Title: Large Scale Queueing Systems. Heavy Traffic Asymptotics and Insights.

Abstract:

Modern service systems, including those arising in health care, cloud computing and call centers, are characterized by the large scale property. A variety of them can be modeled by queueing systems with many parallel servers, operating under a variety of service disciplines. In this talk we will consider two of such widely studied service disciplines: Join-the-Shortest service discipline, and the scheduling policy popularly dubbed Supermarket Model, both in the heavy traffic regime.

For the Join-the-Shortest discipline we establish that in the heavy traffic limit, the system can be described as a two-dimensional Ornstein–Uhlenbeck process, corresponding to the number of idle servers and the number of servers with precisely one waiting customer. All longer queues have asymptotically lower scale.

For the Supermarket Model we give a very detailed description of the dynamics of short queues using a certain "corrected" system of differential equations.

Joint work with Patrick Eschenfeldt (MIT)

Paper links:

“Join the Shortest Queue with Many Servers. The Heavy Traffic Asymptotics”

“Supermarket Queueing System in the Heavy Traffic Regime. Short Queue Dynamics”
https://arxiv.org/abs/1610.03522