Accelerating Performance Inference over Closed Systems by Asymptotic Methods

Giuliano Casale
Imperial College London
London SW7 2AZ, UK
g.casale@imperial.ac.uk

ABSTRACT
Recent years have seen a rapid growth of interest in exploiting monitoring data collected from enterprise applications for automated management and performance analysis. In spite of this trend, even simple performance inference problems involving queuing theoretic formulas often incur computationally bottlenecks, for example upon computing likelihoods in models of batch systems. Motivated by this issue, we revisit the solution of multiclass closed queuing networks, which are popular models used to describe batch and distributed applications with parallelism constraints. We first prove that the normalizing constant of the equilibrium state probabilities of a closed model can be reformulated exactly as a multidimensional integral over the unit simplex. This gives as a by-product novel explicit expressions for the multiclass normalizing constant. We then derive a method based on cubature rules to efficiently evaluate the proposed integral form in small and medium-sized models. For large models, we propose novel asymptotic expansions and Monte Carlo sampling methods to efficiently and accurately approximate normalizing constants and likelihoods. We illustrate the resulting accuracy gains in problems involving optimization-based inference.

ACM Reference format:
DOI: http://dx.doi.org/10.1145/3078505.3078514

1 MODEL
We consider closed queuing networks with K single-server nodes, M = K infinite server nodes, R job classes, and product-form solution [1]. Let N_r be the number of jobs in class r and let N_r = (N_1, ..., N_K), N_r = Σ_i N_r, Let θ_r collect the service demands placed by class-r jobs at node k. The model has state space \mathcal{S}_M = \{n ∈ \mathbb{N}^M | n_k ≥ 0, Σ_i n_k = N_r \}, where n_k is the number of class-r jobs at node k. The normalizing constant of the state probabilities is given by [1]

\[ G_θ(N) = \sum_{n \in \mathcal{S}_M} \prod_{i=1}^K n_i! \prod_{r=1}^R \prod_{k=1}^M g_{kr}^{n_{kr}} n_{kr}^{-1} \]  

We study how to efficiently obtain G_θ(N), motivated by the problem of computing likelihoods for model parameterization, model selection, and statistical inference.

2 MAIN RESULTS
For networks without infinite servers, we show in [2] that an exact integral form for the normalizing constant is given by

\[ G_θ(N) = \frac{(N + K - 1)!}{N_1! \cdots N_R!} \int_\Delta_K \prod_{r=1}^R \prod_{k=1}^M \sum_{θ_k} \left( \sum_{k=1}^M \theta_k u_k \right) \]  

and let θ_k = (θ_k) be the unit simplex. Stemming from this novel integral form, [2] proposes exact and approximate computational methods for G_θ(N). For example, using cubature rules [3], (2) can be interpolated, either exactly or approximately, as

\[ G_θ(N) = \frac{(N + K - 1)!}{\prod_{i=1}^R N_r!} \sum_{i=0}^S \sum_{j=1}^{2^S} \sum_{d=0}^S \sum_{j=1}^{2^S} \frac{g_{kr}^{n_{kr}} n_{kr}^{-1}}{(2S + K - 2i)} \]  

where S specifies the interpolation degree, b ∈ \mathbb{N}^K, b = \sum_{i=1}^K b_i, and \sum_{k=1}^M n_k = N_r.

References