排队论

随机服务系统的理论与应用 • 博士进阶课程

MarkZZZ WeChat: MarkZZZ20XX

课程简介

排队论（Queueing Theory）是研究随机服务系统中等待现象的数学理论，是运筹学、概率论与随机过程交汇的核心领域。本课程以随机过程为语言，系统讲授排队论的经典理论与现代进展。从泊松过程、Markov 链和更新过程出发，深入推导 M/M/1、M/G/1、GI/G/1 等基本模型，进而扩展至排队网络（Jackson 定理完整证明）、非乘积形网络、流体与扩散极限，直至排队控制与应用前沿。

全程强调严格数学推导：Pollaczek-Khinchine（PK）公式从嵌入 Markov 链出发完整推导；c 规则的最优性严格证明；Jackson 网络乘积形定理从流平衡方程到最终结果逐步展开。课程还涵盖重尾分布与次指数性、均场极限等前沿工具，帮助学生建立从模型构建到数学分析的完整能力，并掌握现代排队论在云计算、通信网络、医疗运营与供应链中的应用。

适合人群

- 已掌握概率论与随机过程基础（测度论概率、Markov 链），希望系统学习排队论理论的博士生与研究者
- 从事通信网络、云计算、交通系统、医疗运营、供应链等领域，需要建立随机排队模型进行性能分析的工程师与分析师
- 对随机过程、扩散极限、随机控制等方向感兴趣的数学与统计专业学生
- 希望深入理解 Little 定律、PASTA 定理、Burke 定理等经典结果并掌握其严格证明的学习者

前置知识

- **概率论**：测度论基础（σ-代数、Lebesgue 积分、几乎处处收敛）、特征函数、大数定律与中心极限定理、条件期望
- **随机过程**：离散时间 Markov 链（平稳分布、遍历定理、首达时间）；连续时间 Markov 链的基本概念
- **分析**：Laplace-Stieltjes 变换、生成函数、复变函数基础
- **线性代数**：矩阵指数、谱分解，对推导矩阵解析方法有帮助
- **随机控制基础**（第 11 讲）：动态规划、Bellman 方程有帮助但非必须

1 课程内容

讲次	主题	内容概要
1	概率基础与 PASTA	随机过程分类；连续时间 Markov 链（转移速率矩阵、Kolmogorov 方程）；平稳分布的存在性与唯一性（不可约正常返）；遍历定理；PASTA（泊松到达见时间平均）定理的严格证明；Little 定律 $L = \lambda W$ 及推论
2	泊松过程与更新过程	泊松过程的等价定义（指数间隔、独立增量、Poisson 分布）；分流与合流定理；条件分布与顺序统计；更新过程基本定理（更新定理）；超越分布；Blackwell 更新定理；剩余寿命与年龄分布
3	M/M/1 及其变体	指数分布的无记忆性；生灭过程的流平衡方程；M/M/1 平稳分布完整推导；性能指标（L, L_q, W, W_q）；M/M/c 队列；M/M/1/K 有限容量；Erlang B 公式与 Erlang C 公式的推导及应用
4	M/G/1 队列	嵌入 Markov 链与 Wald 恒等式；Pollaczek-Khinchine（PK）均值公式完整推导；PK 变换公式（母函数方法）；残余寿命与分解定理；M/G/1 变体（多优先级、N-策略、T-策略）
5	GI/G/1 队列	Lindley 递推方程与稳态解的存在性；Spitzer 恒等式；Kingman 近似公式（重流量渐近）；等待时间分布的界；扩散极限：重流量下的布朗运动近似（Iglehart-Whitt 定理）
6	优先级排队与多类顾客	非抢占优先级（NPPS）：守恒律与 PK 公式推广；抢占恢复优先级（PPS）：分析框架；保守性条件（Work Conservation Lemma）；插队方程（Arrival Theorem for Priority Queues）；Processor Sharing 队列；LCFS-PR 策略
7	排队网络：Jackson 定理	开放 Jackson 网络的模型定义；流平衡方程（交通方程）的建立；乘积形平稳分布的完整证明（局部流平衡）；Burke 定理（输出过程为泊松）；闭合 Jackson 网络与 BCMP 定理概述
8	非乘积形网络与近似分析	乘积形条件的本质；分解近似（Decomposition Approximation）；均场（Mean-Field）方法与 N-服务器中的浓度现象；Whittle 网络与部分平衡；扩散近似在网络中的应用
9	流体模型与扩散极限	排队流体模型（Fluid Queue）：On/Off 输入；重流量理论（Heavy Traffic Theory）回顾；Reflected Brownian Motion（RBM）的构造；Harrison-Reiman 定理；多维扩散极限与网络应用

讲次	主题	内容概要
10	重尾分布与次指数性	重尾分布的定义（次指数分布、正规变化）；Cramer-Lundberg 破产模型；重尾队列（M/G/1 with heavy-tailed service）；Lévy 过程简介；稳定分布与吸引域；重尾对排队性能的影响（$M/Pareto/1$ 举例）
11	排队控制与动态优化	最优调度问题框架（MDP 建模）；$c\mu$ 规则的严格证明（交换论证与价值函数方法）；动态定价（Dynamic Pricing）模型；多类队列调度的 Whittle 指数方法；准入控制（Admission Control）与可达条件
12	应用前沿与研究方向	云计算与数据中心排队：JSQ、Power-of-two 策略；通信网络：丢包分析与 QoS 保证；医疗运营：急诊排队、手术室调度；供应链中的库存与排队耦合；机器学习与排队论的结合（学习型调度）；开放研究问题与前沿方向概览

2 参考资料

1. Kleinrock, L. *Queueing Systems, Vol. 1: Theory.* Wiley, 1975.
2. Kleinrock, L. *Queueing Systems, Vol. 2: Computer Applications.* Wiley, 1976.
3. Kelly, F.P. *Reversibility and Stochastic Networks.* Wiley, 1979. (Cambridge University Press 再版, 2011.)
4. Asmussen, S. *Applied Probability and Queues.* 2nd ed., Springer, 2003.
5. Gross, D., Shortle, J.F., Thompson, J.M. & Harris, C.M. *Fundamentals of Queueing Theory.* 4th ed., Wiley, 2008.
6. Baccelli, F. & Brémaud, P. *Elements of Queueing Theory.* 2nd ed., Springer, 2003.
7. Bramson, M. *Stability of Queueing Networks.* Springer, 2008.
8. Harrison, J.M. *Brownian Motion and Stochastic Flow Systems.* Wiley, 1985.
9. Whitt, W. *Stochastic-Process Limits.* Springer, 2002.
10. Robert, P. *Stochastic Networks and Queues.* Springer, 2003.
11. Latouche, G. & Ramaswami, V. *Introduction to Matrix Analytic Methods in Stochastic Modeling.* SIAM, 1999.
12. Kumar, P.R. & Meyn, S.P. "Stability of queueing networks and scheduling policies." *IEEE Trans. Autom. Control* 40(2):251–260, 1995.