Staffing and Scheduling to Differentiate Service in Multiclass Time-Varying Service Systems

Yunan Liu
Edward P. Fitts Department of Industrial and Systems Engineering
North Carolina State University

January 23, 2019

Abstract: Queueing theory is a field driven by applications. But unfortunately, there still remains a large gap between tractable theoretical studies and practical applications, such as call centers and health care systems, which have many realistic features (e.g., time-varying arrivals, customer abandonment, general probability distributions, and complicated network structures). In response to these challenges, we study a practical queueing system having multiple customer classes, nonstationary customer arrivals, and customer abandonment. We will develop effective staffing rules (number of servers) and scheduling policies (assigning newly idle servers to a waiting customer from one of the classes), with the objective of achieving differentiated service levels for each customer class.

One notable motivation of this research is the Canadian triage and acuity scale (CTAS) guideline that classifies patients in the emergency department (ED) into five acuity levels. In particular, CTAS requires that level i patients need to be seen by a physician within w_i minutes $100\alpha_i\%$ of the time", with $(w_1, w_2, w_3, w_4, w_5) = (0, 15, 30, 60, 120)$ minutes, and $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (0.98, 0.95, 0.9, 0.85, 0.8)$. Our goal is to devise new control principles (via staffing and scheduling) to guarantee that $\mathbb{P}(W_i(t) > w_i) \leq \alpha_i$, that is, the probability that a class-i customer waits more than w_i does not exceed α_i at all times for all classes. Our new joint staffing and scheduling policy is both time dependent (which copes with the time variability in arrival pattern) and state dependent (which dynamically captures the stochastic variability in service times and arrival times). Effectiveness of our policy is substantiated by heavy-traffic asymptotic optimality theorems. We also conduct computer simulation experiments to provide engineering confirmation and to gain insights.

Talk will take place from 2:30PM - 3:30PM, Watt Family Innovation Center, Room 203

Bio: Yunan Liu is an associate professor at the Industrial and Systems Engineering Department and an associate faculty member of the Operations Research Center of North Carolina State University. His research interests include queueing theory, stochastic modeling, applied probability, simulation, and their applications in service systems including call centers, healthcare, manufacturing, and blockchain systems. He received his M.S. and Ph.D. in Operations Research from Columbia University and B.S. in Electrical Engineering from Tsinghua University.