas Kahn, Université de Lille 1, France, jonas.kahn@math.univ-lille1.fr

We introduce a new partial order on the class of stochastically monotone Markov kernels having a given stationary distribution π on a given finite partially ordered state space $\mathcal{X}$. When $K \preceq L$ in this partial order we say that K and L satisfy a *comparison inequality*. We establish that if $K_1, \ldots, K_t$ and $L_1, \ldots, L_t$ are reversible and $K_s \preceq L_s$ for $s = 1, \ldots, t$, then $K_1 \cdots K_t \preceq L_1 \cdots L_t$. In particular, in the time-homogeneous case we have $K^t \preceq L^t$ for every t

if K and L are reversible and $K \preceq L$, and using this we show that (for suitable common initial distributions) the Markov chain Y with kernel K mixes faster than the chain Z with kernel L, in the strong sense that *at every time t* the discrepancy—measured by total variation distance or separation or L^2-distance—between the law of Y_t and π is smaller than that between the law of Z_t and π.

Using comparison inequalities together with specialized arguments to remove the stochastic monotonicity restriction, we answer a question of Persi Diaconis by showing that, among all symmetric birth-and-death kernels on the path $\mathcal{X} = \{0, \ldots, n\}$, the one (we call it the *uniform chain*) that produces fastest convergence from initial state 0 to the uniform distribution has transition probability $1/2$ in each direction along each edge of the path, with holding probability $1/2$ at each endpoint.

We also use comparison inequalities

(i) to identify, when π is a given log-concave distribution on the path, the fastest-mixing stochastically monotone birth-and-death chain started at 0, and

(ii) to recover and extend a result of Peres and Winkler that extra updates do not delay mixing for monotone spin systems.

Among the fastest-mixing chains in (i), we show that the chain for uniform π is slowest in the sense of maximizing separation at every time.

Strong Stationary Duality for Diffusions

Vince Lyzinski, The Johns Hopkins University, USA, lyzinski@ams.jhu.edu
James A. Fill, The Johns Hopkins University, USA, jimfill@jhu.edu

Strong stationary duality has had a wide-ranging impact on Markov chain theory since its conception by Diaconis and Fill in 1990. Its diverse applications range from perfect sampling extensions of Markov Chain Monte Carlo to the establishment of cutoff phenomena for wide classes of Markov chains. We extend the idea of strong stationary duality to one-dimensional diffusion processes and in doing so recover some classical Markov chain results in the diffusion setting.

Optimal On-Line Selection of an Alternating Subsequence: A Central Limit Theorem

Alessandro Arlotto, Duke University, USA, alessandro.arlotto@duke.edu
J. Michael Steele, University of Pennsylvania, USA, steele@wharton.upenn.edu

In this talk, we analyze the optimal policy for the sequential selection of an alternating subsequence from a sequence of n independent observations from a continuous distribution, and we prove a central limit theorem for the number of selections made by that policy. The proof exploits the backward recursion of dynamic programming and assembles a detailed understanding of the associated value functions and selection rules. The methods used here suggest a profitable approach to the asymptotic analysis of other finite-horizon Markov decision problems. We also discuss how this result relates to classical central limit theorems for functionals of time non-homogeneous Markov chains.

Optimality Conditions for Total-Cost Partially Observable Markov Decision Processes with General State and Action Spaces

Eugene A. Feinberg, Stony Brook University, USA, Eugene.Feinberg@stonybrook.edu
Pavlo O. Kasyanov, National Technical University of Ukraine, Ukraine, kasyanov@i.ua
Michael Z. Zgurovsky, National Technical University of Ukraine, Ukraine, zgurovsm@hotmail.com

This talk describes sufficient conditions for the existence of optimal policies for Partially Observable Markov Decision Processes (POMDPs) with general state and action spaces. The objective criterion is either minimization of the expected total discounted costs or minimization of the expected total nonnegative costs. It is well-known that a POMDP

can be reduced to a Completely Observable Markov Decision Process (COMDP) with the state space being the sets of belief probabilities for the POMDP. Thus, a policy is optimal in POMDP if and only if it corresponds to an optimal policy in the COMDP. Here we provide sufficient conditions for the existence of optimal policies for COMDPs and therefore for POMDPs. We also derive optimality equations and convergence of value iterations for COMDPs.

Session MA8 - Statistical modelling and analysis for complex data, Chair: Yi in JVdC H

Nature-Inspired Metaheuristic Algorithms for Finding Efficient Experimental Designs

W. K. Wong, University of California, Los Angeles, USA, wkwong@ucla.edu

Optimal design theory and ideas are increasingly applied to many research disciplines. I present a brief overview of optimal design methodology along with its statistical foundation in the context of dose response studies.

Particle swarm optimization (PSO) is then introduced to find optimal designs for potentially any model and any criterion. The method works quite magically and frequently finds the optimal solution or a nearly optimal solution after a few iterations. There is virtually no assumption required for the method to perform well and the user only needs to input a few easy to work with tuning parameters. Using several popular models in the biological sciences, I demonstrate how PSO can find different kinds of optimal designs quickly, including mini-max types of optimal designs where effective algorithms to find such designs have remained elusive until now.

Response adaptive randomization in the presence of mismeasurement

X. Li, University of Minnesota Duluth, USA, xuanli@d.umn.edu

Response adaptive randomization represents a major advance in clinical trial methodology that helps balance the benefits of the collective and the benefits of the individual and improves efficiency without undermining the validity and integrity of the clinical research. Response adaptive designs use information so far accumulated from the trial to modify the randomization procedure and deliberately bias treatment allocation in order to assign more patients to the potentially better treatment. Some important issues and methods of response adaptive design of clinical trials in the presence of mismeasurement will be presented in this talk. We formulate response adaptive designs when the responses may be imperfectly measured. We consider the optimal allocations under various objectives, investigate the asymptotically best response adaptive randomization procedure, and discuss mismeasurement effects on the optimal allocation. We also derive explicit expressions for the variance-penalized criterion with imperfectly responses and propose new target proportions of treatment allocation under the criterion.

Bias analysis for misclassification in a multicategorical exposure in a logistic regression model

Y. Liu, Sasktel and TRlabs, Canada, yal026@mail.usask.ca
J. Liu, University of Saskatchewan, Canada, liu@math.usask.ca

We extend the analysis of asymptotic biases on the logistic regression model parameters in Davidov et. al. (2003) from a binary misclassified exposure to a multicategory misclassified exposure. The formulas are derived under the assumption of differential misclassification and can be simplified under the nondifferential misclassification. Further, in the logistic regression model with the misclassified exposure only, we can determine the signs of the asymptotic biases under nondifferential misclassification assumption. To better understanding of how the bias is influenced by the misclassification model, we use a numerical example. It shows that the magnitude of bias increases as the amount of misclassification error increases in the context of nondifferential misclassification. While in the context of differential misclassification, the change of bias magnitude is more complex without a clear trend.

The most powerful test and the optimal design of response adaptive clinical trials

Y. Yi, Memorial University of Newfoundland, St.John's, Canada, Yanqing.Yi@med.mun.ca

Response adaptive clinical trials have the advantages of allocating more patients to the potential better treatment but lead to dependent samples. The dependency in the samples may cause the loss of statistical power. This talk will introduce the most powerful test for response adaptive clinical trials and explore its properties. Based on the most powerful test statistic, the effect of the adaptation of treatment allocation on statistical power will be discussed and the optimal design of response adaptive clinical trials will be explored.

Session MA9 - Queues and Health, Chair: Örmeci in Cabildo I

Nursing Home Bed Planning with Medicaid Bed-Hold Payment Policies

Otis B. Jennings, Columbia University, New York, United States of America, otisbjennings@gmail.com
William A. Massey, Princeton University, New Jersey, United States of America, wmassey@princeton.edu
Jerome Niyirora, Binghamton University, State University of New York, United States of America, jn43@buffalo.edu

We focus on a profit-optimal bed need analysis by using a queueing model to approximate the number of beds requested for the duration of a multi-year planning period. Assuming that the potential residents are willing to remain on a wait-list, our analysis is based on a multi-server queueing model with abandonment, where the residual time to abandon equals its service time. Using a cost function based on certificate of need regulations, we optimize the number of beds over the transient behavior of the system. We model the demand by a non-homogeneous Poisson process and the customer service times by identically distributed and mutually independent, positive random variables that need not be exponentially distributed.

With these assumptions, we can give explicit formulas for the transient mean number of beds in use and a regulated profit function. We then develop an initial algorithm that uses our queueing model to address statistical issues related to the nursing home problem. Finally, we apply our stochastic analysis to estimate the optimal number of beds needed for long term planning. This estimation is based on a given projected increase in the demand for skilled nursing home care and the regulated revenue structure.

Dynamic Rate Coupled Processors with Abandonment: Approximation and Computation

Jamol Pender, Princeton University, United States of America, jpender@princeton.edu

We consider a two dimensional time varying tandem queue with coupled processors. We assume that jobs arrive to the first station as a non-homogeneous Poisson process. When each queue is non-empty, jobs are processed separately like an ordinary tandem queue. However, if one of the processors is empty, then the total service capacity is given to the other processor. This problem has been analyzed in the constant rate case by leveraging Riemann Hilbert theory and two dimensional generating functions. Since we are considering time varying arrival rates, generating functions cannot be used. Thus, instead of the generating functions we choose to exploit the functional Kolmogorov forward equations for the two dimensional queueing process. By using the functional forward equations, it is necessary to approximate the queueing distribution in order to compute the relevant expectations and covariance terms for the mean and variance. To this end, we expand our two dimensional Markovian queueing process in terms of a two dimensional Hermite polynomial sequence and use this expansion as a surrogate distribution. Finally, we are able to show that we can estimate probabilistic quantities of the two dimensional queueing process such as the mean, variance, and probability that each queue is empty with good accuracy.

Length-of-Stay and Optimal Portfolio of Surgical Procedures

L. Örmeci, Department of Industrial Engineering, Koç University, lormeci@ku.edu.tr

H. Bavafa, The Wharton School, University of Pennsylvania, USA, bavafa@wharton.upenn.edu
S. Savin, The Wharton School, University of Pennsylvania, USA, savin@wharton.upenn.edu

In this study, we provide analytical insights on how to optimally allocate hospital operating capacity between various types of elective surgical procedures. Our focus is on the interaction between two major constraining resources: operating rooms and hospital recovery beds. In our model, each procedure type has an associated revenue, deterministic case duration and stochastic length of stay. Length of stay is the time it takes for patients to recover from the surgery and leave the hospital; we consider discrete, arbitrary distributions of this variable. In this setting, the surgeries performed today affect the availability of hospital beds for the following days. We describe the optimal mix of procedures in the presence of a service-level constraint on hospital beds, an operating room capacity constraint, and procedure demand constraints.

Monday 2:00pm - 3:30pm

Session MB1 - Random Graphs and Complex Networks: Epidemics, Chair: Hasenbein in JVdC A

Rapid Detection of Viruses in Large Contact Networks

John Hasenbein, University of Texas at Austin, USA, jhas@mail.utexas.edu
Jinho Lee, University of Texas at Austin, USA, jinholee@utexas.edu
David Morton, University of Texas at Austin, USA, morton@mail.utexas.edu

In this talk we consider the problem of detecting a virus propagating through a contact network. We are interested in large social networks graphs with thousands or perhaps millions of nodes. The objective is to place a constrained number “honeypots" at a subset of nodes in order to either minimize the expected time until detection or maximize the probability of detection by a certain time. The decision problem can be formulated as a stochastic integer program which is inherently intractable. However, under a variety of virus propagation models, the objective function can be shown to be sub- or super-modular, justifying the implementation of greedy and lazy greedy heuristics. We also investigate solution via Monte Carlo approximation methods. To obtain realistic problem instances we mined a database containing sampled call data, for millions of users, from one of Asia’s largest service providers. From a giant component of 1.2 million callers, we sub-sampled smaller networks via a modified k-core procedure in order to test our proposed solution methodologies. We also investigated various connectivity properties of this social network as the k-core size and sampled time interval was varied.

Diffusions and cascades in random networks

Marc Lelarge, INRIA, marc.lelarge@ens.fr
Emilie Coupechoux, Laboratoire d’Informatique de Paris 6, emilie.coupechoux@lip6.fr

The spread of new ideas, behaviors or technologies has been extensively studied using epidemic models. Here we consider a model of diffusion where the individuals’ behavior is the result of a strategic choice. We study a simple coordination game with binary choice and give a condition for a new action to become widespread in a random network. Our results differ strongly from the one derived with epidemic models and show that connectivity plays an ambiguous role: while it allows the diffusion to spread, when the network is highly connected, the diffusion is also limited by high-degree nodes which are very stable. In a second part, we study a model of random networks that has both a given degree distribution and a tunable clustering coefficient. We analyze the impact of clustering on the cascades (size and frequency).

Critical random graphs and epidemics

Johan van Leeuwaarden, Eindhoven University of Technology, j.s.h.v.leeuwaarden@TUE.nl
Elie Aidekon, Université Pierre et Marie Curie, elie.aidekon@upmc.fr
Remco van der Hofstad, Eindhoven University of Technology, rhofstad@win.tue.nl
A.J.E.M. Janssen, Eindhoven University of Technology, a.j.e.m.janssen@tue.nl
Sandra Kliem, Universität Duisburg-Essen, sandra.kliem@uni-due.de
Florian Simatos, Eindhoven University of Technology, f.simatos@tue.nl

In this talk we discuss connections between critical random graphs and epidemics. We derive tail asymptotics for the cluster sizes of the standard ER graph and SIR model, and discuss extensions to inhomogeneous settings and epidemics with multiple stages.

Dynamic networks in growing populations and epidemics thereon

Tom Britton, Stockholm University and University of Florida, tomb@math.su.se
Tatyana Turova, Lunds University, tatyana@maths.lth.se
Pieter Trapman, Stockholm University, ptrapman@math.su.se

In the talk we define and analyse a model for a (growing) dynamic population in which a social network evolves: individuals create new connections and old connections disappear randomly in time. We also study what might happen if an infectious disease is introduced into the community and show that three different behaviors may happen: the epidemic dies out, the epidemic grows but slower than the community grows, or it grows faster and an endemic equilibrium appears.

Session MB2 - Green Scheduling 2, Chair: Righter / Down in JVdC B

Scheduling and staffing strategic servers

S. Doroudi, Carnegie Mellon University, Pittsburgh, PA, sdoroudi@andrew.cmu.edu
R. Gopalakrishnan, California Institute of Technology, Pasadena, CA, ragad3@caltech.edu
A. R. Ward, University of Southern California, Los Angeles, CA, amyward@marshall.usc.edu
A. Wierman, California Institute of Technology, Pasadena, CA, adamw@caltech.edu

Traditionally, research focusing on the design of scheduling and staffing policies for service systems has modeled servers as having fixed (maybe heterogeneous) rates. However, in reality, service systems are often staffed with people, and thus the rate a server chooses to work may be impacted by the scheduling and staffing policies used by the system. In this paper, we present a model for such 'strategic servers' that choose their rate in order to maximize a tradeoff between an 'effort cost', which captures the idea that servers exert more effort when working at a faster rate, and a 'value of idleness', which captures the idea that servers would prefer to be idle as much as possible. With this model, we revisit classic scheduling and staffing questions in multi-server service systems and highlight that policies designed to be optimal for non-strategic servers need to be adjusted when strategic servers are considered.

Task Assignment in a Server Farm with Switching Delays and General Energy-Aware Cost Structure

E. Hyytiä, Aalto University, Finland, esa.hyytia@aalto.fi
R. Righter, University of California Berkeley, USA, rrighter@ieor.berkeley.edu
S. Aalto, Aalto University, Finland, samuli.aalto@aalto.fi

We consider the task assignment problem to parallel servers with switching delay, where servers can be switched off to save energy. However, switching a server back on involves a constant server-specific setup delay. We will use one

step of policy iteration from a starting policy such as Bernoulli splitting, in order to derive efficient task assignment (dispatching) policies that minimize the long-run average cost. To evaluate our starting policy, we first analyze a single work-conserving M/G/1 queue with a switching delay and derive a value function with respect to a general cost structure. Our costs include energy related switching and running costs, as well as performance-related costs associated with both means and variability of waiting time and latency. The efficiency of the resulting dispatching policies is illustrated with numerical examples.

Session MB3 - Rare-Event Simulation, Chair: Fu / Henderson / Zhou in JVdC C

Rare-event Simulations for Random Differential Equations

Jingchen Liu, Columbia University, U.S.A., jcliu@stat.columbia.edu

Differential equation is an important tool to describe deterministic systems. In practice, systematic and/or random errors often exist such that some coefficients are not measured precisely, which induces uncertainties and sometimes puts the entire system in risk or under a failure situation. In this talk, we consider the extreme behavior and rare-event simulation problems for some of these random differential equations whose coefficients are described by spatially varying processes. This analysis has its applications in material science, fluid dynamics, ocean-earth sciences, etc.

Rare-event probability Estimation via Empirical Likelihood

Zdravko I. Botev, University of New South Wales, Sydney, Australia, botev@unsw.edu.au

Consider the problem of estimating the probability that the sum of d correlated lognormal random variables exceeds a given large threshold γ. In other words, consider estimating

$$\ell = \mathbf{P}(\mathsf{e}^{X_1} + \cdots + \mathsf{e}^{X_d} > \gamma) \quad , \quad (X_1, \ldots, X_d) \sim \mathsf{N}(\mathbf{0}, \Sigma),$$

where $\mathsf{N}(\boldsymbol{\mu}, \Sigma)$ denotes the multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ. This problem arises in some risk insurance models. We describe a new Monte Carlo method for estimating ℓ, in which ℓ is part of a nonparametric (empirical) maximum likelihood estimation problem. We provide numerical results that show that the currently used estimators of ℓ give inaccurate results, even though they have been shown to have vanishing relative error properties. We thus argue that vanishing relative error estimators can sometimes perform poorly in practice, and that our proposed algorithm is the first to provide an accurate and reliable estimator of ℓ.

Rare event simulation for the tail probability of the maximum of dependent random variables

L. Rojas-Nandayapa, The University of Queensland, Australia, l.rojasnandayapa@uq.edu.au

We present a new estimator for the tail probability of the maximum of general dependent random variables. This estimator has a very simple form, yet it can achieve a substantial variance reduction. We illustrate that the proposed estimator is highly flexible so it can be applied over a wide variety of dependence structures and marginal behaviors. Moreover, since the proposed estimator is in *crude* form, we can utilize standard variance reduction techniques that yield further significant improvements. We consider a number of examples where we analyze the asymptotic efficiency properties of the proposed estimator in both its crude and improved versions. The Gaussian case is studied in detail.

On Stochastic Approximations, Quasi-stationary Distributions, and Detection of Moving Targets

J.H. Blanchet, Columbia University, New York NY, USA, jose.blanchet@columbia.edu

P.W. Glynn, Stanford University, Stanford CA, USA, glynn@stanford.edu
S. Zheng, Columbia University, New York NY, USA, johnZ622@gmail.com

Consider a suitably regular Markov process, such as a Langevin diffusion, and consider the problem of computing the limiting distribution obtained by conditioning the process on staying inside, say, a compact set for an arbitrary long time. Such limiting distribution is known as the quasi-stationary distribution (QSD). It arises, as we shall see, in many applications ranging from Biology to one that is of particular interest to us, namely, national security. Since the QSD is obtained by conditioning on an event that is extremely rare, it might come as a surprise that one can design efficient Monte Carlo methods for such estimation problem. Traditional Monte Carlo procedures for estimating a QSD are based on the so-called Fleming-Viot interacting particle system, which provides an estimator that is asymptotically biased. We consider a simple estimator based on the theory of stochastic approximations. Interestingly, we note a phase transition occurring for the validity of the corresponding Central Limit Theorem related to the spectral gap of the associated Markov generator.

Session MB4 - Applied Probability in Finance, Economics and Insurance I, Chair: Cadenillas in JVdC D

Optimal Government Debt Control: Explicit Formula for the Optimal Debt Ceiling

A. Cadenillas, Ajou University and University of Alberta. abel@ualberta.ca
R. Huaman, Department of Mathematics, University of Alberta. huamnagu@ualberta.ca

Motivated by the current debt crisis in the world, we consider a government that wants to find the optimal control of its debt ratio. The debt generates a cost for the country. The government can reduce the debt ratio, but there is a cost associated with this reduction. We obtain a solution for the government debt problem. In particular, we obtain an explicit formula for the optimal debt ceiling.

Optimal Execution of a VWAP Order: a Stochastic Control Approach

Christoph Frei, University of Alberta, Canada, cfrei@ualberta.ca

We consider the optimal liquidation of a position of stock (long or short) where trading has a temporary market impact on the price. The aim is to minimize both the mean and variance of the order slippage with respect to a benchmark given by the market VWAP (volume weighted average price). In this setting, we introduce a new model for the relative volume curve which allows simultaneously for accurate data fit, economic justification and mathematical tractability. Tackling the resulting optimization problem using a stochastic control approach, we derive and solve the corresponding Hamilton-Jacobi-Bellman equation to give an explicit characterization of the optimal trading rate and liquidation trajectory. The talk is based on joint work with Nicholas Westray (Deutsche Bank AG).

Characterization of the minimal penalty of a convex risk measure with applications to robust utility maximization problem for a market model based on Lévy processes.

L. Pérez-Hernández Departamento de Economía y Finanzas, Universidad de Guanajuato, DCEA Campus Guanajuato, C.P. 36250, Guanajuato, Gto, lperezhernandez@yahoo.com

The minimality of the penalization function associated with a convex risk measure is analyzed. First, in a general static framework, we provide necessary and sufficient conditions for a penalty function defined in a convex and closed subset of the absolutely continuous measures with respect to some reference measure $\mathbb{P}$ to be minimal. When the probability space supports a Lévy process, we establish results that guarantee the minimality property of a penalty function described in terms of the coefficients associated with the density processes. The set of densities processes is

described and the convergence of its quadratic variation is analyzed. The class of equivalent local martingale measures is characterized in terms of the parameters of the price process, and the connection with convex risk measures is also presented. Then the robust utility maximization problem for a market model based on Lévy processes is analyzed. The interplay between the form of the utility function and the penalization function required to have a well posed problem is studied, and for a large class of utility functions it is proved that the dual problem is solvable as well as the existence of optimal solutions.

On Sannikov's principal-agent problem

Kelly Sin-Man Choi, University of California, Berkeley, USA, kelly.smchoi@berkeley.edu
Xin Guo, University of California, Berkeley, USA, xinguo@ieor.berkeley.edu

The principal-agent problem is one of the classical problems in economics. Recently Sannikov (2008) proposed a general continuous time stochastic control problem for which an analytical solution was provided and its optimality was demonstrated. This piece of work has attracted significant amount of research interest. In this talk, we will provide rigorous mathematical analysis on the value function with a viscosity solution approach.

Session MB5 - Asymptotic Analysis of Stochastic Networks, Chair: Ramanan in JVdC E

On the dynamic control of matching queues

I. Gurvich, Northwestern University, USA, i-gurvich@kellogg.northwestern.edu
A. R. Ward, University of Southern California, USA, amyward@marshall.usc.edu

We consider the optimal control of matching queues with dynamically arriving jobs. The objective is to minimize cumulative job holding costs over a finite time horizon. In the special case of linear (and equal across classes) holding costs, this is equivalent to maximizing the number of matched jobs. We introduce a multi-dimensional imbalance process that captures the number of additional jobs required so that *some* control policy could have matched all jobs that have arrived by that time (thus leaving all queues empty). The imbalance process facilitates the construction of a lower bound. Under a so-called match-pooling condition, we devise a discrete-review matching policy that asymptotically – as the arrival rates become large – achieves the imbalance-based lower bound.

A free boundary problem arising in order book dynamics

Ł. Kruk, Maria Curie-Skłodowska University, Poland, lkruk@hektor.umcs.lublin.pl
K. Ramanan, Brown University, USA, kavita_ramanan@brown.edu

We consider a system of two boundary value problems for the Poisson equation in a domain $D_\gamma \subset \mathbb{R}^2_+$ bounded by the coordinate axes and the free boundary curve γ, which is symmetric about the ray $\{(x, x), x \geq 0\}$. The problems in the system are coupled by an additional requirement that their solutions coincide on γ. This free boundary problem was introduced by Roşu (2009), who used it to approximate utilities of buyers and sellers in a model of order book dynamics with small "granularity" (i.e., the ratio of the interest rate and the customer arrival rates). We establish existence and uniqueness of a classical solution to the above problem and analyticity of the corresponding free boundary. A stochastic interpretation of this problem is also discussed.

From local to global stability in stochastic processing networks through quadratic Lyapunov functions

A.B. Dieker, Georgia Institute of Technology, USA, ton.dieker@isye.gatech.edu
J. Shin, IBM Research, USA, jshin@us.ibm.com

We construct a generic, simple, and efficient scheduling policy for stochastic processing networks, and provide a general framework to establish its stability. Our policy is randomized and prioritized: with high probability it prioritizes jobs which have been least routed through the network. We show that the network is globally stable under this policy if there exists an appropriate quadratic 'local' Lyapunov function that provides a negative drift with respect to nominal loads at servers. Applying this generic framework, we obtain stability results for our policy in many important examples of stochastic processing networks: open multiclass queueing networks, parallel server networks, networks of input-queued switches, and a variety of wireless network models with interference constraints. Our main novelty is the construction of an appropriate 'global' Lyapunov function from quadratic 'local' Lyapunov functions, which we believe to be of broader interest.

Talk 4: Large deviations of mean-field interacting particle systems

P. Dupuis, Brown University, USA, Paul_Dupuis@brown.edu
K. Ramanan, Brown University, USA, Kavita_Ramanan@brown.edu
W. Wu, Brown University, USA, Wei_Wu@brown.edu

We establish a sample path large deviations principle for the empirical measure of finite-state mean-field weakly interacting particle systems. The rates of the empirical measure process diminish to zero as the boundary is approached and, hence, existing theorems on large deviations do not apply. Instead, we use a weak convergence approach to establish the result. We also discuss the implications of this large deviation principle for the long-time stability of such particle systems.

Session MB6 - Advances in stationary analysis of Markov chains, Chair: Adan in JVdC F

Optimal routing of customers in polling systems

V.G. Kulkarni, University of North Carolina, Chapel Hill, NC 27599, USA, vkulkarn@email.unc.edu
N. Lee, University of North Carolina, Chapel Hill, NC 27599, USA, leent@email.unc.edu
I.J.B.F. Adan, Eindhoven University of Technology, The Netherlands, i.j.b.f.adan@tue.nl
A.A.J. Lefeber, Eindhoven University of Technology, The Netherlands, a.a.j.lefeber@tue.nl

We consider a simple polling system with two stations, no switchover times and exhaustive service discipline. The service times are iid exponential and arrival processes are Poisson. The waiting cost in the station being served is more than the waiting cost in the station that is not being served. When a customer arrives, he is routed to one of the two stations, based on the queue-lengths at the two stations and the position of the server. We consider both the individually optimal and socially optimal policies minimize the expected waiting costs. We consider fluid model approximation to derive the socially optimal costs and derive linear switching curve policies.

Stationary analysis of the shortest queue polling model

I.J.B.F. Adan, Eindhoven University of Technology, The Netherlands, i.j.b.f.adan@tue.nl
O.J. Boxma, Eindhoven University of Technology, The Netherlands, o.j.boxma@tue.nl
S. Kapodistria, Eindhoven University of Technology, The Netherlands, s.kapodistria@tue.nl
V.G. Kulkarni, University of North Carolina, Chapel Hill, NC 27599, USA, vkulkarn@email.unc.edu

We consider a two-node polling model in which customers upon arrival join the shorter of two queues. Customers arrive according to a Poisson process and the service times in both nodes are independent and identically distributed random variables having the exponential distribution. The two-dimensional process of the numbers of customers at the queue where the server is and at the other queue is a two-dimensional Markov process. We derive its equilibrium distribution using two methodologies: the compensation approach and a reduction to a boundary value problem.

Explicit Solutions and Other Properties of Successively Lumpable Quasi Skip Free Processes

L.C. Smit, Leiden University, The Netherlands, lsmit@math.leidenuniv.nl
F.M. Spieksma, Leiden University, The Netherlands, spieksma@math.leidenuniv.nl
M.N. Katehakis, Rutgers University, Newark and New Brunswick, NJ, USA, mnk@rutgers.edu

We consider the class of Quasi-skip-free processes (QSF), a generalization of the quasi-birth-and-death processes. Their probability state transition law does not permit transitions to a state in a level two or more units away from the current states level in one direction. We use a simple condition under which a QSF is successively lumpable (SL-QSF) and the steady state distribution can be calculated explicitly and rapidly. These processes have applications in many areas of applied probability including queueing theory, reliability and the theory of branching processes.
We use this successive lumpability property to derive explicit solutions and bounds for the steady state probabilities of general state space SL-QSFs, and to obtain a number of simplified derivations for results that are much more difficult to establish otherwise. Further, we discuss a procedure to decompose QSFs into separate SL-QSFs and we use the method of successive lumping to calculate discounted rewards in a QSF with a fixed policy.

A New Random-Product Representation For The Stationary Distribution Of A Markov Chain

P. Buckingham, Clemson University, Clemson, SC, USA, pbuckin@clemson.edu
B. Fralix, Clemson University, Clemson, SC, USA, bfralix@clemson.edu

We present a new random product representation for the stationary distribution of a regular, ergodic Continuous-time Markov chain (CTMC), and we use it to derive stationary distributions associated with CTMCs having a transition rate diagram that contains a small number of cycles. We also give an analogous representation for the distribution of a regular, irreducible CTMC at an independent exponential time.

Session MB7 - Markov Decision Processes, Chair: Feinberg in JVdC G

Markov Population Decision Chains with Constant Risk Posture

Pelin G. Canbolat, Koç University, Industrial Engineering, Turkey

The purpose of this work is to formulate and efficiently solve stochastic problems involving populations via Markov population decision chains with constant risk posture. A Markov population decision chain concerns the control of a population of individuals in different states by assigning an action to each individual in the system in each period. The progeny of an individual is a vector specifying the number of its immediate descendants in each state. In every transition, each individual earns a random reward and generates a random progeny. The decision maker maximizes expected (finite- or infinite-horizon) system utility under the following assumptions: (i) The utility function exhibits constant risk posture, (ii) progeny vectors of distinct individuals are independent, and (iii) progeny vectors of individuals in a state who take the same action are identically distributed. The main result is that it is possible to solve the problem with the original (finite) set of state-action pairs without augmenting it to include information about the population in each state or any other aspect of the system history. Possible applications include management of animal populations, crops or forests, control of the spread of an infectious disease, and government decisions such as taxation, demographic planning, education, health services and criminal justice systems.

Convergence properties of approximating Markov Decision Processes with unbounded jump rates

Flora M. Spieksma, University of Leiden, Mathematical Institute, Netherlands, spieksma@math.leidenuniv.nl

Modern queueing applications require allowing the jump rates to be unbounded. This yields a non-uniformisable MDP, to which value iteration cannot be applied. Truncation methods have to be applied, but these generally destroy the

structural properties of the optimal policy. Smoothed rate truncation however preserves these. For numerical evaluation of the structural properties of optimal policies, smoothed rate truncation therefore seems a more suitable method than mere truncation. Interestingly, the preservation of structural properties in a value approximation procedure seems to imply certain monotonicity properties of the optimal n-stage policy as a function of the iteration step n. We will discuss various examples exhibiting these phenomena as well conditions under which optimal policies and value function for the approximating Markov decision Process converge to the optimal policy and value function of the original process.

Sufficiency of Markov Policies for Continuous-Time Markov Decision Processes

Eugene A. Feinberg, Stony Brook University, Applied Mathematics and Statistics, USA,
eugene.feinberg@stonybrook. edu
Manasa Mandava, Stony Brook University, Applied Mathematics and Statistics, USA,
manasa.mandava@stonybrook. edu
Albert N. Shiryaev, Steklov Mathematical Institute, Probability Theory and Mathematical Statistics, Russia,
albertsh@mu.ras.ru

One of the basic facts of the theory of Markov Decision Processes is that for any policy there exists a randomized Markov policy with the same marginal state-action distributions. This fact implies that for major objective criteria, including expected total costs, expected total discounted costs, and average costs per unit time, it is sufficient to restrict the set of all policies to the set of randomized Markov policies. This is also true for problems with multiple criteria and constraints. In this talk we present similar results for Continuous-Time Markov Decision Processes (CTMDPs). The proof consists of two steps. The first step describes the structure of solutions to Kolmogorov's equations for nonhomogeneous jump Markov Processes. The second step applies these results to CTMDPs.

Cash-Flow Based Dynamic Inventory Management

Michael N. Katehakis, Rutgers University, Management Science and Information Systems, USA,
mnk@rci.rutgers.edu
Benjamin Melamed, Melamed@business.rutgers.edu

We model a firm that uses its capital position to invest on a single-product inventory, in an environment that allows the firm to utilize debt to finance increased order quantities while excess cash can be deposited at a bank to earn interest. The demand is random and could be non-stationary over periods. The objective is to maximize the expected value of the capital at the end of a finite planning horizon. We show that the optimal policy is determined by a sequence of two threshold critical values. Furthermore, we develop two myopic ordering policies which respectively provide upper and lower bounds for each threshold values. Based on these bounds, an efficient algorithm is provided to compute the two threshold values. Finally with numerical studies we provide further managerial insights.

Session MB8 - Network Routing, Chair: Honnappa in JVdC H

Self-Optimising State-Dependent Routing In Parallel Queues With Batch Servers

I. Ziedins, University of Auckland, New Zealand, i.ziedins@auckland.ac.nz

It is well-known that adding extra capacity to queues in networks where individuals choose their own route can sometimes severely degrade performance, rather than improving it. We will discuss two examples of queueing networks containing batch service queues where this is the case under probabilistic routing, but where under state- dependent routing the worst case performance is no longer seen in numerical examples. This raises the more general question of whether giving arrivals more information about the state of the network can lead to better performance, and the performance of state-dependent routing with other types of queue, such as processor sharing.

This is joint work with Heti Afimeimounga, Lisa Chen, Mark Holmes, Bill Solomon, and latterly, Niffe Hermansson and Elena Yudovina.

Wait or share service? Customers' choice at the equilibrium

Erim Kardeş, Ozyegin University, Turkey, erim.kardes@ozyegin.edu.tr

Consider a duopoly, where there are two identical firms but one processes orders according to a first-come first-served discipline and the other serves all orders simultaneously by allocating an equal capacity to each customer. Considering customers who individually wish to minimize the expected time until the end of their service, which firm would attract a larger share of the market? We model the situation in terms of a stochastic game among the customers. If a Nash equilibrium exists, it is shown that it is the unique equilibrium point. When customers are lost due to full buffers with a probability of approximately zero, it turns out that the firm sharing its resources must serve at a speed at least 15%faster than its competitor to guarantee a minimum market share of 50%.

Is the Curb 80%Full or 20%Empty? Analysis of San Francisco's Parking Experiment

R.C. Hampshire, Carnegie Mellon University, hamp@cmu.edu

Many cities are deploying smart parking information systems. The goal of smart parking initiatives is to reduce the number of cruising drivers. As a rule of thumb, smart parking interventions typically set average occupancy targets for each block throughout the day. We seek to explore how drivers experience these average occupancy targets. Using data from one year of observations from 17,000 sensors, we analyze how average occupancy is experienced by drivers in San Francisco.

Our contributions include the development of a empirical methodology based on the Erlang C formula to estimate driver cruising and parking demand from average parking occupancy. In addition, we develop a strategy to estimate the censored arrival rate for each block. Finally, we estimate the expected number of blocks a driver must cruise before finding a space, and some preliminary evidence for impacts of SFpark over one year.

Timing and Routing Games in Transitory Generalized Jackson Networks

H. Honnappa, University of Southern California, Los Angeles CA, honnappa@usc.edu
R. Jain, University of Southern California, Los Angeles CA, rahul.jain@usc.edu

How do customers passing through a network of FIFO queues, such as passengers at an airport, choose a time to arrive and a route through the network to minimize their delay? What is the interplay between these choices? We study a game theoretic model of arrival timing and routing in feedforward generalized Jackson networks that models this scenario. We assume a finite population of customers (hence, *transitory*) who arrive and route themselves such that the individual sojourn times are minimized, at equilibrium. For tractability, we study the system in the fluid regime.

[1] introduced the single server *concert queueing* game and proved that the unique Nash equilibrium arrival profile is a uniform distribution function. Routing games, too, have been studied extensively in the literature mostly assuming separable delay functions. In general, customers' arrival timing decisions will be affected by their routing decisions (and vice versa), and in this work we provide what is, to the best of our knowledge, the first analysis of this complex interplay.

We first extend the arguments from [1] to tandem and trellis network topologies (a trellis network has a number of parallel networks connected in tandem), and prove that the unique equilibrium arrival profile is a uniform distribution function; the Wardrop equilibrium routing profile, too, is unique. Furthermore, the equilibrium is fairly efficient, with a price of anarchy upper-bounded by 2. For general network topologies, we use the fact that the arrival and routing decisions form a multi-stage extensive form game to prove that a sub-game perfect equilibrium exists and it is *essentially* unique. We note that this analysis extends standard multi-commodity flow games to transient settings and non-separable delay functions.

References

[1] R. Jain, S. Juneja, and N. Shimkin. The Concert Queueing Game: To Wait or To be Late. Discrete Event Dynamic Systems, 21(1):103134, 2011.

Session MB9 - Queues I, Chair: Ding in Cabildo I

Queueing System Topologies with Limited Flexibility

John N. Tsitsiklis, jnt@mit.edu
Kuang Xu, kuangxu@mit.edu
Laboratory for Information and Decision Systems (LIDS), MIT, Cambridge, MA, USA

We study a multi-server model with n *flexible* servers and rn queues, connected through a fixed bipartite graph, where the level of flexibility is captured by the graph's average degree, $d(n)$. Applications in content replication in data centers, skill-based routing in call centers, and flexible supply chains are among our main motivations.

We focus on the scaling regime where the system size n tends to infinity, while the overall traffic intensity stays fixed. We show that a large capacity region (robustness) and diminishing queueing delay (performance) are jointly achievable even under very limited flexibility ($d(n) \ll n$). In particular, when $d(n) \gg \ln n$, a family of random-graph-based interconnection topologies is (with high probability) capable of stabilizing all admissible arrival rate vectors (under a bounded support assumption), while simultaneously ensuring a diminishing queueing delay, of order $\ln n/d(n)$, as $n \to \infty$. Our analysis is centered around a new class of virtual-queue-based scheduling policies that rely on dynamically constructed partial matchings on the connectivity graph.

A one-dimensional diffusion model for overloaded queues with customer abandonment

Shuangchi He, National University of Singapore, 1 Engineering Drive 2, Singapore 117576,
heshuangchi@nus.edu.sg

We use an Ornstein–Uhlenbeck (OU) process to approximate the queue length process for a GI/GI/$n+$M queue. This simple one-dimensional diffusion model is able to produce accurate performance estimates in two overloaded regimes: In the first regime, the number of servers is large and the mean patience time is comparable or longer than the mean service time; in the second regime, the number of servers can be arbitrary but the mean patience time is much longer (i.e., on a higher order) than the mean service time. We formulate these two regimes into an asymptotic framework where a sequence of queues is considered. The mean patience time goes to infinity in both asymptotic regimes, whereas the number of servers approaches infinity in the first regime but does not change in the second. The OU process is proved to be the diffusion limit for queue length processes in both asymptotic regimes. A crucial tool for proving the diffusion limit is a functional central limit theorem for the superposition of renewal processes. We prove that the superposition of n independent, identically distributed stationary renewal processes, after being centered and scaled in both space and time, converges to a Brownian motion as n goes large.

Stabilizing policies for probabilistic matching systems

Burak Büke, The University of Edinburgh, The United Kingdom, B.Buke@ed.ac.uk
Hanyi Chen, The University of Edinburgh, The United Kingdom, H.Chen-29@sms.ed.ac.uk
Miklós Rasónyi, The University of Edinburgh, The United Kingdom, Miklos.Rasonyi@ed.ac.uk

In this work, we introduce a novel queueing model where two types of users (customers and suppliers) arrive at a system and, instead of waiting to access a resource, the arriving users wait in the system to be matched with a potential candidate from the other queue. This new model is useful for analyzing the behavior of web portals which match the people who provide a specific service (suppliers) with the people who demand the service (customers), such as

employment portals, matrimonial and dating sites, rental portals, and multipurpose web portals (e.g. Craigslist.org and Gumtree.com). As an example, in an employment portal, the employees (suppliers) arrive at the portal at random times and upon arrival they scan the job postings available on the website. If they can find a potential match among these postings, they get hired and leave the system. If they cannot find a match, they post their resume on the website and wait until an arriving employer (customer) hires them. Employers also behave in a similar manner.

We assume that users arrive at a probabilistic matching system according to independent Poisson processes. Each given customer can match with a specific supplier with probability q independent of other customers and suppliers. Once a customer (supplier) arrives if there is at least one match for the supplier (customer), s/he chooses one arbitrarily and leaves the system immediately. If there are no matching suppliers (customers), s/he subscribes to the system and waits to be picked by an arriving supplier (customer). We prove that if a probabilistic matching system system is not controlled, it is unstable (non-ergodic). We suggest three admission control policies to stabilize probabilistic matching systems under different assumptions. The key factor differentiating this novel model from the literature is the matching probability q. Hence, we analyze the performances of the proposed policies, especially focusing on how the matching probability affects different measures. We also outline how the analysis of probabilistic matching system differs from the conventional queueing systems, and introduce some open problems related to these systems.

A Fluid Model for Overloaded Queues with Scoring-Based Priority Rules

Yichuan Ding, Sauder School of Business, University of British Columbia, BC, Canada, daniel.ding@sauder.ubc.ca
Peter Glynn, Department of Management Science and Engineering, Stanford University, CA, USA
glynn@stanford.edu

We consider a queueing system with multitype customers and servers. Whenever a server attempts to deliver service, each customer is assigned a score that depends on customer type, duration of waiting time, and server type. Service is then provided to the customer with the highest score. We characterize the behavior of such a system using a fluid limit process, which has two important features: (1) the service rate in the transient state coincides with the max-flow of a parameterized network, so can be efficiently computed using the so-called GGT algorithm; (2) the service rate at the steady state coincides with the minimal-cost max-flow of a capacitated network, so can be computed within polynomial time. Thanks to these properties, we could compute the transient dynamics as well as the stationary state of the fluid limit process efficiently, and predict the performance of the system when a scoring policy has been implemented. As a byproduct, our method can determine whether a service network allows global first-come-first-serve (FCFS), a question raised by Talreja and Whitt(2008). We illustrate the application of our model in the context of cadaver kidney allocation. In particular, the fluid model we developed can predict the steady-state allocation outcome of the scoring policy proposed by the United Network of Organ Sharing (UNOS) in 2008.

Monday 4:00pm - 5:30pm

Session MC1 - Random Graphs and Complex Networks: Algorithms, Chair: Litvak in JVdC A

Ranking algorithms on directed random networks

Ningyuan Chen, IEOR, Columbia University, nc2462@columbia.edu
Nelly Litvak, University of Twente, N.Litvak@utwente.nl
Mariana Olvera-Cravioto, IEOR, Columbia University, molvera@ieor.columbia.edu

The probabilistic analysis of information ranking algorithms on directed random networks, e.g., Google's PageRank, has recently led to natural approximations based on stochastic fixed-point equations whose explicit solutions can be constructed on weighted branching trees and whose tail behavior can be described in great detail. In this talk we present a model for generating directed random graphs with prescribed degree distributions where we can prove that the rank of

a randomly chosen node does indeed converge to the solution of the corresponding fixed-point equation as the number of nodes in the graph grows to infinity. The proof of this result is based on classical random graph coupling techniques combined with the now extensive literature on the behavior of branching recursions on trees. The results we present are applicable to a wide class of linear algorithms on directed graphs, and have the potential to be extended to other max-plus type of recursions.

Stochastic network models inspired by Szemerédi's Regularity Lemma, with applications to big data analysis and compression

Hannu Reittu, Technical Research Centre of Finland VTT, Finland, Hannu.Reittu@vtt.fi
Fülöp Bazsó, Wigner Research Centre for Physics, Budapest, Hungary, bazso@mail.kfki.hu
Ilkka Norros, VTT, Finland, ikka.noros@vtt.fi

Szemerédi's celebrated Regularity Lemma (SRL) is a cornerstone of modern graph theory and even beyond, like the proof of the famous Green-Tao theorem in number theory shows. Roughly speaking, SRL states that any large graph can be approximately substituted by a bounded number of pseudo-random bipartite graphs. Although SRL has numerous theoretical applications, there are almost no applications to real life networks. However, SRL indicates a structure that is very convenient and potentially useful for practice. In our suggested method we fix an integer k, the order of considered partitions of the node set, and try to find a partition that fits with a pseudo-random "regular structure". The objective is to minimize the overall description length of the data (MDL). We have created corresponding efficient algorithms and examined several real life networks, and in most cases the method works well and a reasonable regular structure can be found. The classes of nodes have usually some meaningful interpretation, like customers with characteristic behaviour etc. The algorithm scales well, since the structure can be found from a sample of the network, while the rest of the network is classified in linear time. This could open a way to use such a sorting method also for the analysis of a very big data in the form of matrices.

Taboos in the computation of the Effective Graph Resistance

Floske Spieksma, Leiden University, spieksma@math.leidenuniv.nl

The effective graph resistance is an effective measure for robustness of an undirected graph G. It is defined by considering the graph as an electric circuit with resistor of 1 Ohm at each edge. Kirchhoffs laws allow to calculate the potential difference per unit current for each pair of nodes. The effective graph resistance R_G is then the sum of the potential differences over all pairs of vertices.

Different characterisations of R_G exist. One is an expression in terms of the eigenvalues of the Laplacian L, which is the difference of degree and incidence matrix. Noticing that L is minus the q-matrix of a continuous time Markov process on the vertices, it appears that one may also compute R_G by considering the associated jump Markov chain. In particular, the potential difference $R(a,b)$ per unit current between vertices a and b can be calculated on basis of the taboo jump matrix in which either transitions to vertex a or vertex b are disabled. This gives rise to an efficient algorithm for computing tight upper and lower bounds for $R(a,b)$.

Session MC2 - Green Scheduling 3, Chair: Righter / Down in JVdC B

Green Computing by Putting Idle Servers in Sleep

Shengquan Wang, University of Michigan, Dearborn, Michigan, U.S.A., shqwang@umd.umich.edu
Jian-Jia Chen, Karlsruhe Institute of Technology, Karlsruhe, Germany, j.chen@kit.edu
Xue Liu, McGill University, Mintreal, Canada, xueliu@cs.mcgill.ca

Reducing the power consumption while maintaining the response time constraint has been an important goal in server system design. One of the techniques widely explored in the literature to achieve this goal is dynamic voltage/frequency

scaling (DVFS). However, DVFS is not efficient in modern systems where the overall power consumption includes a large portion of static power consumption. In this talk, we aim to address how to reduce the static power consumption by dynamic power management (DPM) with sleep model in addition to DVFS. To maximize the sleep efficiency, we propose PowerSleep, a smart power-saving scheme by carefully choosing an execution speed for the server with DVFS and sleep periods while putting the system in the sleep power mode with DPM. By modeling the system with M/G/1/PS queuing model and further significant extensions, we present how to minimize the mean power consumption of the server under the given mean response time constraint. Simulation results show that our smart PowerSleep scheme significantly outperforms the simple power-saving scheme which adopts sleep mode.

Energy Efficient Virtual Appliance Deployment Framework

Nelson Lee, University of North Carolina at Chapel Hill, USA, leent@email.unc.edu
Vidyadhar Kulkarni, University of North Carolina at Chapel Hill, USA, vkulkarn@email.unc.edu

We propose an energy efficient virtual appliance deployment framework that is applicable to large scale computing systems. This study is motivated by Virtual Computing Lab (VCL) service, which provides a remote access to students allowing them to reserve and use a virtual computer with their desired set of applications. Ideally, each set of applications, stored as a virtual appliance, is preloaded to a server before any user requests it. However, this goal is unrealistic due to the large amount of application sets. Using Erlang-B model, we provide a server scheduling and virtual appliance deployment model that results solutions satisfying performance requirement with energy efficiency. We also evaluate our policies with actual data collected from the VCLs at UNC-Chapel Hill and North Carolina State University.

Efficiently Operating Energy-limited Wireless Nodes

A. Mohapatra, Texas A&M University, USA, arupa@tamu.edu
N. Gautam, Texas A&M University, USA, gautam@tamu.edu

We consider a node in a multi-hop wireless network that is responsible for transmitting messages in a timely manner while being prudent about energy consumption. The node makes distributed decisions based on local information such as queue lengths of packets in input buffers, available energy and environmental conditions. The decisions include scheduling packets on the output buffers that would be transmitted at the next opportunity. We discuss stability issues, characterize optimal policies and present numerical examples to illustrate our findings.

Session MC3 - Simulation: Estimation, Chair: Fu / Henderson / Zhou in JVdC C

Confidence Intervals For Quantiles

M. K. Nakayama, New Jersey Institute of Technology, U.S.A., marvin@njit.edu
J.M. Calvin, New Jersey Institute of Technology, Newark, New Jersey, USA

The p-quantile of a random variable is the constant for which exactly p of the mass of its distribution lies to the left of the quantile; e.g., the median is the 0.5-quantile. Quantiles are widely used to assess risk. For example, a project manager may want to determine a time T such that the project has a 95% chance of completing by T, which is the 0.95-quantile. In finance, where a quantile is known as a value-at-risk, analysts frequently measure risk with the 0.99-quantile of a portfolio's loss; thus, there is a 1% chance that the loss will be greater than this value. For complex stochastic models, analytically computing a quantile usually is not possible, so simulation is employed. In addition to providing a point estimate for a quantile, we also want to measure the simulation estimate's error, and this is typically done by giving a confidence interval (CI) for the quantile. Indeed, the U.S. Nuclear Regulatory Commission requires that licensees of nuclear power plants demonstrate compliance using a "95/95 criterion," which entails ensuring

(with 95% confidence) that a 0.95-quantile lies below a mandated limit. In this talk we present some approaches for constructing CIs for a quantile estimated via simulation. Standardized time series (STS) is a class of methods originally proposed by Schruben (1983) to construct a CI for the steady-state mean of a stochastic process. We consider applying STS techniques to construct a CI for a quantile.

Statistical bias correction for stochastic optimization

F. Vázquez-Abad, Hunter College of the City University New York, felisav@hunter.cuny.edu

Stochastic approximation is used to solve a (possibly constrained) non-linear problem under noisy observations. The underlying model has a constant parameter λ which is unknown. Although this is not a parameter of interest for the optimization, the procedure may produce asymptotically unbiased estimates, even if some kind of concurrent estimation is used to estimate λ (as in the so-called quasi-static method). We propose a statistical correction and show that under general assumptions it reduces the bias and speeds convergence. For a particular application in flow networks, it turns out that our corrected procedure also has reduced variance.

Sampling within Algorithmic Recursions

Raghu Pasupathy, Virginia Tech, USA, pasupath@vt.edu
Fatemeh Hashemi, Virginia Tech, USA, hashemi@vt.edu
Soumyadip Ghosh, IBM T.J. Watson Research, USA, ghoshs@us.ibm.com

We consider the context of algorithmic recursions that involve quantities that need to be estimated using simulation. A typical example is the stochastic approximation type recursion where the displacement term (or “Newton decrement") at any iteration involves possibly estimating the Jacobian and Hessian of an underlying function. Within such contexts, we trade-off the deterministic error due to recursion with the “stochastic" error due to sampling. The analysis yields a characterization of the amount of sampling that should be done in such algorithmic processes to ensure consistency, and more interestingly, efficiency. The analysis is aimed at guiding recursions that dynamically sample based on the history of the process, examples of which we will discuss if time permits.

Exact Estimation vs Exact Simulation

P.W. Glynn, Stanford University, USA, glynn@stanford.edu
C.H. Rhee, Stanford University, USA, chrhee@stanford.edu

Exact simulation (or perfect simulation) of Markov chains has been an active area of research since the early 1990s. The goal of an exact simulation algorithm is to generate, in finite time, a sample from the stationary distribution of the Markov chain, based only on the ability to simulate the dynamical behavior of the chain. While exact simulation can be implemented in many interesting examples, there is no known algorithm capable of implementing exact simulation without imposing strong conditions on the underlying chain. In this talk, we discuss how relaxing the problem to one of exact estimation makes this problem much more tractable. An exact estimation algorithm focuses on constructing unbiased estimators for functionals of the stationary distribution, rather than insisting on sampling from the stationary distribution itself. With this relaxation in place, we describe how exact estimation can be implemented for Harris recurrent Markov chains and non-Harris recurrent contracting chains.

Session MC4 - Optimal Stopping and American Options, Chair: Ludkovski in JVdC D

Snell envelope with small probability criteria

P. Hu, Oxford University, United Kingdom, peng.hu@oxford-man.ox.ac.uk

We present a new algorithm to compute the Snell envelope in the specific case where the criteria to optimize is associated with a small probability or a rare event. This new approach combines the Stochastic Mesh approach of Broadie and Glasserman with a particle approximation scheme based on a specific change of measure designed to concentrate the computational effort in regions pointed out by the criteria. The theoretical analysis of this new algorithm provides non asymptotic convergence estimates. Finally, the numerical tests confirm the practical interest of this approach.

Joint work with Pierre Del Moral and Nadia Oudjane.

Sequential Regression Methods for Optimal Stopping

M. Ludkovski, University of California Santa Barbara, USA, ludkovski@pstat.ucsb.edu

We propose a new *sequential* Monte Carlo algorithm for computing Snell envelopes that improves the original approach of Longstaff and Schwartz. Our method employs Active Learning to adaptively refine the stochastic mesh so as to focus on correct classification of the stopping region. Moreover, the use of sequential regression can quantify the uncertainty in approximating the stopping boundaries, providing online guidelines on the amount of computational effort needed. Numerical experiments benchmarking the performance of the proposed algorithm vis-a-vis existing approaches will be presented.

Joint work with R. Gramacy.

Robust Optimal Stopping under Volatility Uncertainty

S. Yao, University of Pittsburgh, USA; songyao@pitt.edu
E. Bayraktar, University of Michigan, USA; erhan@umich.edu

We analyze a robust optimal stopping problem in a financial market with volatility uncertainty. This is a zero-sum controller-stopper game in which the stopper is trying to maximize its pay-off against an adverse player which tries to minimize this payoff by choosing the probability measure from a set $\mathcal{P}_t$ of measures who are not necessarily equivalent. In particular, we analyze the *upper Snell envelope* $\overline{Z}$ of the reward process Y and by comparing it with the Snell envelope of Y under each individual probability $\mathbb{P}$, we show that $\overline{Z}$ is an $\underline{\mathscr{E}_t} \stackrel{\triangle}{=} \inf_{\mathbb{P}\in\mathcal{P}_t} \mathbb{E}_{\mathbb{P}}[\cdot]-$supermartingale, and a $\underline{\mathscr{E}_t}-$*martingale* up to the first time τ^* when $\overline{Z}$ meets Y. Consequently, τ^* is the optimal stopping time for the robust optimal stopping problem.

Session MC5 - Systems with parallel skilled-based service, Chair: Adan in JVdC E

Matching Queues with Flexible Servers

Mor Harchol-Balter, Carnegie Mellon University, Pittsburgh, PA, harchol@cs.cmu.edu
Kristen Gardner, Carnegie Mellon University, Pittsburgh, PA, ksgardne@cs.cmu.edu
Sherwin Doroudi, Carnegie Mellon University, Pittsburgh, PA, sdoroudi@andrew.cmu.edu
Alan Scheller-Wolf, Carnegie Mellon University, Pittsburgh, PA, awolf@andrew.cmu.edu

We are motivated by online games where players seek partners of rating (level) similar to their own to play against, but where the players become more "flexible" over time if they haven't yet been matched. We illustrate a new method, Recursive Renewal Reward (RRR), for analyzing such systems.

Dynamic bipartite matching models

A. Bušić, INRIA and École Normale Supérieure, France, ana.busic@inria.fr

V. Gupta, University of Chicago Booth School of Business, USA, varun.gupta@chicagobooth.edu
J. Mairesse, CNRS, Univ Paris Diderot, France, jean.mairesse@liafa.univ-paris-diderot.fr

Matching models arise in many applications, such as healthcare, transportation, and energy. We consider a bipartite matching queueing model, where customers and servers play symmetrical roles. There is a finite set C, resp. S, of customer, resp. server, classes. Time is discrete and at each time step, one customer and one server arrive in the system according to a joint probability measure μ on $C \times S$, independently of the past. Also, at each time step, pairs of *matched* customer and server, if they exist, depart from the system. Authorized *matchings* are given by a fixed bipartite graph. A *matching policy* decides how to match when there are several possibilities. Customers/servers that cannot be matched are stored in a buffer. The evolution of the model can be described by a discrete time Markov chain. We study this model under various admissible matching policies including: ML (Match the Longest), MS (Match the Shortest), FIFO (match the oldest), RANDOM (match uniformly), and PRIORITY.

Routing to Minimize Waiting and Callbacks in Large Call Centers

A. R. Ward, University of Southern California, USA, amyward@marshall.usc.edu
D. Zhan, University of Southern California, USA, Dongyuan.Zhan.2015@marshall.usc.edu

In a call center, agents may handle calls at different speeds, and also may be more or less successful at resolving customers' inquiries, even when only considering customers calling with similar requests. One common measure of successful call resolution is whether or not the call results in the customer calling back. This presents a natural trade-off between speed and quality, where quality is defined by the percentage of the agent's calls that result in callbacks. The relevant control is the routing; that is, the decision concerning which agent should handle an arriving call when more than one agent is available. In an inverted-V model setting, we formulate an optimization problem with the dual performance objective of minimizing customer wait time and minimizing the callback rate. We solve this optimization problem asymptotically in the Halfin-Whitt many-server limit regime, interpret its solution as a routing control for the discrete-event system, and show via simulation that the interpreted routing control is on the efficient frontier. In particular, any routing control that has a lower average wait time (callback rate) must also have a higher callback rate (average wait time).

Structure of FCFS infinite matching

G. Weiss, The University of Haifa, Israel, gweiss@stat.haifa.ac.il
I.J.B.F. Adan, Eindhoven University of Technology, The Netherlands, iadan@tue.nl

We consider two random infinite sequences of items, each consisting of several types, with a bipartite compatibility graph of allowed matches, and investigate FCFS (first come first served) matching of the two sequences. This model is a great simplification of service systems with skill based routing, but it captures many important features. It is also relevant to organ transplants, stock exchange buy and sell bids, and cloud computing. This seemingly intractable model has in fact an embarrassingly simple structure. We define a Markov chain that describes the evolution of the matching process, show that it is reversible and has simple Bernoulli product form steady state distribution, and list various performance measures that can be calculated.

Session MC6 - Matrix Analytic Methods, Chair: Van Houdt in JVdC F

A constructive proof of {PH}ase type characterization theorem

Illes Horvath, Budapest University of Technology and Economics
Miklos Telek, Budapest University of Technology and Economics, telek@hit.bmu.hu

O'Cinneide's characterization result about PH representation of Matrix Exponential (ME) distributions has been available for more than 20 years. Based on this theorem Commault and Mocanu recommended a monocyclic representation

of PH distributions.
Utilizing the elements of these results we provide a constructive proof of the characterization theorem. Specifically, we present a procedure to generate a PH representation for any Matrix exponential function exhibiting the dominant eigenvalue and the positive density conditions and we prove that the applicability of the procedure. Our proof is rather elementary using on basic function properties (instead of complex mathematical constructions such as polytopes).

Queueing models with matrix-exponential distributions and rational arrival processes

Bo Friis Nielsen, The Technical University of Denmark, Denmark, bfn@imm.dtu.dk
Nigel Bean, University of Adelaide, Australia, nigel.bean@adelaide.edu.au

Traditional proofs of the matrix geometric solutions for queues of the GI/M/1 type use probabilistic arguments related to the sample paths of Markov chains. Such an approach is not directly applicable when the components with which we are modelling consist of matrix-exponential distributions and rational arrival processes. We have previously presented a method based on a modification of a last entrance argument introduced by Ramaswami. In this talk we will focus on an approach based on results for Markov chains on general state spaces when considering state changes in the embedded Markov chain of the queue.

Queues with Customer Interjections

Qi-Ming He, University of Waterloo, Canada, q7he@uwaterloo.ca
Alireza A. Chavoushi,

We study queueing systems with customer interjections. Customers are distinguished into normal customers and interjecting customers. All customers join a single queue waiting for service. A normal customer joins the queue at the end and an interjecting customer tries to cut in the queue. The waiting times of normal customers and interjecting customers are studied. Two parameters are introduced to describe the interjection behavior: the percentage of customers interjecting and the tolerance level of interjection by individual customers. The relationship between the two parameters and the mean and variance of waiting times is characterized analytically and numerically. Issues for further investigation will be discussed.

Delay and Energy Efficiency of Tree Algorithms with Free Access

Robbe Block,University of Antwerp, robbe.block@ua.ac.be
Benny Van Houdt,University of Antwerp, benny.vanhoudt@ua.ac.be

This paper presents a branching process approach to determine the main performance measures of a variety of conflict resolution algorithms known as tree algorithms with free access. In particular we present an efficient approach to calculate the mean delay, number of transmission attempts, collision resolution interval length and energy usage with arbitrary precision. A detailed discussion of the approach will be presented for the basic Q-ary tree algorithm under the standard information theoretical model, but the approach can be readily extended to many other settings, such as channels with variable packet lengths, channel errors, collision detection, etc.

Session MC7 - Markov Decision Processes II, Chair: Van de Ven in JVdC G

Monotonic successive approximations in queueing systems

H. Blok, Leiden University, The Netherlands, blokh1@math.leidenuniv.nl
S. Bhulai, VU University Amsterdam, The Netherlands, s.bhulai@vu.nl
F.M. Spieksma, Leiden University, The Netherlands, spieksma@math.leidenuniv.nl

In many queueing models with arrival or departure control it can be shown that a policy with a switching curve is optimal. Nonetheless, it can still be difficult to compute the switching curve and thus the optimal policy. In this presentation we propose a fast method for approximating such a curve in the following way. Suppose the queueing model can be modelled as a Markov decision process for which successive approximation converges. Often the iteration step has a monotonic property in the sense that the switching curve either increases or decreases in the iteration step of the succesive approximation scheme, depending on the initialisation. As the number of iterations grows, this provides an increasingly tighter bound on the set of potentially optimal curves. We will discuss the method and provide several applications.

Stochastic Sequential Assignment Problem with Threshold Criteria

Sheldon H. Jacobson, University of Illinois, Urbana, Illinois, shj@illinois.edu
Golshid Baharian,University of Illinois, Urbana, Illinois, gbahari2@illinois.edu

The stochastic sequential assignment problem (SSAP) allocates distinct workers to sequentially arriving tasks with stochastic parameters to maximize the expected total reward. In this paper, the assignment of tasks is performed under the threshold criterion, which seeks a policy that minimizes the probability of the total reward failing to achieve a target value. A Markov decision-process approach is employed to model the problem, and sufficient conditions for the existence of a deterministic Markov optimal policy are derived, along with fundamental properties of the optimal value function. An algorithm to approximate the optimal value function is presented, and convergence results are established.

Simulation-Based Approximate Dynamic Programming Approaches for Semiconductor Manufacturing Operations

Xiaoting Chen, Department of Electrical Engineering and Computing Systems, University of Cincinnati, USA
Emmanuel Fernandez, Department of Electrical Engineering and Computing Systems, University of Cincinnati, USA
W. David Kelton, Department of Operations, Business Analytics and Information Systems, University of Cincinnati, USA
Dr. Jose A. Ramirez-Hernandez, Operations Research & Decision Support American Airlines, Inc.

We present decision and control models for operations in reentrant line manufacturing (RLM) systems, e.g., in semiconductor manufacturing. The Markov Decision Process approach used commonly is set up to minimize the total work-in-process (WIP), which in turn indirectly minimizes cycle time (CT), the real objective sought. By viewing the problem in a novel way, we re-formulate it as one that seeks to select the best cost function leading to optimal cycle times. We also present results of a Markov Decision Process model, and extended simulation studies, based on a benchmark problem, using a simulation-based approximate dynamic programming method.

Modeling battery energy storage

P. Harsha, IBM Research, USA, pharsha@us.ibm.com
M. Petrik, IBM Research, USA, mpetrik@us.ibm.com
P.M. van de Ven, IBM Research, USA, pmvandev@us.ibm.com

Energy storage is an essential component of emerging smart grids, and can be used to incorporate renewables into the grid or for balancing electricity demand across the day. The increased attention for such storage has resulted in a wide range of models for storage technology, each highlighting particular aspects of storage behavior, while ignoring others.

We focus on battery energy storage, and first discuss how battery type and charging pattern (charging rate, fraction of the capacity utilized) affect battery properties like lifetime, capacity and efficiency, and propose various approaches to modeling this interaction. We then consider a setting where the storage is used for energy arbitrage, and study for these different battery models the profit-maximizing policy for battery charging and replacement. We evaluate the performance and demonstrate that battery modeling has a significant effect on the optimal policy.

Session MC8 - Networks, Chair: Lewis in JVdC H

Randomization Approaches for Network RM with Choice Behavior

S. Kunnumkal, Indian School of Business, Hyderabad, India, sumit_kunnumkal@isb.edu

We present new approximation methods for the network RM problem with customer choice. Our methods are sampling-based and so can handle fairly general customer choice models. We only assume that customers are endowed with an ordered list of preferences among the products and choose the most preferred alternative among the available ones. The starting point for our methods is a dynamic program that allows randomization. An attractive feature of this dynamic program is that the size of its action space is linear in the number of itineraries. We present two approximation methods that build on this dynamic program and use ideas from the independent demands setting.

Necessary conditions for the invariant measure of a random walk to be a sum of geometric terms

Y. Chen, University of Twente, The Netherlands, y.chen@utwente.nl
R.J. Boucherie, University of Twente, The Netherlands, r.j.boucherie@utwente.nl
J. Goseling, University of Twente, The Netherlands, j.goseling@utwente.nl

We consider the invariant measure of homogeneous random walks in the quarter-plane. In particular, we consider measures that can be expressed as an infinite sum of geometric terms. We present necessary conditions for the invariant measure of a random walk to be a sum of geometric terms. We demonstrate that each geometric term must individually satisfy the balance equations in the interior of the state. We show that the geometric terms in an invariant measure must have a pairwise-coupled structure. We further show that the random walk cannot have transitions to the North, Northeast or East. Finally, we show that for an infinite sum of geometric terms to be an invariant measure at least one coefficient must be negative. This paper extends our previous work for the case of finitely many terms to that of countably many terms.

The Independence Number of a Maximal Outerplanar Graph

Thomas M. Lewis, Furman University, Greenville, SC, USA, tom.lewis@furman.edu

We show that the expected size of the independence number of a maximal outerplanar graph on n vertices is asymptotic to $(1 - e^2)n/2$ as $n \to \infty$. We also show that there is a corresponding strong law of large numbers.

Session MC9 - Queues II, Chair: Shin in Cabildo I

Staffing Service Systems with Load Dependent Service Rate

J. Dong, Columbia University, US, jd2736@columbia.edu
P. Feldman, University of California Berkeley, US, feldman@haas.berkeley.edu
G.B. Yom-Tov, Technion - Israel Institute of Technology, Israel, gality@tx.technion.ac.il

Most operations management literature assumes that service times are independent of the load of the system. However, empirical evidence suggests that the two are correlated. Several factors could contribute to the correlation. During heavily loaded intervals, fatigue may cause agents to slow down, while pressure may cause them to speed up. On the customer side, correlation between load and service time is well established. For example, patients' condition may worsen if treatment is delayed in health care facilities, resulting in longer stays. Speedup by itself will not worsen performance (measured by delays and abandonments), while slowdown may not only decrease customer service level

but also increase agents' workload dramatically. Hence, we concentrate on the latter, and examine how the dependence between service rate and work load affects the operational performance of the system. We do that by developing and analyzing fluid and diffusion approximations of an Erlang-A model with load-dependent service times. We propose methods that help stabilize and improve system performance. We show that if load sensitivity is moderate, a specific correction to the square-root staffing formula is required to achieve good service levels in the QED regime, while if load sensitivity is large such a method is not adequate, and other interventions are preferable.

Queuing with Future Information

Joel Spencer, Courant Institute of Mathematical Sciences, New York University, spencer@courant.nyu.edu
Madhu Sudan, Microsoft Research New England, madhu@mit.edu
Kuang Xu, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, kuangxu@mit.edu

We study an admissions control problem, where a queue with service rate $1-p$ receives incoming jobs at rate $\lambda \in (1-p, 1)$, and the decision maker is allowed to redirect away jobs up to a rate of p, with the objective of minimizing the time-average queue length.

We show that the amount of *information about the future* has a significant impact on system performance, in the heavy-traffic regime. When the future is unknown, the optimal average queue length diverges at rate $\sim \log_{\frac{1}{1-p}} \frac{1}{1-\lambda}$, as $\lambda \to 1$. In sharp contrast, when all future arrival and service times are revealed beforehand, the optimal average queue length converges to a finite constant, $(1-p)/p$, as $\lambda \to 1$. We further show that the finite limit of $(1-p)/p$ can be achieved using only a *finite* lookahead window starting from the current time frame, whose length scales as $\mathcal{O}\left(\log \frac{1}{1-\lambda}\right)$, as $\lambda \to 1$. This leads to the conjecture of an interesting duality between queuing delay and the amount of information about the future.

A Quasi-Skip-Free processes with "quasi" product form stationary distribution

Dwi Ertiningsih, Leiden University, The Netherlands, dwiertiningsihd@math.leidenuniv.nl
Flora Spieksma, Leiden University, The Netherlands, spieksma@math.leidenuniv.nl
Laurens Smit, Leiden University, The Netherlands, lsmit@math.leidenuniv.nl

In this talk we will discuss QSF (Quasi-Skip-Free) processes. These are a generalisation of QBD (Quasi-Birth-Death) processes, where transitions are allowed across several levels in one direction. In particular, we will study so-called skip-free to the left processes that are bounded to the right. We assume that each level of the QSF process contains one exit state. This means that each level contains precisely one state with a positive jump rate to the next lower level.

This implies that jump matrix to the next lower level contains only single non-zero row. Under homogeneity and irreducibility assumptions, we will show that there exists an invariant measure which is of product form as a function of the level. In the case of an ergodic process, the stationary distribution has this product form property.

We will discuss several applications, such as $PH/M/1$ and $PH/M/1$-batch service queues.

Approximation of serial lines with multiple servers and finite buffer

Yang Woo Shin, Changwon National University, Korea, ywshin@changwon.ac.kr
Dug Hee Moon, Changwon National University, Korea, dhmoon@changwon.ac.kr
Ho Sueb Lee, Changwon National University, Korea, hslee@changwon.ac.kr

We consider the serial lines with finite buffers and parallel unreliable servers whose service time is exponential. The blocking after service blocking protocol is assumed. A server is either up (operational) or under repair(broken-down) at any time. We assume the operation dependent failure that is, each server at each stage can be failed only while the

server is working. The time to failure and time to repair are exponentially distributed. The system is approximated based on the decomposition method. To reflect the dependence between consecutive nodes, we use the subsystem with three service stations and two buffer spaces instead of two service stations and one buffer space mostly used in almost all decomposition technique. The subsystems are approximated by two stages tandem queue with two buffers and state dependent arrival rates and service rates. The approximate subsystem is modeled with level dependent quasi-birth-and-death process. Arrival and service rates of the subsystem are calculated iteratively.

Tuesday 8:30am - 10:00am

Session TA1 - Random Graphs and Complex Networks: Connectivity, Chair: Britton in JVdC A

Giant Component in Random Multipartite Graphs with Given Degree Sequences

Sidhant Misra, MIT Sloan School of Management, sidhant@mit.edu
David Gamarnik, MIT Sloan School of Management, gamarnik@mit.edu

We study the problem of the existence of a giant component in random multipartite graphs. We consider a random multipartite graph with p parts generated according to a given degree sequence $n_i^{\mathbf{d}}$ which denotes the number of vertices in part i of the multipartite graph with degree given by the vector $\mathbf{d}$. We assume that the degree sequence converges to a probability distribution. Under certain additional regularity assumptions, we characterize the conditions under which, with high probability, there exists a component of linear size. We also characterize the size of the giant component when it exists. We use the exploration process of Molloy and Reed combined with techniques from the theory of multidimensional Galton-Watson processes to establish this result.

Retransmission Delays over Correlated Channels

Predrag R. Jelenković, Department of Electrical Engineering, Columbia University, predrag@ee.columbia.edu
Evangelia Skiani, Department of Electrical Engineering, Columbia University, valia@ee.columbia.edu

High variability and frequent failures characterize most of the existing communication networks where retransmission-based failure recovery represents a primary approach for successful data delivery. Recent work has shown that, when data sizes have infinite support, retransmissions can cause power law delays and instabilities even if all traffic and network characteristics are super-exponential. The prior studies have considered a baseline model where the channel switches independently between the available and unavailable states. However, communication channels are highly correlated in reality, and in this paper, we extend the prior work on the i.i.d. model to the dependent case.

We use modulated processes, e.g. Markov modulated, to capture the channel dependencies. We study the number of retransmissions and delays under the assumption that the hazard functions of the distributions of data sizes and channel statistics are proportional, conditionally on the channel state. Our results show that the tails of the retransmission and delay distributions are dominated by the longer channel availability periods, informally implying that the “best case wins". Furthermore, we provide explicit formulas that approximate uniformly the entire bodies of the retransmission and delay distributions, respectively. The exact approximations demonstrate the effects of every channel state on the main bodies of the distributions. These main body approximations may be even more relevant than the tails for specific target applications. The accuracy of our analytic approximations is validated via simulation experiments.

The largest component of a hyperbolic model of complex networks

Tobias Müller, Mathematical Institute, Universiteit Utrecht, t.muller@uu.nl

We consider the largest component of a model of random graphs that was introduced by Krioukov et al. in 2010. In this model, we distribute N points at random inside a disk of radius R in the hyperbolic plane, and we join two points by an edge if their (hyperbolic) distance is at most R. Here R is chosen to depend on N in a specific way, and the points follow a 'quasi-uniform' distribution controlled by a parameter α (where $\alpha = 1$ yields the uniform distribution).

As observed by Krioukov et al., and recently verified rigorously by Gugelmann et al., the model exhibits a number of phenomena usually associated with complex networks, including clustering and a power law degree sequence with exponent $2\alpha + 1$.

In joint work with M. Bode and N. Foutoulakis, we are able to show that $\alpha = 1$ is the threshold for the emergence of "giant". That is, when $\alpha > 1$ all components are sublinear, while if $\alpha < 1$ then there is a component of linear size. We are also able to show that $\alpha = \frac{1}{2}$ is the threshold for connectivity. That is, when $\alpha > \frac{1}{2}$ then the graph is disconnected, while if $\alpha < \frac{1}{2}$ then it is connected. (all statements holding with probability tending to one as N tends to infinity) The behavior at the threshold appears to be rather different than in all other models of random graphs we are aware of.

(Based on joint work with M. Bode and N. Fountoulakis)

Degree-degree dependencies in random graphs with heavy-tailed degrees

Nelly Litvak, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, N.Litvak@utwente.nl

Mixing patterns in large self-organizing networks, such as the Internet, the World Wide Web, social and biological networks are often characterized by degree-degree dependencies between neighbouring nodes. One of the problems with the commonly used assortativity coefficient is that in disassortative networks its magnitude decreases with the network size. This makes it impossible to compare mixing patterns, for example, in two web crawls of different size. As an alternative, we have recently suggested to use rank correlation measures, such as Spearman's rho. Numerical experiments have confirmed that Spearman's rho produces consistent values in graphs of different sizes but similar structure, and it is able to reveal strong (positive or negative) dependencies in large graphs. In particular, applied to Web crawls, Spearman's rho has revealed much stronger negative degree-degree dependencies in Web graphs than was previously thought. We analytically investigate degree-degree dependencies for scale-free graph sequences, and provide mathematical proofs for the previously obtained numerical results. We start with a simple model of two heavy-tailed highly correlated random variable X and Y, and show that the sample correlation coefficient converges in distribution either to a proper random variable on $[-1, 1]$, or to zero, and if $X, Y \geq 0$ then the limit is non-negative. We next adapt these results to the assortativity in networks as described by the degree-degree correlation coefficient, and show that it is non-negative in the large graph limit when the degree distribution has an infinite third moment. Then we consider the alternative degree-degree dependency measure, based on the Spearman's rho, and prove that this statistical estimator converges to an appropriate limit under very general conditions. We verify that these conditions hold in common network models, such as configuration model and Preferential Attachment model. We conclude that rank correlations provide a suitable and informative method for uncovering network mixing patterns.

Session TA2 - Applied Probability in Healthcare, Chair: Bayati in JVdC B

A Dynamic Random Graph Model for Kidney Exchange

R. Anderson, Massachusetts Institute of Technology, USA, rma350@mit.edu
I. Ashlagi, Massachusetts Institute of Technology, USA, iashlagi@mit.edu
D. Gamarnik, Massachusetts Institute of Technology, USA, gamarnik@mit.edu
Y. Kanoria, Microsoft Research, USA, ykanoria@gmail.com

Kidney exchange programs enable patients with living but incompatible donors to exchange kidneys in chains initiated by altruistic donors and short cycles. The problem of maximizing the number of transplants in a fixed pool of patients has been well studied. In practice, exchange programs perform this optimization periodically; however, the long run implications of this policy for average patient waiting time are not well understood. To address this issue, we propose a

dynamic random graph model of a kidney exchange program. In each time period, a single node (a donor-patient pair) arrives, and for each pre-existing node in the graph, directed edges (feasible transplants) are added to and from the new node with probability p, i.i.d. We consider two policies for deleting nodes (selecting transplants): either we remove short cycles, or we maintain and advance a single chain. In both cases, we show that the "greedy policy," where we maximize the number of transplants in every time period, results in a long run average patient waiting time that is as good as any batching policy, up to constant factors as $p \to 0$. Additionally, we find that as $p \to 0$, the steady state average patient waiting time using the chains policy is an order of magnitude smaller in p than under the cycles policy, which explains why modern kidney exchanges in the U.S. tend to do most transplants through chains.

Donor-Dependent Scoring Schemes: Shaping the Allocation of Cadaver Kidneys in a New Era

Baris Ata, Kellogg School of Management, Northwestern University, Evanston, IL, USA, b-ata@kellogg.northwestern.edu
Yichuan Ding, Sauder School of Business, University of British Columbia, BC, Canada, daniel.ding@sauder.ubc.ca
Stefanos Zenios, Stanford Graduate School of Business, Stanford University, CA, USA stefzen@stanford.edu

In the United States, candidates on the cadaver kidney transplant waitlist are ranked using a scoring scheme that takes into account characteristics of the donor and the candidate. This paper undertakes a modeling based analysis of a general class of scoring systems. We consider scoring systems that are donor-dependent, but also donor-independent. We use a fluid model to approximate the transplant waitlist, and assume that patients make rational decisions based on the predictions that the fluid model makes. We show that at the equilibrium state, a donor-dependent ranking system increases the chance of survival-based matching between a recipient and a donor, and enables the acceptance of a larger range of kidneys when compared to a donor-independent ranking system. We use the simulation model developed by the Scientific Registry of Transplant Recipients (SRTR) to compare the current policy and our policy. The simulation results show that the donor-dependent policy we proposed improves the total life years from transplant by 12% and reduces the number of discarded kidneys by about 7%, as compared with the current policy, which is more close to a donor-independent policy.

Queues with Delay Sensitive Service Times

Carri W. Chan, Columbia Business School, USA, cwchan@columbia.edu
Vivek F. Farias, Massachusetts Institute of Technology, USA, vivekf@mit.edu
Gabriel Escobar, Kaiser Permanente, USA, gabriel.escobar@kp.org

Mainstream queueing models are frequently employed in modeling healthcare delivery in a number of settings, and further used in making operational decisions for the same. The vast majority of these queueing models assume that the service requirements of a job are independent of the state of the queue upon its arrival. In a healthcare setting, this assumption is equivalent to ignoring the effects of delay experienced by a patient awaiting care. However, it is only natural to conjecture that long delays may have adverse effects on patient outcomes and can potentially lead to longer lengths of stay (LOS) when the patient ultimately does receive care. This work sets out to understand these delay issues from an operational perspective. In particular, we empirically measure how congestion in the Intensive Care Unit (ICU) can lead to delays ICU admission and measure the impact on the patient's ICU LOS. Next, we consider how to incorporate these measured delayed effects into a queueing model and characterize approximations to various quantities of interest when the service time of a job is adversely impacted by the delay experienced by that job. Our findings suggest that this delay effect can be substantial and ignoring it when using queueing models to model healthcare delivery systems may result in significant under-provisioning.

Active Postmarketing Drug Surveillance for Multiple Adverse Events

Joel Goh, Stanford Business School, USA, joelgoh@stanford.edu
Stefanos Zenios, Stanford Business School, USA, stefzen@stanford.edu

Active postmarketing drug surveillance is important for consumer safety. However, existing methods have limitations that prevent their direct use for active drug surveillance. One of the important consideration that has been absent thus far is the modeling of multiple adverse events and their interactions. In this paper, we propose a method to monitor the effect of a single drug on multiple adverse events, which explicitly captures interdependence between events. Our method uses a sequential hypothesis testing paradigm, and employs an intuitive test-statistic. Stopping boundaries for the test-statistic are designed by asymptotic analysis and by reducing the design problem to a convex optimization problem. We apply our method to a dynamic version of Cox's proportional hazards model, and show both analytically and numerically how our method can be used as a test for the hazard ratio of the drug. Our numerical studies further verify that our method delivers Type I/II errors that are below pre-specified levels and is robust to distributional assumptions and parameter values.

Session TA3 - Simulation Optimization 1, Chair: Fu / Henderson / Zhou in JVdC C

Interactive Model-based Search for Global Optimization

Yuting Wang, University of Virginia, yw6h@virginia.edu
Alfredo Garcia, University of Virginia, ag7sr@virginia.edu

Single-thread algorithms for global optimization differ in the way computational effort between exploitation and exploration is allocated. This allocation ultimately determines overall performance. For example, if too little emphasis is put in exploration, a globally optimal solution may not be identified. Increasing the allocation of computational effort to exploration increases the chances of identifying a globally optimal solution but it also slows down convergence. In this paper we propose a new algorithmic design for global optimization based upon multiple interacting threads. In this design, each thread implements a model-based search in which the allocation of exploration versus exploitation effort does not vary over time. Threads interact through a simple acceptance-rejection rule preventing duplication of search efforts. We show that the speed of convergence (both worst-case and average) increases exponentially in the number of threads. Thus, in the proposed algorithmic design, exploration is a complement and not a substitute to exploitation.

Sequential Monte Carlo Multi-model-based Optimization

Xi Chen, Department of Industrial & Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, IL 61801, USA, xchen37@illinois.edu
Enlu Zhou, Department of Industrial & Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, IL 61801, USA, enluzhou@illinois.edu

Model-based optimization methods are stochastic search methods that iteratively find candidate solutions by generating samples from a parameterized probabilistic model on the solution space and update the parameter of the model based on the performance of the candidate solutions. In this work, we propose a multi-model-based method, sequential Monte Carlo multi-model-based optimization (SMCMO), where the candidate solutions are generated from multiple probabilistic models at each iteration. Using multiple models helps to better capture the multi-modal property of the objective function than only using a single model in traditional model-based methods. To develop a mechanism to generate and adaptively propagate multiple models, we view the optimization problem as a parameter estimation problem that iteratively estimates the parameter of the optimal probabilistic model based on the observations or performance of the candidate solutions. We conduct the parameter estimation by sequential Monte Carlo method, which tracks the posterior distribution of the parameter given the history of observations, providing a proper way to determine the diversity of the models based on the spread of the posterior distribution of the parameter. We prove the convergence of SMCMO, and carry out numerical experiments to illustrate its performance.

Global optimization with noise corrupted function evaluations

Jim Calvin, New Jersey Institute of Technology, Newark, New Jersey, USA

In many practical problems, one is interested in approximating the minimum cost associated with some system based on partial information, such as sequentially chosen function evaluations. Sometimes the exact function values are not available, and the optimizer must make do with noisy estimates of the values. For example, if one wishes to minimize the average cost of operating a stochastic system for which no tractable analytic formula is available, then the optimizer might use discrete-event simulation to estimate the cost at a sequence of adaptively chosen parameter values. Typical simulation estimators, when suitably normalized, have an approximately normal distribution for long simulation run lengths. To model this situation, we consider the problem of approximating the minimum of a continuous function using sequentially chosen function evaluations corrupted by normally distributed noise. We describe an optimization algorithm for which the error converges to zero at the optimal rate up to logarithmic factors.

A Bayesian Approach to Stochastic Root Finding

S.G. Henderson, Cornell University, New York NY, USA, sgh9@cornell.edu
P.I. Frazier, Cornell University, New York NY, USA, pf98@cornell.edu
R. Waeber, Cornell University, New York NY, USA, rw339@cornell.edu

We consider the problem of finding the root of a monotone function on the one-dimensional interval [0, 1] when only noisy function evaluations are available. The probabilistic bisection algorithm (PBA) begins with a prior distribution on the location of the root, and successively updates this prior to reflect one's belief about the location of the root. In doing so, it needs to know the probability that the sign of the noisy function evaluation is correct. In practice, this information is not available. We explore a method for circumventing this problem, and develop convergence results. If a certain conjecture holds, then the rate of convergence of the PBA in terms of the number of simulation replications is arbitrarily close to, but slightly slower than, the asymptotic rate of convergence of stochastic approximation. The PBA has the advantages that it performs very stably for finite sample sizes, and can return confidence intervals for the location of the root.

Session TA4 - Quantitative Financial Risk Management, Chair: Zhang / Ghamami in JVdC D

A Systemic Risk Model for Asset Price Contagion

Chen Chen, Industrial Engineering and Operations Research, Columbia University, cc3136@columbia.edu
Garud Iyengar, Industrial Engineering and Operations Research, Columbia University, garud@ieor.columbia.edu
Ciamac C. Moallemi, Graduate School of Business, Columbia University, ciamac@gsb.columbia.edu

We develop a structural model for the analysis of systemic risk in financial markets based on asset price contagion. Specifically, we describe a mechanism of contagion where exogenous random shocks to individual agents in an economy forces portfolio rebalancing and impact asset prices. This, in turn, creates an endogenous chain reaction as downstream agents trade in reaction to price changes. We demonstrate an approach that quantifies the robustness of different financial networks to shocks propagated by asset price contagion.

Systemic Risk with Central Counterparty Clearing

Hamed Amini, Swiss Federal Institute of Technology (EPFL), Switzerland, hamed.amini@epfl.ch

We study a financial network in a stochastic framework. We measure systemic risk in terms of a coherent valuation principle. The framework allows us to examine the effects on systemic risk and price contagion of multilateral clearing via a central clearing counterparty (CCP). We prove existence and uniqueness of an interbank payment equilibrium in conjunction with the price impact on external assets. We find that a CCP not always reduces systemic risk and provide sufficient conditions for the latter to hold. We also propose an optimal capitalization of a CCP based on game theoretic arguments. A real world calibrated numerical study illustrates our findings. This is based on joint work with Damir Filipovic and Andreea Minca.

Efficient Monte Carlo Counterparty Credit Risk Pricing and Measurement

Samim Ghamami, UC Berkeley, USA, samim_ghamami@berkeley.edu
Bo Zhang, IBM Research, USA, bozhang@gatech.edu

Counterparty credit risk (CCR), a key driver of the 2007-08 credit crisis, has become one of the main focuses of the major global and U.S. regulatory standards. Financial institutions invest large amounts of resources employing Monte Carlo simulation to measure and price their counterparty credit risk. We develop efficient Monte Carlo CCR frameworks by focusing on the most widely used and regulatory-driven CCR measures: expected positive exposure (EPE), credit value adjustment (CVA), and effective expected positive exposure (eEPE). Our numerical examples illustrate that our proposed efficient Monte Carlo estimators outperform the existing crude estimators of these CCR measures substantially in terms of mean square error (MSE). We also demonstrate that the two widely used sampling methods, the so-called Path Dependent Simulation (PDS) and Direct Jump to Simulation date (DJS), are not equivalent in that they lead to Monte Carlo CCR estimators which are drastically different in terms of their MSE.

Session TA5 - Many-server queues: Approximations and control, Chair: Gurvich in JVdC E

Sacrificing some optimality in an overload control to achieve rapid recovery

Ohad Perry, Northwestern University, Evanston, IL, ohad.perry@northwestern.edu
Ward Whitt, Columbia University, New York, NY, ww2040@columbia.edu

We consider how two large service pools, each primarily dedicated to one class of customers but is capable of serving both classes, can help each other in a **time-varying environment** involving periods of overloads. In particular, we assume that the arrival rates and number of agents in either service pool are time dependent, and we seek an optimal routing policy ("sharing" of customers among the two pools), assuming a holding cost is incurred on both queues. We suggest a Fixed-Queue Ratio with Activation-and-Release Thresholds (FQR-ART) control, whose aim is to: (i) Activate sharing in one direction when it is optimal to do so, and then keep the two queues at one of two fixed ratios (depending on the direction of sharing); (ii) switch the order of sharing when the direction of overload changes; and (iii) stop sharing when the system returns to normal loading. The FQR-ART control has the desired properties of being automatic and simple to apply. However, we show that there is a need to sacrifice some optimality so as to avoid undesirable sharing and a resulting oscillatory behavior which can lead to severe degradation in the system's performance. To study this nonstationary system we employ a deterministic fluid approximation, which is described implicitly via a time-inhomogeneous ordinary differential equation.

Sizing Step-Down Units in Hospitals

M. Armony, Stern School of Business, New York University, marmony@stern.nyu.edu
C.W. Chan, Columbia Business School, Columbia University, cwchan@columbia.edu
B. Zhu, Courant Institute of Mathematical Sciences, New York University, zhubo@cims.nyu.edu

The use of Step Down units (SDUs) in hospitals has been a matter of controversy in the medical community. On one hand, an SDU alleviates ICU congestion by providing a safe care environment for post-ICU patients before they can be transferred to the general wards. On the other hand, SDUs can take capacity away from the already over-congested ICU. We propose a queueing model that provides SDU sizing guidelines, noting that under some circumstances this optimal size is zero. Our methodology is focused on fluid and diffusion asymptotic analysis. The results are compared against an extensive simulation study.

Many-Server Heavy-Traffic Limits for Queueing Networks with Time-Varying Parameters and Probabilistic Routing

Y. Liu, North Carolina State University, Raleigh, NC, yunan_liu@ncsu.edu
C. C. Huang, North Carolina State University, Raleigh, NC, chuang10@ncsu.edu

We establish a many-server heavy-traffic *functional weak law of large numbers* (FWLLN) and *functional central limit theorem* (FCLT) for the $(G_t/M/s_t + GI)^m/M_t$ queueing network with time-varying parameters (e.g., arrival rates and staffing functions), non-exponential distributions and Markovian routing structure. The analysis is based upon the assumption that all queues of the network alternate between underloaded and overloaded, and are never critically loaded except at isolated switching points. The extension of the FWLLN and FCLT from single-queue $G_t/M/s_t + GI$ models to network models draws heavily on careful treatments to the total arrival processes of the network model, namely, the processes consisting external arrivals and internal feedback flows. The established FCLT limits, i.e., the diffusion processes, are characterized by multi-dimensional Brownian *stochastic differential equations* (SDE's) that are analytically tractable. Useful variance formulas are developed for the purpose of engineering approximation. Besides mathematical proofs, we provide concrete computer simulation experiments as an engineering proof for correctness of these theoretical formulae.

Scaled control in the QED regime

Johan van Leeuwaarden, Eindhoven University of Technology, The Netherlands, j.s.h.v.leeuwaarden@tue.nl
Guido Janssen, Eindhoven University of Technology, The Netherlands, a.j.e.m.janssen@tue.nl
Jaron Sanders, Eindhoven University of Technology, The Netherlands, jaron.sanders@tue.nl

We develop many-server asymptotics in the QED regime for models with admission control. The admission control scales with the size of the system. For a rich class of Markovian models with scaled control, we present refinements to many-server asymptotics and optimality gaps related to square-root staffing, generalizing earlier results for the Erlang B, C and A models.

Session TA6 - Queueing Networks, Chair: D'Auria / Van Leeuwaarden in JVdC F

Mixing properties and delay performance of random-access networks

A. Zocca, Eindhoven University of Technology, a.zocca@tue.nl
S.C. Borst, Eindhoven University of Technology, s.c.borst@tue.nl
J.S.H. van Leeuwaarden, Eindhoven University of Technology, j.s.h.v.leeuwaarden@tue.nl
F.R. Nardi, Eindhoven University of Technology, f.r.nardi@tue.nl

We consider a stylized stochastic model for a wireless random-access network, which yields a product-form stationary distribution of the activity process for the various users and provides useful estimates for the user throughputs. Even in symmetric scenarios, where long-run fairness is guaranteed, it may take a long time for the activity process to move between dominant states, giving rise to potential starvation issues. In order to gain insight in the transient throughput characteristics and associated starvation effects, we examine the behavior of the transition time between dominant activity states and the convergence rate to equilibrium of the activity process in terms of conductance and mixing times, for various interference graphs. In particular, we establish the order of magnitude of the transition time for a symmetric grid network and demonstrate how these results for the transition times can be exploited to obtain delay bounds. We also prove that in some cases the scaled transition time has an asymptotically exponential distribution as the activation rate grows large, and point out interesting connections with related exponentiality results for rare events, renewal theory and meta-stability phenomena in statistical physics.

Online Optimization of Product-Form Networks

J. Sanders, Eindhoven University of Technology, The Netherlands, jaron.sanders@tue.nl
S.C. Borst, Bell Laboratories, Alcatel-Lucent, United States of America, sem@research.bell-labs.com; Eindhoven University of Technology, The Netherlands, s.c.borst@tue.nl
J.S.H. van Leeuwaarden, Eindhoven University of Technology, The Netherlands, j.s.h.v.leeuwaarden@tue.nl

We have developed an online algorithm for optimizing product-form networks through adjustment of system parameters. The algorithm circumvents difficult exact calculations of stationary distributions by measuring empirical frequencies. In essence, the algorithm implements stochastic gradient descent to minimize an objective function, basing decisions on observations of the underlying stochastic process.
There is considerable flexibility in choosing step sizes, lengths of observation periods, types of gradients and objective functions. In order to ensure good performance of the algorithm, it is essential to identify conditions for almost sure convergence to the optimal parameters. We discuss how mixing time results for product-form networks can be exploited to gain insight in these convergence issues.

Beyond strict insensitivity

Matthieu Jonckheere, CONICET, Universidad de Buenos Aires, Argentina, matthieu.jonckheere@gmail.com
Sergio Lopez, UNAM, Mexico.

One of the most notorious historical success (perhaps miracle) of queuing theory is the Erlang formula and its extensions to loss networks. Its applicability relies on its robustness to the call durations distribution: the well known insensitivity property. Attempts to extend this formula to various modern networks (e.g. with best effort traffic) have been only partially successful, due to the increased complexity and intrinsic asymmetries of the corresponding models. We examine here a relaxed robustness property of stationary queuing systems: the insensitivity of the large deviations characteristics to the distribution of job sizes. We show that this property holds for some fundamental models like bandwidth sharing networks under the proportional fair allocation, giving hope to develop efficient performance evaluation tools. In other cases, generic bounds for the large deviations characteristics of the stationary regime are expected.

Congestion in processor-sharing multi-class closed networks: stationary and fluid analysis

Jonatha Anselmi, Basque Center for Applied Mathematics (BCAM), Spain, anselmi@bcamath.org
Bernardo D'Auria, Universidad Carlos III de Madrid, Spain, bernardo.dauria@uc3m.es
Neil Walton, University of Amsterdam, The Netherlands, n.s.walton@uva.nl

The congestion analysis of a queueing network is essential to understand its behavior under high-load conditions, and to find the weak links, i.e. its bottlenecks. The analyzed model is a processor-sharing multi-class Jackson network, and it is known that the set of bottleneck stations is a function of the distribution of customers among the various classes. What is in general not known is the behavior of the subnet made of the stable nodes.
In our study we show that the stable part of the network behaves, in stationary regime, as an open Jackson network where the congested nodes are substituted by Poisson sources of customers with fixed rates that are given by the solution of a strictly concave optimization problem. In addition, analyzing the same network by a fluid limit analysis, we again get that the evolution of the corresponding hydrodynamical system converges to an equilibrium configuration that concentrates all fluid in the bottleneck stations and whose class composition agrees with the one predicted by the probabilistic stationary distribution.

Session TA7 - Optimal Control of Stochastic Inventory Systems, Chair: Squillante in JVdC G

Asymptotic Optimality of Constant-Order Policies for Lost Sales Inventory Models with Large Lead Times

D.A. Goldberg, Georgia Institute of Technology, Georgia, USA, dgoldberg9@isye.gatech.edu
D.A. Katz-Rogozhnikov, IBM Thomas J. Watson Research Center, New York, USA, dkatzrog@us.ibm.com
Y. Lu, IBM Thomas J. Watson Research Center, New York, USA, yingdong@us.ibm.com
M. Sharma, IBM Thomas J. Watson Research Center, New York, USA, mxsharma@us.ibm.com
M.S. Squillante, IBM Thomas J. Watson Research Center, New York, USA, mss@us.ibm.com

Lost sales inventory models with large lead times, which arise in many practical settings, are notoriously difficult to optimize due to the curse of dimensionality. In this talk we show that when lead times are large, a very simple constant-order policy, first studied by Reiman, performs nearly optimally. The main insight of our work is that when the lead time is very large, such a significant amount of randomness is injected into the system between when an order for more inventory is placed and when that order is received, that "being smart" algorithmically provides almost no benefit. Our main proof technique combines a novel coupling for suprema of random walks with arguments from queueing theory.

Matching Supply and Demand in Production-Inventory Systems: Asymptotics and Insights

M.S. Squillante, IBM Thomas J. Watson Research Center, New York, USA, mss@us.ibm.com
Y. Lu, IBM Thomas J. Watson Research Center, New York, USA, yingdong@us.ibm.com
D.D. Yao, Columbia University, New York, USA, yao@columbia.edu

We consider a general class of production-inventory systems based on both lost-sales and backorder inventory models. Such systems require a fundamental understanding of the asymptotic behavior of key performance measures under various supply strategies, as well as the pre-planning of these strategies. Our analysis relies on a thorough study of the asymptotic behavior of a random walk with power drift, which is of independent interest. In addition to providing key insights, our analysis leads to approximations of the corresponding optimization problem that yield simple solutions which are close to optimal. We also establish an equivalence between the lost-sales and backorder models when both have the same penalty cost that becomes large. Numerical comparisons of various approaches are presented, quantifying the accuracy of our approximations with respect to the exact solution.

Session TA8 - Estimation, Chair: Pasupathy in JVdC H

Estimating Waiting Times with the Time-Varying Little's Law

Song-Hee Kim, Columbia University, New York, NY, sk3116@columbia.edu
Ward Whitt, Columbia University, New York, NY, ww2040@columbia.edu

When waiting times cannot be observed directly, Little's law can be applied to estimate the average waiting time by the average number in system divided by the average arrival rate, but that simple indirect estimator tends to be biased significantly when the arrival rates are time-varying and the service times are relatively long. Here it is shown that the bias in that indirect estimator can be estimated and reduced by applying the time-varying Little's law (TVLL). If there is appropriate time-varying staffing, then the waiting time distribution may not be time-varying even though the arrival rate is time-varying. Given a fixed waiting time distribution with unknown mean, there is a unique mean consistent with the TVLL for each time t. Thus, under that condition, the TVLL provides an estimator for the unknown mean wait, given estimates of the average number in system over a subinterval and the arrival rate function. Useful variants of the TVLL estimator are obtained by fitting a linear or quadratic function to arrival data. When the arrival rate function

is approximately linear (quadratic), the mean waiting time satisfies a quadratic (cubic) equation. The new methods are shown to be effective in estimating the bias in the indirect estimator and reducing it, using simulations of multi-server queues and data from a call center.

Copula-based reliability estimation for multicomponent repairable systems via time-truncated NHPP with power law intensity

M. S. Aminzadeh, Department Mathematics, Towson University,
Maryland, USA, maminzadeh@towson.edu

We present estimation of $R = P(S_{N_1(T)+1} < S_{N_2(T)+1})$ for two dependent non-homogenous Poisson processes (NHPPs), $\{N_1(t), t > 0\}$ and $\{N_2(t^*), t^* > 0\}$ under time-truncated sampling scheme. $S_{N_j(T)+1}$, $j = 1, 2$ denotes the arrival time of the first event past a truncated time T. Clayton copula and Power Law intensity function for NHPP are used. A modified multistage inference functions for margins (MMIFM) method is proposed that provides equal-sized samples extracted from the initial samples of arrival times from the two NHPPs. The extracted samples are used to maximize the part of the likelihood function that measures the dependency structure between the two processes. The accuracy of the estimator for R is verified by simulations studies.

"Online" Quantile and Density Estimators

Raghu Pasupathy, Virginia Tech, VA, USA,pasupath@vt.edu
Soumyadip Ghosh, IBM Research, NY, USA, ghoshs@us.ibm.com

The standard estimator $\hat{q}_\alpha(n)$ for the α-quantile q_α of a random variable X, given n observations from the distribution of X, is obtained by inverting the empirical cumulative distribution function (cdf) constructed from the obtained observations. It is well-known that $\hat{q}_\alpha(n)$ requires $O(n)$ storage, and that the mean squared error of $\hat{q}_\alpha(n)$ (with respect to q_α) decays as $O(n^{-1})$. In this talk, we present an alternative to $\hat{q}_\alpha(n)$ that seems to require dramatically less storage with negligible loss in convergence rate. The proposed estimator, $\tilde{q}_\alpha(n)$, relies on an alternative cdf that is constructed by accumulating the observed random variates into variable-sized bins that progressively become finer around the quantile. The size of the bins are strategically adjusted to ensure that the increased bias due to binning does not adversely affect the resulting convergence rate. We will present an "online" version of the estimator $\tilde{q}_\alpha(n)$, along with results on its consistency, convergence rates, and storage requirements. If time permits, we will discuss analogous ideas for density estimation.

Tuesday 11:30am - 1:00pm

Session TB1 - Random Graphs and Complex Networks: Social Networks, Chair: Olvera-Cravioto in JVdC A

Super Star Model: Predicting the Structure of Retweet Graphs in Twitter

Tauhid Zaman, MIT Sloan School of Management, zlisto@mit.edu
Shankar Bhamidi, University of North Carolina, bhamidi@email.unc.edu
J. Michael Steele, The Wharton School of the University of Pennsylvania, steele@wharton.upenn.edu

Twitter is a micro-blogging site where users post messages known as "tweets" and repost the tweets of others via an act known as "retweeting". We observe the following interesting phenomenon in the retweet graphs associated with a wide array of topics in Twitter: there is always a single vertex whose degree is on the order of the graph size. We refer to this vertex as a super star. To model this phenomenon we propose the super star model, which is a variation off the preferential attachment model that allows for one vertex (the super star) to have degree that grows linearly in the graph

size. This model predicts a relationship between the super star degree and the degree distribution of non-super star vertices. We find that this relationship exists in real Twitter graphs. Furthermore, we show that the super star model predicts the degree distribution of these Twitter graphs more accurately than preferential attachment.

Super Star Model: The Ghost in the Machine

J. Michael Steele, The Wharton School of the University of Pennsylvania, steele@wharton.upenn.edu
Shankar Bhamidi, University of North Carolina, bhamidi@email.unc.edu
Tauhid Zaman, MIT Sloan School of Management, zlisto@mit.edu

The super star model is a variation of the preferential attachment model that allows for one vertex (the super star) to have degree that grows linearly in the network size. Although the model specification is simple, the analysis of the model is subtle and features emerge that make it quite distinct from the plain vanilla preferential attachment model. This talk focuses on the tools that support the analysis of the super star model and many related network models that can be built by surgery on multi-type branching processes.

The Logarithmic Dimension Hypothesis

Anthony Bonato, Department of Mathematics, Ryerson University, abonato@ryerson.ca

On-line social networks (OSNs) such as Facebook and LinkedIn have transformed our paradigms of communication and social interaction. The recent Geometric Protean or GEO-P model exploits a hypothesized underlying social space (sometimes called *Blau space*), where nodes are close via a prescribed metric if they share similar attributes. In this model, nodes are linked with probability based on both their relative proximity and via a ranking scheme. A key prediction of this model is that a small number of network statistics (including network order, power law exponent, average degree, and diameter) give an estimate on the dimension of the network; that is, the least number of attributes needed to identify agents in the networks. The dimension is predicted to be approximately the logarithm of the number of nodes. This so-called *Logarithmic Dimension Hypothesis* (*LDH*) would therefore predict a relatively small number of attributes needed to identify agents, and may have many potential applications to social networks.

We present evidence for the LDH based model selection techniques exploiting machine learning. Such techniques have been used successfully in the past for protein interaction networks, and more recently in OSNs by Janssen et al. By exploiting a simplified version of the GEO-P model which shares most of its asymptotic properties, we provide evidence for the dimensionality of Blau space for the Facebook 100 graphs and samples from LinkedIn. Additional evidence for the conjecture is provided by considering eigenvalue distributions.

Co-Evolutionary Models of Community Structure in Networks

Pawel Pralat, Department of Mathematics, Ryerson University, pralat@ryerson.ca

Community structure and network segregation-where two or more groups of homogeneous actors form dense subgraphs with relatively few linkages across groups-is a salient feature of many real-world networks. While much work has been done on approaches to identify communities within networks, relatively few studies have developed theoretical models of how and why community structure emerges. Even fewer studies have accounted for the coevolution of nodal attributes and link structures-a commonly-observed phenomenon within many complex, self-organizing networks. This research proposes a general model of network self-organization driven by both homophily (where nodes with similar attributes are more likely to be connected) and contagion processes (where connected nodes become more similar to one another over time). While previous work has shown that even very small levels of homophily are sufficient to produce community structures in networks when attributes are held constant, we find that adding even very small levels of contagion are sufficient to prevent community structures from emerging-even in the presence of relatively strong homophily processes. This research paves the way for more complex models of co-evolving attributes and link structures, a necessary step in the development of better theories of community structure in networks.

Session TB2 - Stochastic Control and Revenue Management in Smart Grids, Chair: Zhang in JVdC B

Voltage and Reactive Power Control for Power Loss Minimization Using Approximate Stochastic Annealing

Eugene Feinberg, Applied Math, Stony Brook University, Stony Brook, NY, eugene.feinberg@stonybrook.edu
Jiaqiao Hu, Applied Math, Stony Brook University, Stony Brook, NY, jqhu@ams.sunysb.edu
Eting Yuan, Applied Math, Stony Brook University, Stony Brook, NY, etingyuan@gmail.com

Efficient management of voltage profiles and reactive power in power distribution systems plays an important role towards this goal. In this paper, we propose to use an approximate stochastic annealing (ASA) algorithm for solving the voltage and reactive power control (VVC) problem. The objective is to determine the proper settings of capacitor banks and transformer taps in a power distribution system to minimize daily energy losses. Several types of constraints, such as voltage constraints and operation limits constraints on transformer load tap changers (LTCs) and shunt capacitors (SCs), are considered in our model. In this work, the original ASA algorithm is adapted for solving the VVC problem by sampling from a sequence of probability distributions over the space of all possible configurations of LTCs and SCs. A Lagrangian Relaxation-Dynamic Programming (LR-DP) algorithm is also proposed to obtain upper and lower bounds on the performance of the optimal solution. The performance of the ASA algorithm is illustrated on a well-known PG&E 69-bus distribution network. Our testing results indicate that the ASA algorithm may yield solutions very close to optimum within a modest amount of computational time.

Dynamic electricity retail pricing in uncertain environments

Liyan Jia, Electrical and Computer Engineering, Cornell University, Ithaca, NY 14850, liyanjia1987@gmail.com
Lang Tong, Electrical and Computer Engineering, Cornell University, Ithaca, NY 14850, ltong@ece.cornell.edu

The problem of optimizing dynamic electricity retail price for residential consumers is considered. A two stage retail market structure is modeled as a dynamic game between the retailer and the consumers. Based on the optimal demand response obtained from the thermal dynamic loads, the trade-off between consumer surplus and retail profit is characterized by a concave and non-increasing Pareto front. It is shown that each point on the Pareto front corresponds to an equilibrium point in the dynamic game with a particular payoff function, and any consumer surplus-retail profit pair above the Pareto front is not attainable by any dynamic pricing scheme. Effects of renewable energy are also considered.

An inventory theory framework for the analysis of probabilistic demand response schemes

Pavithra Harsha, IBM Watson Research, Yorktown Heights, NY 10598, pharsha@us.ibm.com
Mayank Sharma, IBM Watson Research, Yorktown Heights, NY 10598, mxsharma@us.ibm.com
Ramesh Natarajan, IBM Watson Research, Yorktown Heights, NY 10598, nramesh@us.ibm.com
Soumyadip Ghosh, IBM Watson Research, Yorktown Heights, NY 10598, ghoshs@us.ibm.com

This work has two main contributions. First we conceptually describe probabilistic demand response schemes and argue that such schemes are particularly well-suited for dynamically managing demand in the residential consumption sector. However, these schemes are relatively new and there is little or no theoretical guidance for their development and evaluation. Our second contribution is therefore to describe a methodology for analyzing these schemes which incorporates the respective costs of supply generation, spinning reserve, and storage or deferment, into an operational objective function that is similar to the well-known newsvendor model, but with suitable modifications for the smart grid context. We derive sufficient conditions on the distributions of the responsive load in order to optimize this operational objective as a function of the supply generation level, and the magnitude of the response signal. These conditions are used to examine the suitability of various probabilistic demand response schemes in the single-period and multiple-period contexts.

Dynamic Pricing for Reliably Shaping of Electricity Demand

Xiaoxuan Zhang, IBM Watson Research, Yorktown Heights, NY 10598, zhangxiaoxuan@live.com
Jayant Kalagnanam, IBM Watson Research, Yorktown Heights, NY 10598, jayant@us.ibm.com
Dmitriy Katz-Rogozhnikov, IBM Watson Research, Yorktown Heights, NY 10598, dkatzrog@us.ibm.com
Mark Squillante, IBM Watson Research, Yorktown Heights, NY 10598, mss@us.ibm.com
Soumyadip Ghosh, IBM Watson Research, Yorktown Heights, NY 10598, ghoshs@us.ibm.com

We propose two dynamic pricing schemes to address these challenges in retail electricity setting, i.e., non-differential and differential (customized) dynamic pricing. For the non-differential dynamic pricing, all the residential customers are offered with the same incentive at each period, and the incentive is time-varying; For the customized dynamic pricing, each customer is offered with an incentive price depending on their demand variability and reliability of their responsive load. The goal is to determine the aggregate electricity demand reduction with high reliability (low risk) that maximizes the utility companyÕs long-term average revenue. Both dynamic incentive pricing schemes improve the utility companyÕs revenue by utilizing virtual generation, i.e., demand reduction, and achieve improvements in both total systemic costs and load reduction over existing schemes. In addition, customized pricing proves to be superior to the non-differential pricing in the sense of reliable aggregate demand shift and improved long-term average revenue, since it is based on the customer segmentation according to their reduction reliability and achieves an improved incentive resource allocation among all the customers. The residential customersÕ real-time price elasticity is estimated from the smart metering data of the OlyPen project. Extensive simulation experiments show that for each single period, customers with higher elasticity and lower variance should be paid with higher incentive rates; and along multiple periods, customers with smaller likelihood of shifting their load and greater inclination to consume less over the entire horizon should be given higher rebates. The proposed dynamic pricing mechanism also improves the social welfare for both electricity suppliers and buyers.

Session TB3 - Simulation Optimization 2, Chair: Fu / Henderson / Zhou in JVdC C

Maximizing Quantitative Traits in the Mating Design Problem

Susan R. Hunter, School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA, hunter@cornell.edu
Benjamin McClosky, Nature Source Genetics, 33 Thornwood Drive, Ithaca, NY 14850, USA, bmcclosky@naturesourcegenetics.com

Maximizing Quantitative Traits in the Mating Design Problem

We consider a version of the mating design problem in which breeders allocate a breeding budget to a set of parent pairs to maximize the expected maximum trait observed in the progeny population. In this context, the only parent pairs that receive nonzero breeding budget at optimality in the mating design problem lie on a Pareto set. Since the performance of each parent pair is assessed through Monte Carlo simulation, identifying the Pareto set is a bi-objective simulation optimization problem. We derive an asymptotically optimal simulation budget allocation to estimate the Pareto set of parent pairs that, in our numerical experiments, out-performs MOCBA in reducing misclassifications. This estimated Pareto set is used as an input to the mating design problem, which is an integer program.

Ranking and Selection with Tight Bounds on Probability of Correct Selection

Peter I. Frazier, Cornell University, USA, pf98@cornell.edu

We consider a classic problem from simulation optimization, indifference-zone ranking and selection. Although this problem is well-studied, existing procedures have loose bounds on solution quality, leading them to sample more than necessary. We construct the first sequential elimination procedure whose bounds on worst-case probability of

correct selection (for more than two alternatives) are tight. Tight bounds allow this procedure to deliver solutions with fewer simulation samples. We assume independent normal samples, and consider settings with both common known variance, and heterogeneous unknown variance. The procedure, called the Bayes-inspired Indifference Zone (BIZ) procedure, avoids the Bonferonni inequality, instead using a novel symmetry based in Bayesian analysis, which also provides new results about hitting probabilities of geometric Brownian motion to the faces of a polytope.

Chance-constrained selection of the best

B.L. Nelson, Northwestern University, U.S.A., nelsonb@northwestern.edu
L.J. Hong, Hong Kong University of Science and Technology, Hong Kong, hongl@ust.hk
J. Luo, Hong Kong University of Science and Technology, Hong Kong, lawren290911@gmail.com

Selecting the solution with the largest or smallest mean of a primary performance measure from a finite set of solutions while requiring secondary performance measures to satisfy certain constraints is called constrained selection of the best (CSB). We consider CSB problems whose secondary performance measures are subject to probabilistic constraints. We design procedures that first check the feasibility of all solutions and then select the best among all of the sample feasible solutions; we prove the statistical validity of these procedures for variations of the problem; and we show via numerical results that they are efficient.

Stochastic Dynamic Decision Simulation: Markov Chains in Multi-stage Stochastic Programming

Suvrajeet Sen, University of Southern California, Los Angeles, CA, s.sen@usc.edu

Stochastic Programming (SP) grew out of the need to incorporate probabilistic information into constrained optimization models such as Linear Programming. Unfortunately, the nature of probabilistic information that is currently allowed in Multi-stage Stochastic Programming leaves much to be desired. For instance, most sampling-based algorithms are unable to guarantee asymptotic convergence for instances in which the stochastic process exhibits dependence between stages of the multi-stage model. In this talk, we will illustrate that the Multi-stage Stochastic Decomposition (MSD) algorithm ensures asymptotic convergence (wp1) even if stagewise independence is violated. Indeed for Markov chains, we can show that the MSD approximations can be streamlined in a manner that requires at most m approximations per stage, where m denotes the number of states of the Markov chain. Interestingly, no further discretization of the state-space is necessary in the multi-stage SLP setting.

Session TB4 - Stochastic Calculus and Finance, Chair: Chigansky in JVdC D

On the Markov property of some Brownian martingales

J.Y. Fan, Monash University, Australia, jieyen.fan@monash.edu
K. Hamza, Monash University, Australia, kais.hamza@monash.edu
F.C. Klebaner, Monash University, Australia, fima.klebaner@monash.edu

Let h_n be the (probabilists') Hermite polynomial of degree n. Let $H_n(z,a) = a^{n/2}h_n(z/\sqrt{a})$ and $H_n(z,0) = z^n$. It is well-known that $H_n(B_t,t)$ is a martingale for every n. In this paper, we show that for $n \geq 3$, $H_n(B_t,t)$ is not Markovian. We then give a brief discussion on mimicking $H_n(B_t,t)$ in the sense of constructing martingales whose marginal distributions match those of $H_n(B_t,t)$.

On a class of Stochastic Implied Volatility models

F.C. Klebaner, Monash University, Australia, fima.klebaner@monash.edu

Stochastic Implied volatility models aim at constructing a pair of stochastic processes, the price process and the implied volatility process, such that the Black-Scholes formula holds true for all strikes and maturities. There is no result that assures existence of such model in general. We show a class of such processes where existence follows by construction, however parameters are given as solutions of some differential equations and explicit formulae are hard to find.

Mixed fractional Brownian motion: the filtering perspective

P.Chigansky, The Hebrew University of Jerusalem, Israel, pchiga@mscc.huji.ac.il
M. Kleptsyna, Universite du Maine, France, marina.kleptsyna@univ-lemans.fr

We present an alternative approach to studying the properties of the mixed fractional Brownian motion, based on the filtering theory of Gaussian processes. It yields a new proof of the regularization theorem due to P. Cheridito and resolves an open question, concerning statistical estimation of the unknown drift parameter observed in mixed fractional noise.

Session TB5 - Fluid limits and its applications, Chair: Van Houdt in JVdC E

Mean field approximation meets stochastic model checking

L. Bortolussi, University of Trieste, Italy, luca.bortolussi@gmail.com

Mean field approximation techniques are well established approaches to analyse large-scale stochastic processes, especially population Markov models. Some years ago, they have entered the arena of quantitative formal methods, and they provided a consistent way to define fluid semantics for stochastic process algebras (formal languages that can be used to describe population Markov models, and more generally continuous time Markov chains, in terms of interacting agents, and reason algebraically about them).
Here, we will discuss how mean field approximation can play a role in another successful area of quantitative formal methods, namely stochastic model checking. In particular, we will focus on a subclass of Continuous Stochastic Logic (CSL), considering formulae that expresses properties of single agents in a population model, and we will present an approximate model checking algorithm based on mean field approximation. We will also discuss model checking CSL formulae against time-inhomogeneous CTMC models, as this turns out to be the core procedure needed for fluid approximation. Finally, we will consider a class of global properties that can be analysed by linear noise approximation, a higher order fluid approximation.

Some fluid models in performance and dependability analysis

L.Aspirot, Universidad de la República, Montevideo, Uruguay, `aspirot@fing.edu.uy`
E. Mordecki, Universidad de la República, Montevideo, Uruguay, `mordecki@cmat.edu.uy`
G. Rubino, INRIA, Rennes, France, `gerardo.rubino@inria.fr`

Many probabilistic models, designed for evaluating systems' performance or dependability, lead to complex and "large" mathematical objects that can only been analyzed using simulation. The bad side of this situation is the lack of qualitative insight that is often associated with analytical methods. Fluid limits and mean field theory, are possibles approaches that can be tried in some cases (typically, when we deal with the so-called population models), allowing to transform the original stochastic process into a very "small" deterministic one.
We will discuss technical details of two on-going studies where we analyze performance aspects of communication networks and dependability properties of basic models of repairing services. Specifically, we explore the relations between the stochastic and the related deterministic models, in two cases: Bit-Torrent-like P2P networks with some variants in the way the network is structured, and the basic Machine Repairman Model. We will focus on the convergence of the stochastic model towards the deterministic one as the size of the former goes to infinity, and the relations between the stationary distributions of the stochastic models (when their sizes vary) and the fixed points of the deterministic related models.

A Mean Field Model for a Class of Garbage Collection Algorithms in Flash-based Solid State Drives

Benny Van Houdt, University of Antwerp, Belgium, benny.vanhoudt@ua.ac.be

Garbage collection (GC) algorithms play a key role in reducing the write amplification in flash-based solid state drives, where the write amplification affects the lifespan and speed of the drive. This paper introduces a mean field model to assess the write amplification and the distribution of the number of valid pages per block for a class $\mathcal{C}$ of GC algorithms. Apart from the RANDOM GC algorithm, class $\mathcal{C}$ includes two novel GC algorithms: the d-CHOICES GC algorithm, that selects d blocks uniformly at random and erases the block containing the least number of valid pages among the d selected blocks, and the RANDOM++ GC algorithm, that repeatedly selects another block uniformly at random until it finds a block with a lower than average number of valid blocks.
We show that the d-CHOICES GC algorithm has a write amplification close to that of the GREEDY GC algorithm even for small d values, e.g., $d = 10$, and offers a more attractive trade-off between its simplicity and its performance than the WINDOWED GC algorithm introduced and analyzed in earlier studies.

Session TB6 - Transient Functionals of Structured Markov Chains, Chair: Taylor in JVdC F

BRAVO for QED Queues

Yoni Nazarathy, The University of Queensland, Australia, y.nazarathy@uq.edu.au
Daryl J. Daley, The University of Melbourne, Australia, dndaley@gmail.com
Johan van Leeuwaarden, Eindhoven University of Technology, The Netherlands, j.s.h.v.leeuwaarden@tue.nl

The BRAVO effect stands for "Balancing Reduces Asymptotic Variance of Outputs". It is a stochastic queueing systems phenomenon where the long term asymptotic variability of the number of departures is significantly reduced in critically loaded (balanced) cases in comparison to non-critically loaded cases. For example, in the single server finite capacity M/M/1/K queue, the asymptotic variability is around 2/3 when the arrival and service rates are equal, yet is around 1 (as a Poisson process) otherwise. Formally the asymptotic variability is the limit of the ratio of the variance of the number of departures and the mean number of departures.

In this talk we consider the BRAVO effect for multi-server queues and show that it occurs under the quality and efficiency driven (QED) scaling regime. The QED scaling regime based on the many-server asymptotics of Halfin and Whitt (1981) is a well studied asymptotic analysis regime in which the number of servers and the load are increased in a manner such that the delay probability goes to a constant. We consider finite multi-server Markovian queues under a form of QED scaling and obtain a BRAVO effect. Hence BRAVO for QED.

The role of the deviation matrix in asymptotic functionals of Markov chains

Sophie Hautphenne, The University of Melbourne, Australia, sophiemh@unimelb.edu.au
Guy Latouche, Université libre de Bruxelles, Belgium, latouche@ulb.ac.be
Peter Taylor, The University of Melbourne, Australia, taylorpg@unimelb.edu.au

In this talk, we look at the deviation matrix of continuous-time Markov processes, and we discuss the role played by this matrix in asymptotic functionals of these processes. In particular, we use properties of the deviation matrix to revisit the BRAVO effect occurring in the asymptotic variance of the output process of the single-server finite capacity M/M/1/K queue (see Yoni Nazarathy's talk), and to obtain explicit expressions for the second order approximation of the variance curve. In light of the deviation matrix, we also analyse asymptotic functionals of other queueing systems such as the multi-server finite capacity M/M/K/K queue (also called the Erlang loss system).

A time-dependent study of birth-death processes, via the knockout queue

B. Fralix, Clemson University, Clemson, SC, USA, bfralix@clemson.edu

We present a new approach towards studying the time-dependent behavior of birth-death processes, whose birth rates and death rates are nonincreasing and nondecreasing, respectively, with respect to the state variable. This approach involves interpreting such a birth-death process as the queue-length process of what we refer to as the 'knockout queue', which is a queueing system where new arrivals are allowed to 'knockout', or eliminate customers upon arrival. Applications to both multiserver and finite-capacity queues will be presented.

Session TB7 - Control of Queues, Chair: Down in JVdC G

Optimal Admission Control for Tandem Loss Systems

D. F. Silva, School of ISyE, Georgia Institute of Technology, Atlanta, GA, 30332, dfsi3@gatech.edu
B. Zhang, IBM T. J. Watson Research Center, 1101 Kitchawan Road, Route 134, Yorktown Heights, NY 10598, bozhang@gatech.edu
H. Ayhan, School of ISyE, Georgia Institute of Technology, Atlanta, GA, 30332, hayhan@isye.gatech.edu

We study a system of two queues in tandem with finite buffers, Poisson arrivals to the first station, and exponentially distributed service times at both stations. Losses are incurred either when a customers is rejected at the time of arrival to the first station or when the second station is full at the time of service completion at the first station. The objective is to determine the optimal admission control policy that minimizes the long-run average cost.

A few new views on queues

Jingui Xie, Department of Industrial Engineering, Tsinghua University, Beijing, China, xiejingui@tsinghua.org.cn
Gabriel Zayas-Cabán, Center for Applied Mathematics, Cornell University, Ithaca, NY, USA, gz63@cornell.edu
Linda V. Green, Graduate School of Business, Columbia University, New York, NY, USA, lvg1@columbia.edu
Mark E. Lewis, School of Operations Research and Information Engineering, Cornell University, Ithaca, NY, USA, mark.lewis@cornell.edu

In many healthcare systems, the care of patients consists of two phases of service: assessment and treatment. Often these are carried out by the same medical provider and so there is a question as to how to prioritize the work in order to balance initial delays for care with the need to discharge patients in a timely fashion. We model a hospital emergency room (ER) triage and treatment process as a tandem queue with a single server. We explore alternative service disciplines under various scenarios and identify optimal policies for each.

Optimal rate for a queueing system in heavy traffic with superimposed On-Off arrivals.

Arka P. Ghosh, Iowa State University, Ames, IA, USA, apghosh@iastate.edu

We consider a control problem for a queueing system with heavy-tailed On-Off process arrivals and constant-rate service (control).The control problem is to find the optimal value of the service rate which minimizes an infinite horizon discounted cost function. The main result of the paper guarantees the existence of an optimal rate as well as specifies an explicit range of possible values of this optimal rate. As a part of the analysis, we also formulated and solved an approximating control problem driven by fractional Brownian motion. A key ingredient of the proof (and a result of independent interest) is an asymptotic maximal bound on the second moment of the centered cumulative On-Off process, which is also derived.

Dynamic Scheduling of a $GI/GI/1+GI$ Queue with Multiple Customer Classes

A. R. Ward, University of Southern California, USA, amyward@marshall.usc.edu
J. Kim, University of Southern California, USA, Jeunghyun.Kim.2015@marshall.usc.edu

We consider a dynamic control problem for a $GI/GI/1+GI$ queue with multiclass customers. The customer classes are distinguished by their inter-arrival time, service time, and abandonment time distributions. There is a cost $c_k > 0$ for every class k customer that abandons the queue before receiving service. The objective is to minimize average cost by dynamically choosing which customer class the server should next serve each time the server becomes available (and there are waiting customers from at least two classes).

It is not possible to solve this control problem exactly, and so we formulate an approximating Brownian control problem. The Brownian control problem incorporates the entire abandonment distribution of each customer class. We solve the Brownian control problem under the assumption that the abandonment distribution for each customer class has an increasing failure rate. We then interpret the solution to the Brownian control problem as a control for the original dynamic scheduling problem. Finally, we perform a simulation study to demonstrate the effectiveness of our proposed control.

Session TB8 - Stochastic Models, Chair: Jonckheere in JVdC H

Spatial stochastic models of heterogeneous cellular networks with repulsively deployed base stations

miyoshi@is.titech.ac.jp
I. Nakata, Tokyo Institute of Technology, Japan, nakata.i.aa@m.titech.ac.jp

We consider spatial stochastic models of downlink heterogeneous cellular networks (HCNs) with multiple tiers, where the base stations of each tier have a particular spatial density, transmit power, path-loss exponent and bias towards admitting mobile users. Existing works on such spatial models of HCNs assume, due to the tractability, that the base stations are deployed according to homogeneous Poisson point processes (PPPs). This means that the base stations are deployed independently with each other and their spatial correlation is ignored. In this talk, we propose two spatial models for the analysis of downlink HCNs, in both of which the base stations are deployed according to α-Ginibre point processes (α-GPPs). The α-GPP is one of the determinantal point processes and accounts for the repulsion between the base stations. Besides, the degree of repulsion can be adjusted according to the value of $\alpha \in (0, 1]$ and it is known that the α-GPP converges in law to a homogeneous PPP as $\alpha \to 0$. For such proposed models, we derive computable integral representations for the coverage probability of a typical mobile user—the probability that the downlink signal-to-interference-plus-noise ratio (SINR) for the typical user achieves a target threshold. We also exhibit the results of some numerical experiments.

Asymptotic Results for the First and Second Moments of Discrete-Time Bulk-Renewal Process

J.J. Kim, Royal Military College of Canada, Canada, s25412@rmc.ca
M.L. Chaudhry, Royal Military College of Canada, Canada, chaudhry-ml@rmc.ca

A simple and elegant solution to determine the asymptotic results for the renewal density as well as for the first and second moments of the number of renewals for the discrete-time bulk-renewal process is presented. The method of generating function is used to find the constant term in the second moment. In classic texts such as Feller (1968) and Hunter (1983), the constant term is missing. A recent paper by Van der Weide et al. (2007) states that it is not clear how to get the constant term using generating functions and as such they present this result using a different approach. Recently, Chaudhry and Fisher (2011) have responded to this problem by providing the asymptotic results for the renewal density as well as for both the first and second moments using generating functions. The purpose of this

note is to extend their results to bulk-renewal process in discrete-time and give an elegant derivation of the asymptotic results.

Queues and risk models with simultaneous arrivals

E.S. Badila, Eindhoven University of Technology, The Netherlands, e.s.badila@tue.nl
O.J. Boxma, Eindhoven University of Technology, boxma@win.tue.nl
J.A.C. Resing, Eindhoven University of Technology resing@win.tue.nl
E.M.M. Winands, University of Amsterdam, e.m.m.winands@uva.nl

We focus on a particular connection between queueing and risk models in a multi-dimensional setting. We first consider the joint workload process in a queueing model with parallel queues and simultaneous arrivals at the queues. For the case that the service times are ordered (from largest in the first queue to smallest in the last queue) we obtain the Laplace-Stieltjes transform of the joint stationary workload distribution. Using a multivariate duality argument between queueing and risk models, this also gives the Laplace transform of the survival probability of all books in a multivariate risk model with simultaneous claim arrivals and the same ordering between claim sizes. (Joint work with Onno Boxma, Jacques Resing and Erik Winands).

Particle systems and quasi-stationary distributions

P. Groisman, University of Buenos Aires, Argentina, pgroisma@dm.uba.ar
M. Jonckheere, Conicet, Universidad of Buenos Aires, Argentina, mjonckhe@dm.uba.ar

Nothing lasts forever. However many phenomena can be well described by a process which enters a quasi-stationary state before eventually vanishing. Since the pioneering work of Kolmogorov and Yaglom, a lot of work has been dedicated to understand the quasi-stationary behavior of Markov processes through the invariant distributions (Quasi-Stationary distributions) of the conditioned evolution. Unlike invariant distributions, QSD are solutions of a non-linear equation and there can be 0, 1 or an infinity of them. Also, they cannot be obtained as Cesaro limits of Markovian dynamics. These facts make the computation of QSDs a nontrivial matter.

We study different particles systems (Branching particles, Branching with selection, Fleming Viot systems) allowing to simulate QSD distributions. We also derive and explain some links with the existence of traveling waves for some specific PDEs.

Session TB9 - Two-dimensional Markov Processes: Fluid Queues and Random Walks, Chair: Nguyen in Cabildo I

Markov modulated two node fluid network: Tail asymptotics of the stationary distribution

M. Miyazawa, Tokyo University of Science, Japan, miyazawa@is.noda.tus.ac.jp

We consider a Markov modulated two node fluid network with unlimited buffers. Exogenous fluid input and potential release rates at each node are modulated by a finite state Markov chain. A constant fraction of the output from one node is transferred to the other, and these fractions are assumed to be not modulated. Thus, this model is a network generalization of standard Markov modulated single and tandem fluid queues. We describe it by a continuous time Markov modulated two dimensional reflecting process on the quadrant. Similar to a semi-martingale reflecting Brownian motion, regulators play a key role to handle reflection here.

We are interested in tail asymptotic behaviors of the stationary distribution of this continuous time reflecting process, provided it is stable. For this, we derive the convergence domain of the moment generating function of the stationary distribution in terms of the modeling primitives. This gives an upper bound for the tail decay rate of the marginal distribution in an arbitrary direction. This bound is obtained as the solution of an optimization problem for a linear

objective function with a finite number of linear and quadratic constrains. We conjecture it to be the decay rate. Some technical evidence is provided for this conjecture.

Tail asymptotics of the stationary distribution of a two dimensional reflecting random walk with unbounded upward jumps

Masahiro Kobayashi, Tokyo University of Science, Japan, m_kobayashi@is.noda.tus.ac.jp
Masakiyo Miyazawa, Tokyo University of Science, Japan, miyazawa@is.noda.tus.ac.jp

We consider a two dimensional reflecting random walk on the nonnegative integer quadrant. This random walk is assumed to be skip free in the direction to the boundary of the quadrant, but may have unbounded jumps in the opposite direction, which are referred to as upward jumps. We are interested in the tail asymptotic behavior of its stationary distribution, provided it exists. Assuming the upward jump size distributions have light tails, we completely find the rough tail asymptotics of the marginal stationary distributions in all directions. This generalizes the corresponding results for the skip free reflecting random walk. We exemplify these results for a two node network with exogenous batch arrivals.

Asymptotic independence of (simple) two-dimensional Markov processes

G. Latouche, Université libre de Bruxelles, Belgium, guy.latouche@ulb.ac.be

We consider a two-dimensional Markov process $\{(X_1(t), X_2(t))\}$ on the nonnegative integers, which we would refer to as a 2-d reflecting random walk with homogeneous transitions: X_1 and X_2 may change by at most one unit at each transition, and the jump probabilities are independent of X_1 and X_2 if both are different from zero, with similar assumption if X_1 or X_2 is equal to zero. The process is assumed to be irreducible and positive recurrent, with stationary distribution $\pi_{n_1,n_2} = \lim_{t\to\infty} \mathrm{P}[X_1(t) = n_1, X_2(t) = n_2]$.

The stationary distribution is said to have *product form* if it is factored as $\pi_{n_1,n_2} = \alpha n_1, \beta n_2$, where α and β are two probability densities. In a first step, we investigate the conditions under which it holds. This is of theoretical interest of course but, in a second step, we show how to use the knowledge to find product form approximations for otherwise unmanageable random walks.

Tuesday 2:30pm - 4:00pm

Session TC1 - Optimization I, Chair: Kalathil in JVdC A

An optimal dynamic funding solution to prevent default costs

U. Ayesta, CNRS-LAAS, Toulouse, France,
M. Erausquin, UPV/EHU, University of the Basque Country, Bilbao, Spain
E. Ferreira, UPV/EHU, University of the Basque Country, Bilbao, Spain
P. Jacko, Basque Center for Applied Mathematics, Bilbao, Spain

There are many situations where either public or private resources might be used to obtain social rewards. How to manage those resources in an optimal manner is a problem widely analysed in the literature. Depending on the available controls, and on the social reward that is wanted to be maximized, different optimal control models have been presented. For instance, Arrow (1970) develops models where control instruments such as income tax rate, government deficit size, and debt-money ratio are controlled in order to maximize functions defined in terms of some social criteria (usually depending on growth rate).

In the last years, and more intensely due to economical and financial situation, the institutions have been facing a new resource allocation problem, which is to minimize the social cost incurred by the economy when a company defaults.

In other words, if there is a measure of the social and/or economic interest for a firm to survive (e.g. sudden increase in unemployment due to collective redundancies), it is interesting for governments and institutions to have tools to make funding decisions considering the dynamics of the firms.

However, to the best of our knowledge, there is no literature describing dynamic mathematical models for this resource allocation problem. This paper represents a first attempt to develop a mathematical method that provides an optimal resource allocation policy, with the goal of minimizing the lomg term social costs incurred by the economy due to default events.

We consider a resource allocation problem, where some institution decides how to share fixed resources to prevent default events among companies. Under a dynamic approach, the problem is solved in the framework of Multi Armed Restless Bandit problems.

A two-state Markovian process determines the evolution of the default risk, and the objective for the institution is to minimize the total cost incurred in the long term due to default events. We prove the existence of an optimal policy, that assigns an index value to each company identifying its priority for resources. The analytical expression for the index is derived, which generalizes the return-on-investment (ROI) index. Finally, a discussion and intepretation of the structure of the optimal policy is presented and some proposals for future research.

Main Bibliografy

Arrow, Kurz, 1970. Public Investment, The Rate of Return, and Optimal Fiscal Policy. Johns Hopkins University Press.

Gittins, J.C., Glazebrook, K., Weber, R., 2011. Multi-Armed Bandit Allocation Indices. Wiley-Blackwell

Nino-Mora, J., 2007. Dynamic priority allocation via restless bandi marginal productivity indices. TOP 15, 161-198.

Minimizing risk measures in bandit problems

J. Y. Yu, IBM Research–Ireland, Damastown, Dublin 15, Ireland, jiayuanyu@ie.ibm.com

Bandit problems model sequential decision problems where not only are the outcomes random, but their distributions are unknown. Most of the existing work on bandit problems assumes a risk-neutral agent, whose objective is maximizing the expected reward. However, many applications have risk-averse objectives, such as minimizing the probability of failure. These problems can be modeled as minimizing one of various risk measures associated with a decision policy. In this work, we consider widely used risk measures, such as the value-at-risk and the expected shortfall, and present an efficient decision policy that estimates the risk measures from the empirical distribution of observed samples. We show a performance guarantee for this policy and apply it to an active learning problem where we selectively seek feedback from human experts.

Decentralized Learning for Multi-player Multi-armed Bandits: An Algorithm for Near-Logarithmic Regret

Dileep Kalathil, Naumaan Nayyar, Rahul Jain
University of Southern California, Los Angeles, (manisser, nnayyar, rahul.jain)@usc.edu

Multi-Armed Bandits (MAB) are a classical model for learning and optimization by a single agent in an unknown environment . In this work we consider the problem of decentralized online learning with multiple players in an unknown environment which we formulate as a Multi-player Multi-armed Bandits model. In this model each player can pick among multiple arms and when a player picks an arm, it gets a random reward. We consider both i.i.d. reward model and Markovian reward model. In the i.i.d. reward model, the reward from each arm is modelled as an i.i.d. process with an unknown distribution with an unknown mean. In the Markovian reward model, each arm is modelled as a Markov chain with an unknown probability transition matrix with an unknown stationary distribution. The players have a joint objective to minimize the cost of learning called *regret* and all players should learn jointly to discover the best arms to play as a team. Since they are all trying to learn at the same time, they may collide when two or more players pick the same arm and neither of them get any reward. There is no dedicated control channel

for coordination or communication among the players. Any other communication between the players is costly and will increase the regret. We propose a decentralized online learning policy called distributed Phased Exploration and Exploitation (dPEE) algorithm that achieves an expected regret that grows at most as *near-*$O(\log T)$. The motivation comes from opportunistic spectrum access by multiple secondary users in cognitive radio networks wherein they must pick among various wireless channels that look different to different users.

Session TC2 - Stochastic systems in applications, Chair: Lu in JVdC B

Risk-Aware Revenue Maximization in Display Advertising

W. D. Heavlin, Google Research, CA, USA, bheavlin@google.com
A. Radovanović, Google Research, NY, USA, anaradovanovic@google.com

Display advertising is the graphical advertising on the World Wide Web (WWW) that appears next to content on web pages, instant messaging (IM) applications, email, etc. Over the past decade, display ads have evolved from simple banner and pop-up ads to include various combinations of text, images, audio, video, and animations. As a market segment, display continues to show substantial growth potential, as evidenced by companies such as Microsoft, Yahoo, and Google actively vying for market share. As a sales process, display ads are typically sold in packages, the result of negotiations between sales and advertising agents.

A key component to any successful business model in display advertising is sound pricing. Main objectives for online publishers (e.g. Amazon, YouTube, CNN) are maximizing revenue while managing their available inventory appropriately, and pricing must reflect these considerations.

This paper addresses the problem of maximizing revenue by adjusting prices of display inventory. We cast this as an inventory allocation problem. Our formal objective (a) maximizes expected revenue using (b) iterative price adjustments in the direction of the gradient of an appropriately constructed Lagrangian relaxation. We show that our optimization approach drives the revenue towards local maximum under *mild* conditions on the properties of the (unknown) demand curve.

The major unknown for optimizing revenue in display environment is how the demand for display ads changes to prices, the classical demand curve. This we address directly, by way of a factorial pricing experiment. This enables us to estimate the gradient of the revenue function with respect to inventory prices. Overall, the result is a principled, risk-aware, and empirically efficient methodology.

Performance Analysis and Scheduling in Big Data Systems

Li Zhang, IBM Thomas J. Watson Research Center, New York, USA, zhangli@us.ibm.com
Jian Tan, IBM Thomas J. Watson Research Center, New York, USA, tanji@us.ibm.com
Xiaoqiao Meng, IBM Thomas J. Watson Research Center, New York, USA, xmeng@us.ibm.com

More and more analytic applications are built in Big Data systems providing important insights from analyzing large amount of data. Understanding the performance issues in these large distributed systems is important for providing efficient resource utilization and satisfactory performance guarantees. Due to the large number of jobs sharing limited resources, an efficient job scheduler is critical. We present analytical and experimental studies of existing job schedulers. We also introduce our own job scheduler with superior performance.

Structural Properties and Heuristic Optimal Policies for Serials Lines with Flexible Workers

Yingdong Lu, IBM T.J. Watson Research Center, Yorktown Heights, NY 10598

We consider a system with a stream of jobs that consist of N tasks that have to be processed in order on N different serving stations. Flexible workers can be assigned to one or several stations. The interarrival times and task distributions are assume to follow general probability distributions. Using stochastic ordering techniques, we obtain structural

properties for optimal policies for various objectives, such as maximize throughput and minimize average sojourn time. These structural properties enable us to propose heuristic policies, and establish performance guarantees for these policies.

Optimal Resource Capacity Management in Stochastic Networks

M.S. Squillante, IBM Thomas J. Watson Research Center, New York, USA, mss@us.ibm.com
A.B. Dieker, Georgia Institute of Technology, Georgia, USA, ton.dieker@isye.gatech.edu
S. Ghosh, IBM Thomas J. Watson Research Center, New York, USA, ghoshs@us.ibm.com

We develop a general framework for determining the optimal resource capacity for each station comprising a stochastic network, motivated by applications arising in computer capacity planning and business process management. The problem is mathematically intractable in general and therefore one typically resorts to either overly simplistic analytical approximations or very time-consuming simulations in conjunction with metaheuristics. In this talk we propose an iterative methodology that relies only on the capability of observing the queue lengths at all network stations for a given resource capacity allocation. We theoretically investigate the proposed methodology for single-class Brownian tree networks, and further illustrate the use our methodology and the quality of its results through extensive numerical experiments.

Session TC3 - Simulation and Estimation, Chair: Lam in JVdC C

Iterative Methods for Robust Estimation under Bivariate Distributional Uncertainty

S. Ghosh, T. J. Watson IBM Research Center, Yorktown Heights, NY 10598, USA, ghoshs@us.ibm.com
H. Lam, Boston University, Boston, MA 02215, khlam@bu.edu

We propose an iterative algorithm to approximate the solution to an optimization problem that arises in estimating the value of a performance metric in a distributionally robust manner. The optimization formulation seeks to find a bivariate distribution that provides the worst-case estimate within a specified statistical distance from a nominal distribution and satisfies certain independence condition. This formulation is in general non-convex and no closed-form solution is known. We use recent results that characterize the local "sensitivity" of the estimation to the distribution used, and propose an iterative procedure on the space of probability distributions. We establish that the iterations of solutions are always feasible and that the sequence is provably improving the estimate. We describe conditions under which this sequence can be shown to converge to a locally optimal solution. Numerical experiments illustrate the effectiveness of this approach for a variety of nominal distributions.

Session TC4 - Markov modulation in risk and queueing processes, Chair: Palmowski / D'Auria in JVdC D

The Sequential Probability Ratio Test revisited: linking statistics and ruin theory

Hansjörg Albrecher, University of Lausanne, Switzerland, hansjoerg.albrecher@unil.ch
P. Asadi, University of Lausanne, Switzerland, peiman.asadi@unil.ch
J. Ivanovs, University of Lausanne, Switzerland, jevgenijs.ivanovs@unil.ch

Wald's classical sequential probability ratio test for one simple hypothesis against an alternative is based on the boundary crossing of an associated random walk. In this talk we connect this test to a problem in ruin theory, leading to explicit expressions for the decision boundaries for sequential testing of Erlang distributions. Information on the mean sample size of the test can be retrieved as well. The approach relies on the use of scale matrices associated to corresponding Markov additive processes. This simplifies and extends earlier results of Teugels & Van Assche.

Risk model with an observer in Markov environment

H. Albrecher, University of Lausanne, Switzerland, hansjoerg.albrecher@unil.ch
Jevgenijs Ivanovs, University of Lausanne, Switzerland, jevgenijs.ivanovs@unil.ch

We consider a spectrally-negative Markov additive process as a model of a risk process in random environment. Following recent interest in alternative ruin concepts, we assume that the ruin occurs when an independent Poissonian observer sees the process negative, where the rate of observations may depend on the state of the environment. This model generalizes Parisian Lévy risk models with exponential implementation delay. Mathematically, we are interested in the joint Laplace transform of the occupation times in the negative half plane corresponding to different states of the environment. We extend the quantity of interest by adding essential structure to it, which allows to employ an approximation method. Surprisingly, the resulting formulas are rather neat.

Approaching Markov-modulated Brownian motions via matrix-analytic methods

Giang Nguyen, The University of Adelaide, Australia, giang.nguyen@adelaide.edu.au

We obtain Markov-modulated Brownian motions (MMBM) as the limit of a family of Markov-modulated linear fluid processes. This weak convergence enables us to apply matrix-analytic methods to Markov-modulated Brownian motions with boundaries. In particular, we revisit the stationary distribution of reflected Markov-modulated Brownian motions, and discuss a direct formula for the stationary distribution of two-sided MMBM.

Heavy-traffic asymptotics for networks of parallel queues with Markov-modulated service speeds

J.L. Dorsman, Eindhoven University of Technology and Centrum Wiskunde & Informatica, The Netherlands, j.l.dorsman@tue.nl
M. Vlasiou, Eindhoven University of Technology and Centrum Wiskunde & Informatica, The Netherlands, m.vlasiou@tue.nl
B. Zwart, Centrum Wiskunde & Informatica, VU University Amsterdam, Eindhoven University of Technology, The Netherlands, and Georgia Institute of Technology, USA, Bert.Zwart@cwi.nl

Motivated by a number of real-life applications, we study queueing networks consisting of parallel single-server queues. The distinguishing feature of these networks is that the service speeds vary over time and are in addition mutually dependent. More specifically, we assume that the service speeds are modulated by a single continuous-time Markov chain. We obtain heavy-traffic limits for the joint workload, waiting-time and queue-length processes. We do so by using a functional central limit theorem approach, which requires the interchange of steady-state and heavy-traffic limits. The marginals of these limiting distributions are shown to be exponential with rates that can be computed by matrix-analytic methods. Moreover, we show how to numerically compute the joint distribution, by viewing the limit processes as multi-dimensional semi-martingale reflected Brownian motions in the non-negative orthant.

Session TC5 - Spatial Stochastic Processes and their Applications, Chair: Ramanan in JVdC E

A spatial model of cancer initiation: basic results

K.Z. Leder, University of Minnesota, USA, lede0024@umn.edu
R. Durrett, Duke University, USA, rtd@math.duke.edu
J. Foo, University of Minnesota, USA, jyfoo@math.umn.edu

The issue of cancer initiation has been studied from a mathematical perspective in several works, however the vast majority of these works have focused on homogeneously mixing populations. In order to understand cancer initiation in a solidly structured tissue it is necessary to introduce a spatial model of this process. In this talk I will describe a spatial Moran model with selection and mutation that is used to model the cancer initiation process. Specifically I will discuss some of the basic properties of the model that are necessary to understand how and when a cancer initiation event occurs. In a second talk Jasmine Foo will discuss further results of this model. This is based on joint work with Rick Durrett and Jasmine Foo.

A spatial model of cancer initiation: approximations and applications

J. Foo, University of Minnesota, USA, jyfoo@umn.edu

In this talk I will continue the discussion of Kevin Leder on the the problem of cancer initiation. Specifically I will introduce an approximation to the spatial Moran process based on the Bramson-Griffeath shape theorem that can be used to determine the asymptotic properties of the time until cancer initiation events occur. This approximation will identify three distinct regimes by which cancer initiation can occur, each regime producing distinct spatial behaviors and initiation times. I will conclude by using these asymptotic results to discuss some estimates of spatial heterogeneity in tumors in solid tissue. Joint work with K. Leder, R. Durrett, and M. Ryser.

Limit theorems for Smoluchowski dynamics associated with critical continuous-state branching processes

G. Iyer, Carnegie Mellon University, USA, gautam@math.cmu.edu
N. Leger, Carnegie Mellon University, USA, nleger@math.cmu.edu
R. Pego, Carnegie Mellon University, USA, rpego@math.cmu.edu

We will exploit a well known connection (Bertoin, Le Gall '06) between critical continuous-state branching processes (CSBP) that are absorbed at 0 and the generalized Smoluchowski coagulation equation to prove a variety of limit theorems. We establish a fundamental connection between the existence of a non-degenerate scaling limit and regular variation (at 0) of the branching mechanism. In particular, we show:

- A non-degenerate scaling limit of a critical CSBP (absorbed at 0) exists if and only if the branching mechanism is regularly varying at 0.
- The CDF of the rescaled process conditioned on survival has a nondegenerate limit if and only if the branching mechanism is regularly varying at 0.

When the scaling limits exist, we can further characterise the nondegenerate scaling limits of arbitrary finite-measure solutions in terms of generalized Mittag-Leffler series.

Conditional phase transitions: some simple examples

P. Rebeschini, Princeton University, USA
R. Van Handel, Princeton University, USA, rvan@princeton.edu

We are interested in the emergence of phase transitions in random fields when we condition on a part of the model. Simple examples demonstrate that such phase transitions can arise conditionally, even when no phase transition exists prior to conditioning. A general understanding of such phenomena is still lacking. Conditional phase transitions are closely related to the development of Monte Carlo algorithms for data assimilation that can surmount the curse of dimensionality, as well as to some classic measure-theoretic problems.

Session TC6 - Funding Opportunities in OR at NSF, Chair: Jacobson in JVdC F

Funding Opportunities in Operations Research at the National Science Foundation

Sheldon H. Jacobson, National Science Foundation, shj@nsf.gov

This presentation discusses opportunities for funding in Operations Research at the National Science Foundation. It also discusses recent changes implemented for proposals submitted to NSF, as well as new issues related to proposal review that would be of interest to anyone interested in seeking funding for their research through the National Science Foundation.

Session TC7 - Stochastic control, Chair: Gupta / Keskin in JVdC G

A study of the impact of long-range dependence on a class of threshold-type stochastic control policies

M. Sharma, IBM T. J. Watson Research Center, Yorktown Heights, NY, mxsharma@us.ibm.com
J.W. Bosman, CWI Amsterdam, The Netherlands, Joost.Bosman@cwi.nl
X. Gao, Georgia Tech, Atlanta, GA, gxf1240@gatech.edu
Y. Lu, IBM T. J. Watson Research Center, Yorktown Heights, NY, yingdong@us.ibm.com
M. S. Squillante, IBM T. J. Watson Research Center, Yorktown Heights, NY, mss@us.ibm.com

Recently, the authors considered a class of general dynamic resource management problems where the goal was to determine the allocation level for each resource type in order to serve the uncertain demand in a way so as to maximize the expected profit over a time horizon of interest. They derived the optimal control policy within a singular stochastic optimal control setting and provided simple expressions for the dynamic threshold levels for each resource over time. In this presentation, we study the performance of the proposed policy in an alternate setting where the demand is assumed to be drawn from the set of stable distributions with Hurst parameter H. While it is clear that the proposed policy is no longer optimal in this setting, we are able to experimentally confirm that the policy continues to exhibit excellent properties in tracking the demand, and outperforms an interval-average based offline policy which has full information on demand. Further, the gap to optimality appears to have an interesting functional dependence on the Hurst parameter. We present some experimental and theoretical results to support and explain our findings.

Approximate Dynamic Programming Approach to Stochastic Matching

N.P. Bhat, Graduate School of Business, Columbia University, nbhat15@gsb.columbia.edu
C.C. Moallemi, Graduate School of Business, Columbia University, ciamac@gsb.columbia.edu

We consider a class of stochastic control problems where the action space at each time can be described by a class of matching or, more generally, network flow polytopes. Special cases of this class of dynamic matching problems include many problems that are well-studied in the literature, such as: (i) online keyword matching in Internet advertising (the 'adwords' problem); (ii) the bipartite matching of donated kidneys from cadavers to recipients; and (iii) the allocation of donated kidneys through exchanges over cycles of live donor-patient pairs. We provide an approximate dynamic program (ADP) algorithm for dynamic matching with stochastic arrivals and departures. Our framework is more general than the methods prevalent in the literature in that it is applicable to a broad range of problems characterized by a variety of action polytopes and generic arrival and departure processes. In order to access the performance of our ADP methods, we illustrate computationally tractable upper bounds on the performance of optimal policies in our setting. We apply our methodology to a series of kidney matching problems calibrated to realistic kidney exchange statistics, where we obtain a significant performance improvement over established benchmarks and, via upper bounds, illustrate that our approach is near optimal.

Higher order Markov random fields for independent sets

David A. Goldberg, Georgia Institute of Technology, USA, dgoldberg9@isye.gatech.edu

The independent sets of a network, i.e. sets of nodes with no internal edges, arise in the optimization of stochastic networks when agents must simultaneously utilize a scarce resource. We consider higher order Markov random fields to study independent sets in large networks with no small cycles. We give sufficient conditions for a second-order homogenous isotropic Markov random field to exhibit long-range boundary independence (i.e. decay of correlations, unique infinite-volume Gibbs measure), and give both necessary and sufficient conditions when the relevant clique potentials of the corresponding Gibbs measure satisfy a log-convexity assumption. We gain further insight into this characterization by interpreting our model as a multi-dimensional perturbation of the hardcore model, and (under a convexity assumption) give a simple polyhedral characterization for those perturbations (around the well-studied critical activity of the hardcore model) which maintain long-range boundary independence. We also characterize (again as a polyhedral set) how one can change the occupancy probabilities through such a perturbation. We then use linear programming to analyze this set of attainable probabilities, showing that although one cannot acheive denser independent sets, it is possible to optimize the number of excluded nodes adjacent to no included nodes.

A Brownian Model of Dynamic Pricing with Demand Model Uncertainty

N. Bora Keskin, The University of Chicago Booth School of Business, Chicago, IL, bora.keskin@chicagobooth.edu

We propose a Brownian formulation to address the trade-off between learning and earning in a dynamic pricing problem that involves demand model uncertainty. Using stochastic control theory, we solve for the optimal policy of the Brownian model. We discuss how this optimal policy can be employed to design dynamic pricing policies in discrete time formulations, and demonstrate the performance of such policies via simulation. This approach provides general guidelines to implement successful price experimentation in practice.

Session TC8 - Novel directions in queueing, Chair: Kapodistria in JVdC H

Service systems with skill based routing, under FCFS policies

I.J.B.F. Adan, Eindhoven University of Technology, The Netherlands, i.j.b.f.adan@tue.nl
M.A.A. Boon, Eindhoven University of Technology, The Netherlands, m.a.a.boon@tue.nl
G. Weiss,University of Haifa, Israel, gweiss@stat.haifa.ac.il

We study an overloaded service system with servers of types $\mathcal{S} = \{s_1, \ldots, s_J\}$, serving customers of types $\mathcal{C} = \{c_1, \ldots, c_I\}$ under FCFS. Customers arrive in Poisson streams, join the queue and then abandon or get served. Service is skill based, which is described by a graph G, where $(i, j) \in G$ if server type s_j is trained to serve customer type c_i. The service duration depends on both server and customer type. At this level of generality, the design in terms of staffing and cross-training decisions is a challenging problem. Based on recent results on infinite matching and on some asymptotic assumptions, we propose an algorithm to determine, for given data, the required levels of staffing to meet target levels of service quality and labor division. The algorithm is validated through a systematic simulation study, showing that it is remarkably robust and accurate. As such, we believe that the algorithm will prove to be useful in aiding the design and effective operation of complex systems with skill based routing.

Using Estimated Patience Levels to Optimally Schedule Customers

A. Bassamboo, Kellogg School, Northwestern University, IL, USA, a-bassamboo@kellogg.northwestern.edu
R.S. Randhawa, Marshall School, University of South California, LA, USA, rrandhaw@marshall.usc.edu

Customer impatience has become an integral component of analyzing queueing systems in extant literature, especially in the context of call centers. A common model of impatience is to assume that customers are endowed with a "patience

clock" and that they abandon when this clock runs out and they are still waiting for service. Further, the heterogeneity in the customer's patience clocks are represented using a common distribution, the impatience or abandonment distribution. That is, customers patience times are i.i.d. draws from this distribution. So, as customers wait in the queue, an update can be formed on their willingness to wait, or patience, as time progresses. So, even though all customers may be identical when they just join the queue, as time progresses, they become differentiated on their further willingness to wait. The main research question in this paper is: What is the optimal way to schedule customers given that as the customers wait in the queue, they reveal additional information about their willingness to wait?

An Aggregation Method For The Tandem Threshold Queue

N. Baer, University of Twente, The Netherlands, n.baer@utwente.nl
R.J. Boucherie, University of Twente, The Netherlands, r.j.boucherie@utwente.nl
A. Al Hanbali, University of Twente, The Netherlands, a.alhanbali@utwente.nl
J.C.W. van Ommeren, University of Twente, The Netherlands, j.c.w.vanommeren@utwente.nl

We consider a tandem system of 2 queues with finite buffer where service rates are controlled by a threshold policy. Customers arrive to the first queue according to a Poisson process with rate λ and require an exponential amount of service at each queue i, $i = 1, 2$. Each queue i has an upper threshold U_i, a lower threshold L_i, a high service rate μ_i and a low service rate ν_i. When the queue length in queue i reaches U_i, the service rate changes from μ_i to ν_i. The service rate switches back from ν_i to μ_i when the queue length decreases to L_i. We present an aggregation algorithm based on Matrix Analytic Methods allowing us to obtain the stationary queue length distribution. Furthermore, we show that even though the aggregation algorithm is designed for finite buffers, it is not necessary for the second queue to be finite.

Factorial approach for the study of the infinite server queue with synchronized reneging

S. Kapodistria, Eindhoven University of Technology, The Netherlands, s.kapodistria@tue.nl
J.A.C. Resing, Eindhoven University of Technology, The Netherlands, j.a.c.resing@tue.nl
T. Phung-Duc, Tokyo Institute of Technology, Japan, tuan@is.titech.ac.jp

The effect of reneging (or customers' abandonments) has received considerable attention from the first years of Queueing Theory. In the classical models, it is assumed that each customer sets his own patience time upon arrival to a system and he abandons it prematurely if this time expires before he has initiated (or completed) its service time. In recent papers, several authors have considered other kinds of reneging, among them the so-called synchronized abandonments that occur when the customers decide concurrently but independently, at given opportunities, whether they will abandon the system or not.
The tractability of models with abandonments is limited, even in the Markovian framework. There are several studies for models with the usual (independent) abandonments and a few studies for models with synchronized abandonments that yield closed-form expressions for some stationary performance measures. We present a new methodological approach for the study of models combining both linear (due to infinite servers) and binomial transitions (arising form synchronized abandonments). We show the path for the stationary and transient analysis of these model and we discuss extensions and generalizations.

Wednesday 10:00am - 11:30am

Session WA1 - Urn models, Chair: Maulik in JVdC A

Pólya-Eggenberger-Friedman Urn Models: A New Approach

Antar Bandyopadhyay, Indian Statistical Institute, New Delhi and Kolkata, India, antar@isid.ac.in
Debleena Thacker, Indian Statistical Institute, New Delhi, India, thackerdebleena@gmail.com

In this talk we will consider the classical and other generalizations including infinite color generalization of the Pólya-Eggenberger-Friedman Urn Models and show that in the *balanced* case the configuration of the urn after n steps can be obtained by sampling the under lying Markov chian at random time which depends on n. We will show that most of the existing results may be derived using this representation. Moreover new results will be derived particularly for the infinite color case.

On a preferential attachment and generalized Polya's urn model

Andrea Collevecchio, Università Cá Foscari Venezia, Italy, collevec@unive.it
Codina Cotar, University College London, United Kingdom, c.cotar@ucl.ac.uk
Marco LiCalzi, Università Cá Foscari Venezia, Italy, gourab@math.ubc.ca

We study a general preferential attachment and Polya's urn model. At each step a new vertex is introduced, which can be connected to at most one existing vertex. If it is disconnected, it becomes a pioneer vertex. Given that it is not disconnected, it joins an existing pioneer vertex with a probability proportional to a function of the degree of that vertex. This function is allowed to be vertex dependent, and is called reinforcement function. We prove that there can be at most three phases in this model, depending on the behavior of the reinforcement function. Consider the set whose elements are the vertices whose cardinality tends a.s. to infinity. We prove that this set either is empty, or it has exactly one element, or it contains all the pioneer vertices. Moreover, we describe the phase transition in the case where the reinforcement function is the same for all vertices. Our results are general, and in particular we are not assuming monotonicity of the reinforcement function. Finally, consider the regime where exactly one vertex has a degree diverging to infinity, and suppose that at a certain stage, a given vertex has a large degree compared to the others. We give a lower bound for the probability that this vertex will be the leading one, i.e. its degree will diverge to infinity. Our proofs rely on a generalization of the Rubin construction given for edge-reinforced random walks, and on a Brownian Motion embedding.

Rates of convergence of color count in balanced urn models

Amites Dasgupta, Indian Statistical Institute, Kolkata, India, amites@isical.ac.in
Krishanu Maulik, Indian Statistical Institute, Kolkata, India, krishanu@isical.ac.in
Gourab Ray, University of British Columbia, Vancouver, Canada, krishanu@isical.ac.in

We consider urn models containing balls of finitely many colors and with balanced replacement matrices, that is, replacement matrices with common row sum. Using martingale techniques, we study the rates of convergence of each color count in almost sure as well as L^p sense, for any $p > 0$. We also provide relationship between the limiting random variables corresponding to various colors thus obtained.

Session WA2 - Finance and Risk I, Chair: Iyer in JVdC B

Dynamic Pricing with Linear Price Correction

Stefanus Jasin, Stephen M. Ross School of Business, University of Michigan, Ann Arbor, sjasin@umich.edu
George Chen, Stephen M. Ross School of Business, University of Michigan, Ann Arbor, georgeqc@umich.edu
Izak Duenyas, Stephen M. Ross School of Business, University of Michigan, Ann Arbor, duenyas@umich.edu

We consider a monopolist dynamic pricing problem with multiple products, multiple resources, time-homogeneous demands, and a finite selling horizon. Two common concerns for dynamic pricing are the potentially expensive re-optimizations and the frequency of price adjustments. To address this, we introduce a pricing heuristic which only requires solving a large convex optimization problem *once* at the beginning of selling horizon and then automatically adjusts the prices over time. That is, apart from the initial optimization, our heuristic can be implemented in online fashion. Not only the inted heuristic reduces the burden of re-optimizations, it only requires that we dynamically adjust

the prices of a *small* subset of products, leaving the prices of the remaining products unchanged throughout the selling season. These are the products which form a "base" with respect to the initial optimization problem. We provide a theoretical bound on expected revenue loss which depends on the choice of the base and a general updating schedule of the products in the base. (We allow the price of different products in the base to be updated at different times.) Given time, we will also discuss some extensions to the setting of time-varying demands. In particular, we will address whether one can still adjust the prices of only a small subset of products in such setting.

Switch-when-safe cone-constrained mean-variance strategies

C. Labbé, HEC-Montreal, Montreal, Canada, chantal.labbe@hec.ca
F. Watier, Université du Québec à Montreal, Montreal , Canada, watier.francois@uqam.ca

Li and Zhou (2006) devised a *switch-when-safe* (SWS) financial strategy in which an investor follows an optimal mean-variance strategy up to a possible random time where the accumulated wealth is large enough so he can, at this point, safely reinvest all of this amount of money in a risk-free bank account in order to achieve his financial target at the end of the investment horizon. They established, for an unconstrained mean-variance portfolio in a continuous-time Black-Scholes market model, that there is at least an 80% probability that the investor will meet his goal with a SWS strategy. Surprisingly, we will show that for cone-constrained mean-variance strategies (which includes no-short selling restrictions) the 80% lower bound probability still holds. Furthermore, we will give an expression for the probability that the investor reaches his target before bankruptcy in terms of a “greediness" parameter.

Feedback Stackelberg Solutions of Infinite-Horizon Stochastic Differential Games

Alain Bensoussan, University of Texas at Dallas, USA, alain.bensoussan@utdallas.edu
Shaokuan Chen, University of Texas at Dallas, USA, Shaokuan.Chen@utdallas.edu
Suresh P. Sethi, University of Texas at Dallas, USA, sethi@utdallas.edu

We present a sufficient condition for a feedback Stackelberg equilibrium of a stochastic differential game on an infinite horizon. This condition gives rise to a system of elliptic partial differential equations involving a static Stackelberg game at the level of Lagrangian. As an example, we consider a linear quadratic problem, obtain the corresponding algebraic Riccati equation, and provide its solution in the scalar case.

Keywords: Differential games, feedback Stackelberg equilibrium, Riccati equation, infinite horizon.

Information and the Value of Execution Guarantees

Krishnamurthy Iyer, University of Pennsylvania, kriyer@seas.upenn.edu
Ramesh Johari, Stanford University, ramesh.johari@stanford.edu
Ciamac C. Moallemi, Columbia Business School, ciamac@gsb.columbia.edu

In many markets, uncertainty about whether a trade is executed can be removed by paying a price premium. We use financial markets as a particular setting in which to study this trade-off. In particular, we assess the role of information in the choice between certain trade at a price premium in an intermediated dealer market and contingent trade in a dark pool. Our setting consists of intrinsic traders and speculators, each endowed with heterogeneous fine-grained private information as to an asset’s value, that endogenously decide between these two venues. We look for an equilibrium in this setting, and address three main questions: First, we illustrate how the choice between certain and contingent trade depends on information available to an individual agent, as well as the overall distribution of information across all agents. Second, we analyze how the premium for certain trade over contingent trade affects the strategic behavior of traders. Finally, we investigate how the presence of the option for contingent trade affects the welfare of the market in the presence of competitive market makers.

Session WA3 - Algorithmic and policy-level applications of probability, Chair: Kharoufeh / Goldberg in JVdC C

Distributionally robust inventory control when demand is a martingale

David A. Goldberg, Georgia Institute of Technology, USA, dgoldberg9@isye.gatech.edu
Linwei Xin, Georgia Institute of Technology, USA, lwxin@gatech.edu

In this talk, we consider the problem of optimally controlling an inventory in the presence of uncertainty. In particular, we suppose that the joint distribution (over time) of the sequence of demands belongs to some unknown set of joint distributions, and wish to pick the control policy which is optimal against a worst-case distribution belonging to this set. Departing from previous work, in which the adversary is typically restricted to product measures, we consider the setting in which the joint distribution must take the form of a martingale with bounded support. We characterize the optimal policy in this setting, and draw some interesting conclusions about how well one can perform under different assumptions on the joint distribution of demand.

Asymptotically Optimal Online Stochastic Bin Packing

Varun Gupta, University of Chicago, USA, guptav@uchicago.edu
Ana Radovanović, Google Research, New York, USA, anaradovanovic@google.com

Motivated by the problem of packing Virtual Machines on physical servers in the cloud, we study the problem of online stochastic bin packing. Items with sizes i.i.d. from an unknown distribution with integral support arrive as a stream and must be packed on arrival in bins of size B, also an integer, with the goal of minimizing the number of bins used. All currently known heuristics for online stochastic bin packing are either optimal for only certain classes of item size distributions, or rely on learning the distribution. The state-of-the-art Sum of Squares heuristic (Csirik et al.) obtains sublinear (in number of items seen) waste for distributions where the expected waste for the optimal offline algorithm is sublinear, but has a constant factor larger waste for distributions with linear waste under OPT.

We present the first distribution agnostic bin packing heuristic that achieves additive $O(\sqrt{n})$ waste compared to OPT for all distributions. Our algorithm is essentially gradient descent on a suitably defined Lagrangian relaxation of the bin packing Linear Program, and extends to multi-dimensional packing constraints as well.

Next, we consider the more general problem of online stochastic bin packing with item departures where the time requirement of an item is unknown at arrival. Our algorithm extends as is to the case of item departures. We also briefly revisit the Best Fit heuristic which has not been studied in the scenario of item departures yet.

Optimal Replacement of a Component in a Partially-Observable Environment

John A. Flory, University of Pittsburgh, USA, jhnflory@gmail.com
Jeff Kharoufeh, University of Pittsburgh, USA, jkharouf@pitt.edu

Motivated by wind energy systems, we present a switching diffusion model to characterize the degradation of a component operating in a random environment. A Markov chain Monte Carlo scheme is used to infer the parameters of the environment model using real degradation observations and to update the component's lifetime distribution. We establish the optimality of threshold replacement policies when the environment is uncertain but degradation is observable using a partially-observable Markov decision process framework. Numerical examples will be presented to illustrate the estimation procedures and optimal replacement policies.

Selfish or Altruistic? The impact of customer routing in a self-service queue

Hakjin Chung, University of Michigan, United States, hakjin@umich.edu

Hyun-soo Ahn, University of Michigan, United States, hsahn@umich.edu
Rhonda Righter, UC Berkeley, United States, rrighter@ieor.berkeley.edu

We consider a firm who can provide service in both full-service and self-service modes and examine how customer's selfishness to improve their own utility influences the optimal service rate for full service. We show that selfish routing increases the traffic toward the full service and resultant delay while the socially optimal customers may choose to avoid the full-service line when the benefit of reducing the negative externalities for other customers outweighs their own benefit. Consequently, the provider tends to build a higher capacity than when customers self-regulate to increase the traffic. We also show that the duplicated efforts to increase the surplus from both customer's and firm sides may result in the lower surplus than when either customer or firm exerts effort.

Session WA4 - Ruin Probability and Related Applications, Chair: Kaishev in JVdC D

Expected discounted loss for a spectrally negative risk process

Esther Frostig, Department of Statistics, University of Haifa, Israel

We consider a spectrally negative risk where upon ruin event the deficit is immediately paid, and the process restarts from level 0. We consider risk processes with and without dividend barrier. In a risk process with a dividend barrier b, all the surplus above b is paid as dividend. We find the expected discounted deficit (loss) and the expected discounted dividends in terms of the scale function of the process. We also obtain the average loss rate and the expected dividends per time unit.

A Risk Model with Reporting Delays

Andrei Badescu, University of Toronto, Canada, badescu@utstat.toronto.edu

We consider an insurance risk model where the time between accidents and reporting occurrences are exponential or Erlang distributed. Furthermore, we assume that when reported, a claim is paid at instant and its size follows a phase-type distribution. By translating the newly formed point process into a particular Markovian Arrival Process (MAP) structure and by using fluid flows, we obtain several ruin related measures. Numerical illustrations are presented in the end.

Evaluating Ruin Probabilities in a Dependent Risk Model

Dimitrina S. Dimitrova, Cass Business School, City University London, UK, d.dimitrova@city.ac.uk
Vladimir Kaishev, **Shouqi Zhao**, Cass Business School, City University London.

We consider a collective finite-horizon ruin probability model with Poisson claim arrivals, dependent claim amounts having any joint distribution and aggregate premium income represented by any non-decreasing positive, real valued function. In this paper, we first establish some enlightening connections between the finite-time ruin probability formulas obtained by Ignatov and Kaishev (2000, 2004) and provide their unified treatment in terms of classical Appell polynomials. Secondly, we consider their efficient numerical implementation and demonstrate that these formulas are useful not only theoretically but also in computing ruin probabilities in related practical applications. As recently pointed out by Das and Kratz (2012), the need to evaluate these ruin probability formulas naturally arises in the context of designing early warning systems against bankruptcy of (insurance) companies. Another application is in computing ruin probabilities and risk capital allocation under the so called dual risk model.

This talk is based on joint work with Vladimir Kaishev and Shouqi Zhao, Cass Business School, City University London.

Bibliography

Das, S., and Kratz, M. (2012). Alarm System for Insurance Companies: A strategy for Capital Allocation. *Insurance:Mathematics and Economics*, 51(1), 53–65.

Ignatov, Z.G., and V.K. Kaishev. (2000). Two-sided Bounds for the Finite-time Probability of Ruin. *Scandinavian Actuarial Journal, 2000*(1), 46–62.

Ignatov, Z.G., and V.K. Kaishev. (2004). A Finite-Time Ruin Probability Formula For Continuous Claim Severites. *Journal of Applied Probability, 41*(2), 570–578.

On Ruin Probability and Related Dual Models

Vladimir K. Kaishev, Cass Business School, City University London, UK, v.kaishev@city.ac.uk
Zvetan Ignatov, Sofia University
Dimitrina Dimitrova, **Shouqi Zhao**, Cass Business School, City University London.

We consider a reasonably general insurance risk model under which, cumulative premium income is modelled by any non-decreasing premium income function, consecutive claims arrive according to a point process and their severities may be dependent with any joint distribution. Under some assumptions on the claim arrival (point) process, we give closed form expressions, in terms of (generalized) Appell polynomials, for some risk related quantities such as the finite-time probability of ruin and the deficit at ruin. We further provide two dual interpretations of this model, one is the so called dual risk model and the second one is a dual queuing-theoretic model. We show that the ruin probabilistic results obtained under the insurance risk model are elegantly transferred to the context of the two dual models considered, thus providing new insights into these dual areas of research.

Session WA5 - Stochastic Systems, Chair: Squillante in JVdC E

Excessive delays in random-access networks

J.S.H. van Leeuwaarden, Eindhoven University of Technology, Eindhoven, The Netherlands, j.s.h.v.leeuwaarden@tue.nl
N. Bouman, Eindhoven University of Technology, Eindhoven, The Netherlands, n.bouman@tue.nl
S.C. Borst, Eindhoven University of Technology, Eindhoven, The Netherlands, s.c.borst@tue.nl

We explore the achievable delay performance in wireless random-access networks. While relatively simple and inherently distributed in nature, suitably designed backlog-based random-access schemes provide the striking capability to match the optimal throughput performance of centralized scheduling mechanisms in a wide range of scenarios. The specific type of activation rules for which throughput optimality has been established, may however yield excessive backlogs and delays. Motivated by that issue, we examine whether the poor delay performance is inherent to the basic operation of these schemes, or caused by the specific kind of activation rules. We first establish lower bounds for the delay in the case of backlog-based activation rules and show that the type of rules for which throughput optimality has been established yield excessive delays. We further examine fixed activation rates and establish lower bounds for the delay and mixing time. The bounds indicate that the delay and mixing time can dramatically grow with the load in certain topologies.

Diffusion scale tightness of invariant distributions of a large-scale flexible service system

A.L. Stolyar, Bell Labs, Alcatel-Lucent, New Jersey, USA, stolyar@research.bell-labs.com

The model is a flexible service system with multiple customer classes, multiple server pools; mean service time of a customer by a server depends on both the customer class and the server pool. We study a *Leaf Activity Priority* scheduling/routing policy, and consider the many-servers asymptotic regime: the customer arrival rates and the number

of servers in each pool tend to infinity in proportion to a scaling parameter r, while the overall system load remains strictly subcritical. Indexing the systems by r, we prove that the family of invariant distributions is tight on the diffusion, i.e. $r^{1/2}$, scale. Namely, the sequence of invariant distributions, centered at the equilibrium point and scaled down by $r^{-1/2}$, is tight. (This strengthens an earlier result, showing $r^{1/2+\epsilon}$-scale tightness for any $\epsilon > 0$.) This, in particular, implies a limit interchange: the limit of diffusion-scaled invariant distributions is equal to the invariant distribution of the limiting diffusion process.

Uniqueness of the stationary distribution of the diffusion approximation of a many-server queue

K. Ramanan, Brown University, Rhode Island, USA, Kavita_Ramanan@brown.edu
M. Aghajani, Brown University, Rhode Island, USA, mohammadreza_aghajani@brown.edu

A three-state Markov process was shown to arise as the diffusion limit for a many-server queue in the Halfin-Whitt asymptotic regime by Kaspi and Ramanan. We identify a more tractable two-state representation of this process that is still Markov, and show that this Markov process has a unique stationary distribution. In addition, we show that this stationary distribution is the limit of stationary distributions of the scaled state processes associated with the N-server queues. Standard Harris recurrence methods are not applicable to prove uniqueness of the stationary distribution because the Markovian process is infinite-dimensional. Instead, the proof relies on an asymptotic coupling approach that is potentially of broader interest. Time permitting, we will also discuss generalizations.

Explicit Solutions for the Stationary Distribution of a General Class of Markov Processes

M.S. Squillante, IBM Thomas J. Watson Research Center, New York, USA, mss@us.ibm.com
J.S.H. van Leeuwaarden, Eindhoven University of Technology, Eindhoven, The Netherlands, j.s.h.v.leeuwaarden@tue.nl
E.M.M. Winands, Rabobank and University of Amsterdam, Amsterdam, The Netherlands, e.m.m.winands@uva.nl

We consider a general class of two-dimensional Markov processes defined over the nonnegative quarter-plane, potentially infinite in both dimensions, and obtain exact solutions for the stationary distribution. More specifically, we derive an explicit solution for the rate matrix $\mathbf{R}$ of the invariant probability vector, as well as for the associated matrix $\mathbf{G}$. The probabilistic interpretations of these fundamental solution matrices allow us to describe their elements in terms of paths on the two-dimensional lattice. Then determining explicit expressions for the matrices becomes equivalent to solving a lattice path counting problem, the solution of which is derived using path decomposition, Bernoulli excursions, and hypergeometric functions. Time permitting, we consider a few applications and related large-deviation decay rates.

Session WA6 - Queueing Models for Computer Systems, Chair: Gautam in JVdC F

The Benefit of Introducing Variability in Quality Based Service Domains

A. Scheller-Wolf, Tepper School of Business, Carnegie Mellon University, awolf@andrew.cmu.edu
Y. Xu, Tepper School of Business, Carnegie Mellon University, yingx1@andrew.cmu.edu
K. Sycara, Robotics Institute, School of Computer Science, Carnegie Mellon University, katia@cs.cmu.edu

We consider a single-server queueing system in which the value customers obtain from service increases with their service time, but decreases with their waiting time. For such a system we show, surprisingly, that given a homogeneous customer population, system utility can be improved by increasing the variability in the system by varying the service rate. This is true even if the service rule is static, i.e., even if the service rate must be decided independent of the state of the system. Specifically, we show it is optimal to segment customers into service grades which are differentiated by their service rate (or equivalently mean service time). For such a system we derive the closed-form

optimal strategies of service rate differentiation, showing that optimal service rates and grade utilizations both form geometric sequences. We also compute the asymptotic system performance (as number of grades increases to infinity) and illustrate the sensitivity of the benefits of differentiation with respect to customers' characteristics: marginal service value, marginal waiting cost and variation of processing time. We find that providing differentiated service can improve system performance by 5%without any additional capacity investment.

Where Queueing Theory Meets Computer System Design

Mor Harchol-Balter, Carnegie Mellon University, Pittsburgh, PA, harchol@cs.cmu.edu

Queueing theory is extremely useful in the design of computer systems. In this talk we will discuss some of the less-traditional queueing models that come up when modeling computer systems. Examples include: closed-loop queueing systems, processor-sharing servers, very high job size variability, extremely high setup times, server speeds that change with load, and fairness considerations.

A Nonparametric Learning Approach to the Pricing and Provisioning Problem in Cloud Services

Yue Tan, Department of Integrated Systems Engineering, The Ohio State University, United States, tan.268@buckeyemail.osu.edu
Cathy H. Xia, Department of Integrated Systems Engineering, The Ohio State University, United States, xia.52@osu.edu

The emerging cloud computing service market aims at delivering computing resources as a utility over the Internet with high quality. The emerging nature of this market makes service demand highly uncertain. Traditional pricing and provisioning solutions that employ classical demand estimation approaches often fail to capture the dependence between the price-sensitive demand and the service level associated with the cloud service. As a result, those approaches can hardly guarantee convergence to the optimal offering. We present a non-parametric learning approach that provides much improved offering solutions. This learning approach could be used to calibrate the model and converge to the optimal solution. Though our model is developed in the context of cloud computing, the methodology can be applied to various applications with model uncertainty.

Stabilizing Queues with Non-Homogeneous and Multi-class Workloads in Data Centers

N. Gautam, Texas A&M University, USA, gautam@tamu.edu
S. Kwon, Texas A&M University, USA, soongeol@tamu.edu

We consider a system with multiple queues and heterogeneous servers where requests from multiple classes arrive in a time-varying fashion. For such a system, our goal is to manage resources so that the resulting performance measures are time-homogeneous. The key benefit of stabilizing the queue length process or sojourn time distribution is that it enables effective analysis by leveraging upon the vast literature for time-homogeneous systems. In addition, systems with smoothed traffic are much more conducive for developing control algorithms. We will first present several scenarios where we tune highly non-homogeneous systems to result in time-stable performance. Then we will illustrate our methodology in the context of energy and performance management in data centers and other distributed computing environments. Our objective is to develop strategies to: (i) assign classes to servers, (ii) determine the number of servers to be powered on, (iii) route requests to appropriate servers, and (iv) create a procedure for speed scaling. The requirement is to develop the aforementioned strategies under: (a) a distributed setting where real-time information is not exchanged between the sub-systems; (b) a necessity for time-stable performance; (c) a preference for simplified operations while maintaining cost-effectiveness and high quality of service. We illustrate our strategies and methodology using numerical examples.

Session WA7 - Estimation and Parameter Uncertainty in Queues, Chair: Nazarathy in JVdC G

Does the Past Predict the Future? The Case of Delay Announcements in Service Systems

Mor Armony, New York University, USA, marmony@stern.nyu.edu
Achal Bassamboo, Northwestern University, USA, a-bassamboo@kellogg.northwestern.edu
Rouba Ibrahim, Rouba Ibrahim, University College London, The United Kingdom, rouba.ibrahim@ucl.ac.uk

We investigate ways of making delay announcements in large service systems, such as call centers. We consider announcing the delay of the last customer to enter service (LES) to delayed customers, upon arrival. Customers typically respond to delay announcements by becoming more or less patient, and their response alters system performance. We explicitly model announcement-dependent customer behavior. We study the accuracy of the LES announcement in single-class multi-server Markovian queueing models with customer abandonment. We supplement our results with simulation.

Bayesian analysis of the $MAP_2/G/1$ queueing system

M.P. Wiper, Universidad Carlos III de Madrid, Spain, michael.wiper@uc3m.es
P. Ramírez Cobo, Universidad de Sevilla, Spain,jrcobo@us.es
R.E. Lillo, Universidad Carlos III de Madrid, Spain, lillo@est-econ.uc3m.es

The second-order Markovian arrival process or MAP_2, has been recently proposed as a versatile model for arrival traffic which allows for dependent arrivals. However, the basic MAP_2 model is not identifiable which leads to problems from an inferential viewpoint. Here, we use a recently developed, identifiable representation of the MAP_2 as the basis for an algorithm which allows us to undertake Bayesian inference for the MAP_2 arrival process. We use our approach to estimate the queue length and virtual waiting time distributions of a stationary $MAP_2/G/1$ queueing system, which is a generalization of the $M/G/1$ system that allows for dependent inter-arrival times. Our procedure is illustrated with applications in internet traffic analysis.

Sequential Staffing in Call Centers with Parameter Uncertainty

John J. Hasenbein, The University of Texas Austin, USA, jhas@mail.utexas.edu
David P. Morton, The University of Texas Austin, USA, morton@mail.utexas.edu
Jing Zan, Zilliant, Inc., USA, jzan@utexas.edu

We consider staffing agents in which an initial forecast of the call volume is given at the beginning of the day. After an observation period, the manager has the ability to update the staffing level based on a revised forecast. The manager operates under QoS constraints which can be very general. The resulting problem is formulated as a two-stage stochastic program. When utilization is the metric, the resulting solution can be written in a closed form analogous to a newsvendor solution.

Session WA8 - Approximation and Computation of Stochastic Models: Recent Advances, Chair: Lam in JVdC H

Extinction Escape Times

K.Z. Leder, University of Minnesota, USA, lede0024@umn.edul

J. Foo, University of Minnesota, USA, jyfoo@math.umn.edu
M. Kelly University of Minnesota, USA, mbkelly@math.umn.edu

Consider a large population of cells that is decaying exponentially according to a subcritical branching process. In the examples we are interested in this decay is driven by the presence of a drug that reduces cellular viability. Without intervention the population will eventually go extinct, however in many important scenarios mutant cells can be created that are resistant to this drug. These mutant individuals then go onto create supercritical branching processes and thereby allowing the total population to escape from extinction. In such a process these mutant individuals may eventually overtake the entire population. In this talk I will discuss approximations to the time when this population takeover occurs under various assumptions on the dynamics of the mutations and growth behavior of the mutant population.

Modeling and Analysis of Rogue Targets Avoiding a Detection

Jose Blanchet, Columbia University, United States, jose.blanchet@columbia.edu
Aya Wallwater, Columbia University, United States, aw2589@columbia.edu

A target travels around a region according to a diffusion process with an instantaneous drift which locally maximizes an objective function and with a noise correction. At the beginning obstacles are placed according to a non-homogeneous Poisson spacial process (all the obstacles are placed at once and independently of the diffusion process). Motivated by applications related to tracking and location of rogue objects, we are interested in studying the conditional distribution of the target given that it has evaded the obstacles for long time

Most of the talk focuses on the problem formulation and the analysis, which provides insights leading to a game in which the captor chooses a deployment policy and the target chooses a confinement region. We also hope to describe a Monte Carlo algorithm that evaluates the cost of a given policy. The algorithm can be shown to be asymptotically optimal (in the sense of variance minimization and running time) in a large deviations regime as the time horizon increases and the number of obstacles per unit area is suitably large.

Perfect sampling for infinite server and loss systems

Jose Blanchet, Columbia University, U.S. , jose.blanchet@columbia.edu
Jing Dong, Columbia University, U.S. , jd2736@columbia.edu

We present the first class of perfect sampling (also known as exact simulation) algorithms for the steady-state distribution of non-Markovian loss networks. We use a variation of Dominated Coupling From The Past for which we simulate a stationary infinite server queue backwards in time and analyze the running time in heavy traffic. We use the infinite server queue as an upper bound process to simulate loss systems. The running time analysis of our perfect sampling algo- rithm for loss systems is performed in the Quality-Driven (QD) and the Quality-and-Efficiency- Driven regimes. In both cases, we show that our algorithm achieves sub-exponential complexity as both the number of servers and the arrival rate increase. Moreover, in the QD regime, our algorithm achieves a nearly optimal rate of convergence.

Stochastic Model for Limit Order Book with Asymptotic Analysis

Jose Blanchet, Columbia University, United States, jose.blanchet@columbia.edu
Xinyun Chen, Columbia University, United States, xc2177@columbia.edu

We develop a stochastic model for the limit order book that describe the joint evolution of the full order book and the stock price. The model is able to capture stylizing features of limit order book from empirical observations. In the end, we show that our model leads to a jump diffusion approximation for the joint evolution of the stock price and the spread size.

8

Author Index

G

H

I

J

K

L

M

N

O

P

R

S

T

V

W

X

Y

Z

Tuesday July 16, 2013

	Track 1	Track 2	Track 3	Track 4	Track 5	Track 6	Track 7	Track 8	Track 9
Room	JVdC A	JVdC B	JVdC C	JVdC D	JVdC E	JVdC F	JVdC G	JVdC H	Cabildo I
TA *8:30am - 10:00am*	*Random Graphs: Connectivity* **Britton** Misra Jelenković Müller Litvak	*Applied Probability in Healthcare* **Bayati** Anderson Ding Chan Goh	*Simulation Optimization 1* **Fu / Henderson / Zhou** Garcia Zhou Calvin Henderson	*Quantitative Financial Risk Management* **Zhang / Ghamami** Chen Amini Ghamami	*Many-server queues: Approximations and control* **Gurvich** Perry Armony Liu Van Leeuwaarden	*Queueing Networks* **D'Auria / Van Leeuwaarden** Zocca Sanders Jonckheere D'Auria	*Opt. Control of Stochastic Inv. Systems* **Squillante** Goldberg Squillante	*Estimation* **Pasupathy** Kim Aminzadeh Pasupathy	
10:15am - 11:15am	**Tutorial - Bill Massey** (Room: Juan Vasquez de Coronado D-E)								
TB *11:30am - 1:00pm*	*Random Graphs: Social Networks* **Olvera-Cravioto** Zaman Steele Bonato Pralat	*Stoch. Control and Revenue Mgmt. in Smart Grids* **Zhang** Feinberg Tong Ghosh Zhang	*Simulation Optimization 2* **Fu / Henderson / Zhou** Hunter Frazier Nelson Sen	*Stochastic Calculus and Finance* **Chigansky** Hamza Klebaner Chigansky	*Fluid limits and its applications* **Van Houdt** Bortolussi Rubino Van Houdt	*Transient Functionals of Structured MCs* **Taylor** Nazarathy Hautphenne Fralix	*Control of Queues* **Down** Ayhan Lewis Ghosh Ward	*Stochastic Models* **Jonckheere** Miyoshi Kim Badila Jonckheere	*2D MPs: Fluid Queues and Random Walks* **Nguyen** Miyazawa Kobayashi Latouche
1:00pm - 2.30pm	**Lunch**								
TC *2:30pm - 4:00pm*	*Optimization* **Kalathil** Ferreira Yu Kalathil	*Stochastic systems in applications* **Lu** Radovanović Zhang Lu Squillante	*Simulation and Estimation* **Lam** Lam	*Markov modulation in risk and queueing processes* **Palmowski / D'Auria** Albrecher Ivanovs Nguyen Dorsman	*Spatial Stochastic Processes and their Applications* **Ramanan** Leder Foo Iyer Van Handel	*Funding Opportunities in OR at NSF* **Jacobson** Jacobson	*Stochastic Control* **Gupta / Keskin** Sharma Bhat Goldberg Keskin	*Novel directions in queueing* **Kapodistria** Adan Bassamboo Baer Kapodistria	
4:15pm - 5:15pm	**Tutorial - Tsachy Weissman** (Room: Juan Vasquez de Coronado D-E)								
6:00pm - 6:45pm	**APS Business Meeting** (Room: Cabildo I)								
6:30pm - 7:30pm	**Refreshments**								
7:30pm - 9:30pm	**Banquet (with musical entertainment)** (Location: Juan Vasquez de Coronado D-E)								

Wednesday July 17, 2013

	Track 1	Track 2	Track 3	Track 4	Track 5	Track 6	Track 7	Track 8	Track 9
Room	JVdC A	JVdC B	JVdC C	JVdC D	JVdC E	JVdC F	JVdC G	JVdC H	Cabildo I
8:30am - 9:30am	**Plenary - Paul Glasserman** (Room: Juan Vasquez de Coronado D-E)								
WA *10:00am - 11:30am*	*Urn models* **Maulik** Bandyopadhyay Cotar Maulik	*Finance and Risk* **Iyer** Jasin Watier Sethi Iyer	*Algorithmic and policy-level appl. of probability* **Kharoufeh** / **Goldberg** Goldberg Gupta Flory Ahn	*Ruin Probability and Related Applications* **Kaishev** Frostig Badescu Dimitrova Kaishev	*Stochastic Systems* **Squillante** Van Leeuwaarden Stolyar Ramanan Squillante	*Queueing Models for Computer Systems* **Gautam** Scheller-Wolf Harchol-Balter Tan Gautam	*Estimation and Param. Uncertainty in Queues* **Nazarathy** Armony Wiper Hasenbein	*Approx. and Comp. of Stoch. Models* **Lam** Leder Wallwater Dong Chen	
11:45am - 1.15pm	**Lunch**								
1:30pm - 7.00pm	**Social event: Visit to the Gold Museum**								