

Towards Pareto Frontier for Polling Systems

Akul Bansal, Veeraruna Kavitha

Industrial Engineering and Operations Research,
IIT Bombay

16 December, 2019

Polling Systems

- ▶ Single server and multi-class queuing systems.
- ▶ Arrivals, service times and switch over times are random.
- ▶ How to allocate a single server among multiple queues?

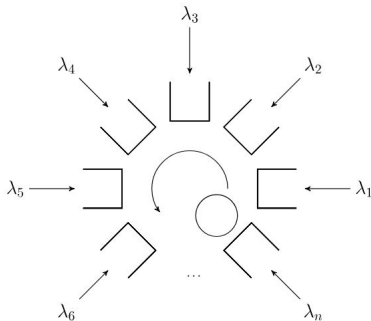


Figure 1: Polling system with n queues and one server ¹

¹Credits: "Polling system" Wikipedia, The Free Encyclopedia

Pareto Frontier

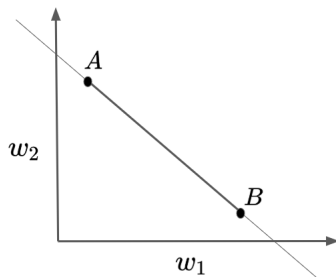


Figure 2: Achievable Region of M/G/1 System with 2 queues

- ▶ What is a solution to multiple-objective problem ?
- ▶ Pareto Frontier: A set of solution values where improving on one objective is not possible without losing on the other objective.

Brief Literature Review

- ▶ Applications of polling systems (Levy et al. 1990, Boon et al. 2011)
- ▶ Specific Schedulers (Koole et al. 1999)
- ▶ Static Optimal Policies (Boxma et al. 1990, Bertsimas et al. 1993)
- ▶ Randomized Policies (Kleinrock et al. 1988)
- ▶ Switching Costs vs. Switching Times (Riemann et al. 1998)
- ▶ Dynamic Optimal Policies (Hofri et al. 1988, Gandhi et al. 1996)
- ▶ Pareto Frontier (Ayush et al. 2014, Kiran et al. 2019)

Motivation

- ▶ Monte-Carlo simulation results for (β, ex) schedulers.
- ▶ (β, ex) scheduler: switch from Q_i , when $w_i\beta_i \leq w_{-i}$, where w_i is waiting time at queue i .

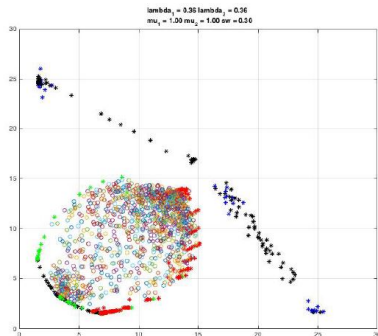


Figure 3: Achievable Region for waiting times (Yogesh and Kavitha 2016)

Polling Systems as Markov Decision Processes (MDP)

- ▶ Markov Decision Processes is a model for sequential decision making
- ▶ Rewards, actions and transition probabilities depend on present state only.

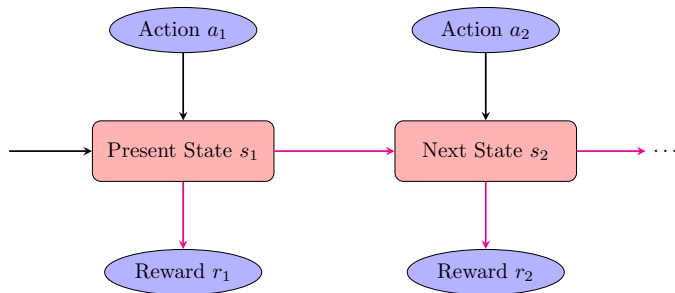


Figure 4: MDP Model

Key Components of MDP

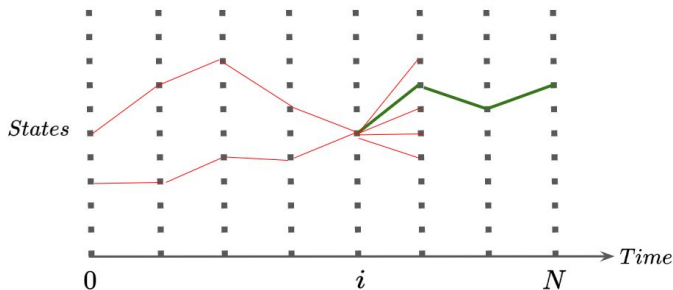
- ▶ Set of decision epochs T
- ▶ Set of possible states S , s_t is state at time t .
- ▶ Set of actions available A_S , a_t is action at time t .
- ▶ Decision rules $d_t : S \rightarrow A_S$
- ▶ Immediate rewards $r_t(s_t, a_t) : S \times A \rightarrow \mathbb{R}$
- ▶ Transition probabilities $p(j|s, a)$ to each future state j , given current state and action s, a .

Optimizing an MDP formulation

- ▶ Find an optimal sequence of decision rules (policies) that prescribe the action to be taken at any time i.e. $\pi = (d_1, d_2, \dots)$.
- ▶ When $N = \infty$, we consider total expected *discounted* reward
$$\max_{\pi} v_{\lambda}^{\pi}(s) = \lim_{N \rightarrow \infty} \mathbb{E} \left(\sum_{i=1}^N \lambda^{i-1} r(X_i, Y_i) \right) \text{ for } \forall s \in S$$

Optimality equations

- ▶ Under discounted objective, how do you determine the optimal policy?
- ▶ For finite horizon N , $\forall s \in S$ and $n < N$ solve the following system of equations: $v(s) = \max_{a \in A_s} \left(r(s, a) + \sum_{j \in S} \lambda p(j|s, a) v(j) \right)$
- ▶ Use *value iteration algorithm* to solve this system of equations.



Modelling 2-Queue Polling Systems as MDP

- Consider a two-queue polling system with M/M/1 queues.

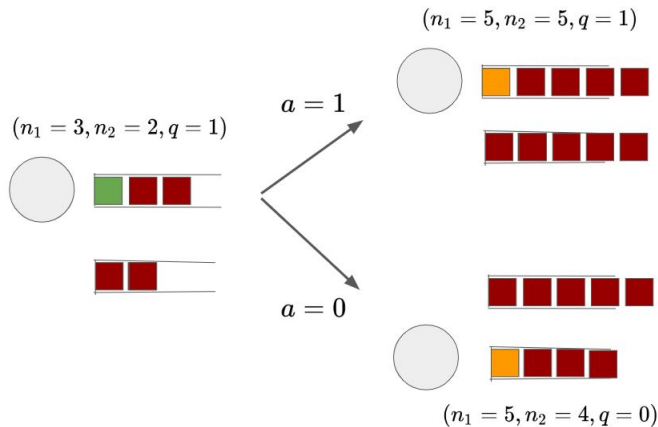


Figure 5: System state under different actions

Polling Systems as MDP - Rewards

- ▶ The objective is to minimize the average number of people in the two queues $\lim_{T \rightarrow \infty} \frac{\int_0^T N_1(t) + N_2(t) dt}{T}$
- ▶ Using PASTA we can write

$$\lim_{T \rightarrow \infty} \frac{\int_0^T N_1(t) dt}{T} = \lim_{K \rightarrow \infty} \frac{\sum_{k=1}^K N_1(\alpha_{1k}^-)}{K} = \lim_{K \rightarrow \infty} \frac{\sum_{k=1}^K N_1(d_{1k}^+)}{K} \text{ where}$$

K is the total number of departures in queue 1.

Polling Systems as MDP - Rewards

- ▶ In the two queue this can be written as

- ▶
$$\lim_{K \rightarrow \infty} \frac{\sum_{k=1}^K N_1(d_k^+) \mathcal{I}_{d_k \in Q_1}}{K} \frac{K}{D_1(K)} + \frac{\sum_{k=1}^K N_1(d_k^+) \mathcal{I}_{d_k \in Q_2}}{K} \frac{K}{D_2(K)}$$

- ▶
$$\frac{D_i(K)}{K} \rightarrow \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

- ▶
$$r(s_t, a_t) = iq \frac{\lambda_1 + \lambda_2}{\lambda_1} + j(1 - q)\alpha \frac{\lambda_1 + \lambda_2}{\lambda_2}$$

Transition Probabilities

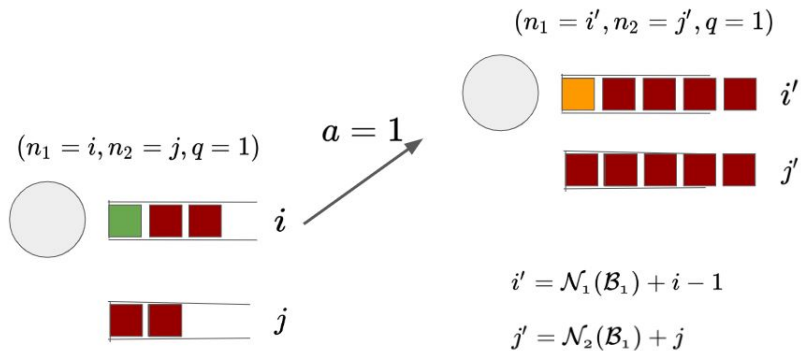


Figure 6: Transition probabilities under stayback

Transition Probabilities

$$i' = \mathcal{N}_1(\mathcal{W}_{12} + \mathcal{B}_1) + i - 1$$

$$j' = \mathcal{N}_2(\mathcal{W}_{12} + \mathcal{B}_1) + j$$

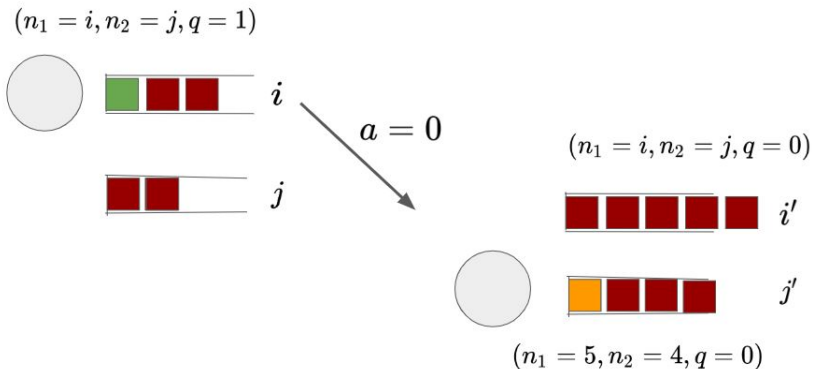


Figure 7: Transition Probabilities under switch

Transition Probabilities

► In general distribution of future states is defined as follows:

$$q' = qa + (1 - q)(1 - a)$$

$$i' = \mathcal{N}_1((1 - a)(qW_{12} + (1 - q)W_{21}) + qB_1 + (1 - q)B_2) + i - q$$

$$j' = \mathcal{N}_2((1 - a)(qW_{12} + (1 - q)W_{21}) + qB_1 + (1 - q)B_2) + j - (1 - q)$$

Simulation Results

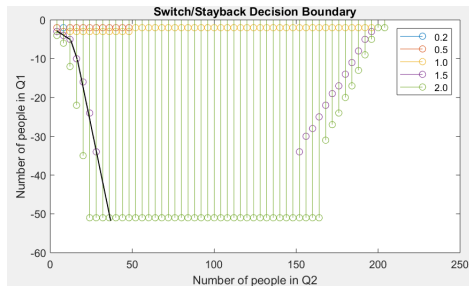


Figure 8: Simulation Results using Value Iteration Algorithm at Queue 1

- Closed-form expression for decision policy inaccessible due to truncation errors.
- Policy is exhaustive at the queue with higher weight.

Main Result

Consider a two queue polling system with number of people in queue i represented by $N_i(t)$, $i \in \{1, 2\}$ at time t . The objective is to minimize the weighted expectation of number of people in the system, that is,

$$\mathbb{E}N_1 + \alpha \mathbb{E}N_2 = \lim_{T \rightarrow \infty} \frac{\int_0^T N_1(t) + \alpha N_2(t) dt}{T} \approx \lim_{T \rightarrow \infty} \int_0^T \beta^t (N_1(t) + \alpha N_2(t)) dt \quad (1)$$

Theorem

When arrivals, service time and walking times have symmetric distribution at the two queues, then for $\alpha > 1$, the policy π that minimizes (1) is exhaustive at queue 2.

Proof Sketch

- ▶ Assume optimal policy $\bar{\pi}$ is non-exhaustive.
- ▶ Design a non-stationary policy $\tilde{\pi}$ which is better than $\bar{\pi}$.
- ▶ For comparison, couple service and waiting times under $\bar{\pi}$ and $\tilde{\pi}$.

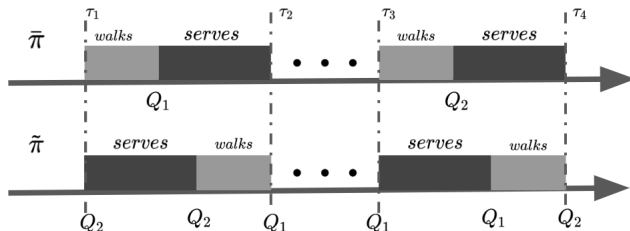


Figure 9: Policy $\bar{\pi}$ and $\tilde{\pi}$ when queue 1 is non-empty on arrival

Proof Sketch

$$\begin{aligned}\int_{\tau_1}^{\tau_4} \tilde{N}_1(t) + \alpha \tilde{N}_2(t) dt &= \int_{\tau_1}^{\tau_4} (\bar{N}_1(t) + 1) + \alpha(\bar{N}_2(t) - 1) dt \\ &= \int_{\tau_1}^{\tau_4} \bar{N}_1(t) + \alpha \bar{N}_2(t) dt + \int_{\tau_1}^{\tau_4} (1 - \alpha) dt\end{aligned}\tag{2}$$

Since, $\alpha > 1$ therefore,

$$\int_{\tau_1}^{\tau_4} \tilde{N}_1(t) + \alpha \tilde{N}_2(t) dt < \int_{\tau_1}^{\tau_4} \bar{N}_1(t) + \alpha \bar{N}_2(t) dt\tag{3}$$

We can therefore conclude that

$$\mathbb{E}^{\tilde{\pi}} N_1 + \alpha \mathbb{E}^{\tilde{\pi}} N_2 < \mathbb{E}^{\bar{\pi}} N_1 + \alpha \mathbb{E}^{\bar{\pi}} N_2\tag{4}$$

Future Work

- ▶ Deriving a new MDP formulation, given exhaustive policy at one of the queues.
- ▶ The goal is to find a closed-form expression or better approximation for the value function $v(s)$.