Course Syllabus

Instructor

Yunan Liu Office: 446 Daniels, Email: yunan_liu@ncsu.edu, Phone: 919-513-7208

Course Description

This is a seminar course on *stochastic modeling with applications in queueing theory*, as a natural continuation of ISE 760. One goal is to help students learn about various application context. A second goal is to focus on a class of mathematical models and analysis techniques that have proven useful in the application context. As is almost always the case in operations research, these models and analysis techniques have many other applications, so that the course can be useful even if you are primarily interested in other applications.

From the mathematical perspective, the course consists of both the conventional single-server queues and the recently developed multi-server queues (and networks of such multi-server queues). Important customer behavior includes *abandoning* (leaving after waiting for a while), retrying (coming back later after abandoning) and returning (coming back for additional service). There may be multiple types of customers and customer service representatives (agents) with different sets of skills.

We hope to achieve the following goals:

- (a) We keep up the progress of learning (reading) at a rate $\lambda = 1$ paper per week.
- (b) Students practise their presentation skill (in preparation for future conference presentations).
- (c) Students practise their teaching skill (in preparation for future independent teaching).
- (d) We hope the learning and thinking can broaden our views and can be turned into potential future research opportunities.

Time and Place

We meet once a week, Wednesday 4:30-6:00, Daniels 479.

Prerequisites

This course is intended for **doctoral** students in operations research and related fields. Student are expected to have completed an introductory course on stochastic models at the level of the first-year doctoral course ISE 760.

Course Workload

Each class consists of (i) a 45-minute lecture and (ii) a 45-minute presentation of research papers. During the first half of each meeting, a student (or the instructor) gives a 45-minute lecture on a specific topic. These topics are beyond those that are covered in ISE 760.

Candidate topics for the mini lectures are:

- Brownian motion (BM), reflected BM (RBM) and Donsker's theorem
- Ornstein-Uhlenbeck (OU) process
- Phase-type distribution: discrete and continuous
- Pollaczek-Khintchine formula and M/GI/1 queues
- Lindley equation and GI/GI/1 queues
- Little's law (LL) and rate conservation law (RCL)
- Poisson arrival sees time average (PASTA)
- Martingale

During the second half of each meeting, a student (or the instructor) presents a research paper. A suggested paper list is given at the end of this syllabus.

There will be **no** other assignments or exams. At the end of the semester, students are expected to finish a mini project. For instance, a project can be the implementation of a specific algorithm in a paper or a brief (say 3-page) survey of one or more research papers.

<u>Recommended Text</u> (not required)

- Ward Whitt, Stochastic-Process Limits. Springer (2002).
- Soren Asmussen, Applied Probability and Queues, 2nd ed. Springer (2003).
- Karl Sigman, Stationary Marked Point Processes: An Intuitive Approach. Chapman and Hall/CRC (1995).
- Pierre Brmaud, Point Processes and Queues: Martingale Dynamics. Springer (1981).

Suggested Research Papers

Suggested research papers are listed below. A student can choose any paper from the list (or even outside the list). A student should present 1-2 papers.

- 1. Staffing queues with time-varying arrivals to achieve system stability:
 - (a) Green, L. V., P. J. Kolesar, W. Whitt. Coping with time-varying demand when setting staffing requirements for a service system. *Production and Operations Management.* 16 13-39 (2007).
 - (b) Jennings, O. B., A. Mandelbaum, W. A. Massey, W. Whitt. Server staffing to meet time-varying demand. *Management Science*. 42 1383-1394 (1996).
 - (c) Feldman, Z., A. Mandelbaum, W. A. Massey, W. Whitt. Staffing of time-varying queues to achieve time-stable performance. *Management Science*. 54 324-338 (2008).

- (d) Liu, Y., W. Whitt. Stabilizing Customer Abandonment in Many-Server Queues with Time-Varying Arrivals. Under revision in *Operations Research*. (2011).
- (e) Yom-tov, G., A. Mandelbaum. Erlang-R: A Time-Varying Queue with ReEntrant Customers. Submitted to *Manufacturing and Service Operations Management*. (2011).
- 2. Many-server heavy-traffic (MSHT) limits of queues:
 - (a) Halfin, S., W. Whitt. Heavy-Traffic Limits for Queues with Many Exponential Servers. Operations Research. 29 567-588 (1981).
 - (b) Whitt, W. Understanding the Efficiency of Multi-Server Service Systems. Management Science. 38(5) 708-723 (1992).
 - (c) Garnette, O., A. Mandelbaum, M. Reiman. Designing a call center with impatient customers. *Manufacturing Service Operations Management.* **4** 208-227 (2002).
 - (d) Whitt, W. Efficiency-driven heavy-traffic approximations for many-server queues with abandonments. *Management Science*. **50** 1449-1461 (2004).
 - (e) Whitt, W. Fluid models for multiserver queues with abandonments. Operations Research. 54 37-54 (2006).
 - (f) Mandelbaum, A., S. Zeltyn. Staffing Many-Server Queues with Impatient Customers: Constraint Satisfaction in Call Centers. *Operations Research.* **57** (2009).
 - (g) Liu, Y., W. Whitt. The $G_t/GI/s_t + GI$ Many-Server Fluid Queue. Under revision at *Queueing Systems*. (2011).
- 3. Conventional heavy-traffic limits of queues:
 - (a) Kingman, J. F. C. On queues in heavy traffic. J. of the Royal Statistical Society. 24 383-392 (1962).
 - (b) Iglehart, D. L., Whitt, W. Multiple channel queues in heavy traffic, II: sequences, networks, and batches. Adv. Appl. Probab. 2 355-369 (1970).
- 4. Network queues:
 - (a) Mandelbaum, A., W. A. Massey, M. I. Reiman. Strong approximations for Markovian service networks. *Queueing Systems*. **30** 149-201 (1998).
 - (b) Talreja, R., W. Whitt. Fluid Models for Overloaded Multi-Class Many-Server Queueing Systems with FCFS Routing. *Management Science*. 54 1513-1527 (2008).
 - (c) Liu, Y., W. Whitt. A Network of Time-Varying Many-Server Fluid Queues with Customer Abandonment. Operations Research. 59 835-846 (2011).
 - (d) Perry, O., W. Whitt. Responding to Unexpected Overloads in Large-Scale Service Systems. *Management Science* 55 1353-1367 (2009).
- 5. Queues with phase-type distributions:
 - (a) Nelson, B.L., Michael R. Taaffe, The Ph_t/Ph_t/∞ Queueing System: Part I The Single Node, INFORMS JOC 16 (3), 2004, pp. 266-274.
 - (b) Nelson, B.L., Michael R. Taaffe, The (Ph_t/Ph_t/∞)^k Queueing System: Part II The Multiclass Network, INFORMS JOC 16 (3), 2004, pp. 275-283

- (c) Puhalskii, A. A., M. I. Reiman. The Multiclass GI/PH/N Queue in the Halfin-Whitt Regime. Adv. Appl. Prob. 32 564-595 (2000).
- (d) Dai, J. G., S. He, T. Tezcan. Many-Server Diffusion Limits for G/Ph/n + GI Queues. Annals of Applied Probability. **20** 1854-1890 (2010).
- (e) Liu, Y., W. Whitt. A Fluid Model for Many-Server Queues with Time-Varying Arrivals and Phase-Type Service Distribution. Working paper.
- 6. Delay announcements:
 - (a) Ibrahim, R., W. Whitt. Real-Time Delay Estimation Based on Delay History. Manufacturing and Service Operations Management. 11 397-415 (2009).
 - (b) Ibrahim, R., W. Whitt. Wait-Time Predictors for Customer Service Systems with Time-Varying Demand and Capacity. Operations Research. 59 1106-1118 (2011).
 - (c) Armony, M., N. Shimkin, W. Whitt. The Impact of Delay Announcements in Many-Server Queues with Abandonment. Operations Research. 57 66-81 (2009)
 - (d) Gad Alon and Achal Bassamboo. The Impact of Delaying Delay Announcements. Operations Research. 59 1198-1210 (2011).
- 7. Queueing in health care:
 - (a) Jennings, O. B., F. de Vericourt. Nurse-to-patient ratios in hospital staffing: a queueing perspective. Working paper (2007).
 - (b) de Bruin, A., G. Koole, B. van Rossum. Modeling the Emergency Cardiac In-Patient Flow: An Application of Queueing Theory. *Health Care Management Science*. **10** 125-137 (2007).
 - (c) de Bruin, A., R. Bekker, G. Koole, L. van Zanten. Dimensioning Hospital Wards Using the Erlang Loss Model. *Annals of Operations Research* forthcoming (2011).
 - (d) Chan, W. C., G. Yom-Tov, G. Escobar. Does Speedup Reduce Congestion? An Examination of Intensive Care Units with Readmissions. Working paper (2011).
 - (e) Bekker, R., and A. de Bruin. Time-Dependent Analysis for Refused Admissions in Clinical Wards. *Annals of Operations Research* (2009).
 - (f) Green, L., S. Savin. Providing Timely Access to Medical Care: A Queueing Model. Under revision for *Operations Research*.
- 8. Survey papers on stochastic-process limits in many-server queues:
 - (a) Pang, G., R. Talreja, W. Whitt. Martingale Proofs of Many-Server Heavy-Traffic Limits for Markovian Queues. *Probability Surveys* 4 193-267 (2007).
 - (b) Dai, J. G., S. He, Queues in Service Systems: Customer Abandonment and Diffusion Approximations. *Tutorials in Operations Research*. 31-59 (2011).
 - (c) Ward, R. A. Asymptotic Analysis of Queueing Systems with Reneging: A Survey of Results for FIFO, Single Class Models. Surveys in Operations Research and Management Science. 16 1-14 (2011).
- 9. Fairness and engineering approximations:
 - (a) Armony, M., A. R. Ward. Fair Dynamic Routing in Large-Scale Heterogeneous-Server Systems, Operations Research. 58 624-637 (2010).

- (b) Whitt, W. Engineering solution of a basic call-center model. *Management Science*. **51** 221-235 (2005).
- 10. Little's law and extensions: ("It's little but it's the law!")
 - (a) Little, J. A Proof of the Queue Formula: $L = \lambda W$. Operations Research. 9 383-387 (1961).
 - (b) Little, J. Little's Law as Viewed on Its 50th Anniversary. Operations Research. 59 536-549 (2011).
 - (c) Glynn, P. W., W. Whitt. A Central-Limit-Theorem Version of $L = \lambda W$. Queueing Systems. **2** 191-215 (1986).
 - (d) Whitt, W. A Review of $L = \lambda W$ and Extensions. Queueing Systems. 9 235-268 (1991).