Motivation: customers’ abandonment have a huge economical impact.
- Information technology: call centers, The Internet,...
- Marketing: online shopping,...
- Hospital triage: mass casualty events,...

Objective: develop resource allocation algorithms to optimize the performance.

Method:
- Model the impatience phenomena mathematically.
- Very complex phenomena → need to use approximations:
  1. Fluid Approximation (based on average evolution of the process),
  2. Lagrangian Relaxation (relax the set of admissible policies).
- Derive insights and easy-to-implement policies.
- Assess the accuracy of the solutions.

Fluid Approximation:
- Average time evolution.
- 2 classes and linear costs.
- Approximate stochastic control problem by a deterministic fluid control problem.

Main insights:
- Optimal policy: small population \( C \mu \) / \( \theta \)
- Good performance in light load and overload.

Lagrangian Relaxation:
- Constraint Relaxation: leads to decomposition and reduction of problem’s dimensionality.
- Allows us to consider: arbitrary cost functions, arbitrary number of classes,...
- Derivation of policies for more general models.

Main insights:
- Linear costs: solution coincides with fluid index.
- Performs well in high load.

Perspectives:
- Establish optimality of indices in heavy traffic.
- Fluid approximation for more than 2 classes.
- Multiple server case: applications into call centers.
- Impact on the internet: how to design new protocols?

Stochastic Model: arrival rate: \( \lambda_k \); abandonment rate: \( \delta_k \);
service rate: \( \mu_k \).
Cost of holding users: \( C_k(n_k) \); abandonment penalty: \( d_k \).

References: