Product-Form Approximation of Queueing Networks with Phase-Type Service

Giuliano Casale, Peter G. Harrison, Maria Grazia Vigliotti*
Department of Computing — Imperial College London
Emails: g.casale@imperial.ac.uk, {pgh, mgv98}@doc.ic.ac.uk

1. INTRODUCTION

The applicability of queueing network models to real-world systems is often enhanced by characterizing service times using phase-type (PH) distributions. Unfortunately, the combinatorial explosion of the underlying state space makes it difficult to evaluate efficiently the resulting models. We propose to tackle this issue by a solution algorithm inspired by the Reversed Compound Agent Theorem (RCAT) [2].

Let the QBD process for the uphill queue be specified by the forward transition matrix $\boldsymbol{F}$, by the local transition matrix $\boldsymbol{L}$, and by the backward transition matrix $\boldsymbol{B}$. Also, let us indicate with $\boldsymbol{\alpha}_n$ the vector of state probability when the queue size is n. Further, let $(\boldsymbol{\beta}, \boldsymbol{T})$ be the representation of the PH distribution at the uphill queue. Then a RCAT product-form exists if there exist a scalar $\bar{q} > 0$ such that $\boldsymbol{\alpha}_n \hat{\boldsymbol{L}} = \boldsymbol{\alpha}_{n+1}\boldsymbol{B}$, $n \geq 1$, where $\hat{\boldsymbol{L}} = \mathrm{diag}(\bar{q}\delta_1, \bar{q}\delta_2, \ldots, \bar{q}\delta_K)$ and δ_k is 1 if the kth column of $-\boldsymbol{T}\boldsymbol{e}\boldsymbol{\beta}$ has at least one nonzero element, 0 otherwise [1].

An interesting observation arising in [1] from the above formula is that, for models with a RCAT product-form, there exist a matrix $\boldsymbol{H} = -\boldsymbol{F}(\boldsymbol{L} + \hat{\boldsymbol{L}})^{-1}$ such that $\boldsymbol{\alpha}_n = \boldsymbol{\alpha}_{n-1}\boldsymbol{H}$, $n \geq 2$. In general, this matrix differs from the rate matrix $\boldsymbol{R}$, even though they must share the same dominating eigenvalue η and $\boldsymbol{\alpha}_n = \boldsymbol{\alpha}_{n-1}\boldsymbol{R}$. Observe now that $\boldsymbol{\alpha}_{n+1} = \boldsymbol{\alpha}_n\boldsymbol{R} = \boldsymbol{\alpha}_n\boldsymbol{H}$, $n \geq 1$, thus $\boldsymbol{\alpha}_n(\boldsymbol{R} - \boldsymbol{H}) = 0$, which implies $\boldsymbol{\alpha}_n(\hat{\boldsymbol{L}} - \boldsymbol{R}\boldsymbol{B}) = 0$, $n \geq 1$. Since the last equation must hold for all $n \geq 1$, it suggests that the general solution admits the form $\boldsymbol{\alpha}_{n+k} = c_{n+k}\boldsymbol{\alpha}_{n+k-1}$, for some sequence of coefficients c_i, $i > 1$. Applying the Cayley-Hamilton theorem to $\boldsymbol{R}$, it can be shown that choosing $c_{n+k} \equiv \eta$ such that $\boldsymbol{\alpha}_{n+1} = \eta\boldsymbol{\alpha}_n$ provides always a solution that is consistent with the matrix geometric one for all $n \geq 1$. In particular, if $\boldsymbol{H}$ and $\boldsymbol{R}$ share only the eigenvalue η, the relation can be proved exactly.

Based on the above observations, one may then approximate the steady state of a queue satisfying RCAT's conditions by the relation $\boldsymbol{\alpha}_n = \eta\boldsymbol{\alpha}_{n-1}$ which, upon imposing $\sum_{n\geq 1}\boldsymbol{\alpha}_n\boldsymbol{1} = \rho$, ρ being the utilization of the queue, gives

$$\alpha_n = \boldsymbol{\alpha}_n\boldsymbol{1} \approx \begin{cases} 1-\rho, & n = 0 \\ \rho(1-\eta)\eta^{n-1}, & n \geq 1 \end{cases} \tag{1}$$

Notice that in the special case where the PH service distribution is exponential, this reduces to $\eta = \rho$ and (1) becomes the familiar $M/M/1$ steady-state queue-length distribution. For non-exponential PH queues, the solution (1) may be interpreted as the equilibrium solution of a load-dependent M/M/1 queue with service rates depending on ρ and η.

2. CLOSED NETWORK APPROXIMATION

We use (1) to define a simple approximate product-form solution, i.e., $\pi(n_1, n_2, \ldots, n_M) \approx G^{-1}\prod_{i=1}^{M} F_i(n_i)$, where $F_i(n)$ is defined identically to α_n for queue i assumed with utilization ρ_i, n_k is the population of jobs at queue k, G is a normalizing constant. The utilizations ρ_i are estimated iteratively starting from their values in a corresponding network of exponential queues. At each iteration, the utilization obtained by application of the approximate product-form is used to refine the guess for ρ_i. Such utilization is computed by exact formulas that involve normalizing constants only, which in turn are determined by a convolution algorithm.

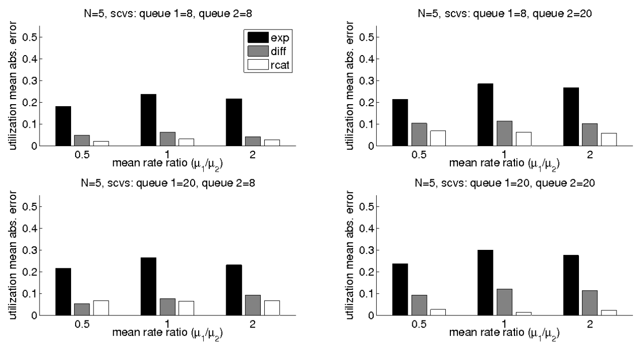

Figure 1: Comparison of exponential ('exp'), diffusion ('diff'), and RCAT approximations.

Figure 1 shows approximation results for the station utilizations in a cyclic model with 2 queues (mean service time μ_1^{-1} and μ_2^{-1}), think time $Z = 1$, and 5 jobs. The results indicate that our approximation is often much better than exponential and diffusion approximations, especially for large squared coefficient of variation (scv) of the service times.

3. REFERENCES

[1] G. Casale, P.G. Harrison, and M.G. Vigliotti. Product-Form Approximation of Tandem Queues via Matrix-Geometric Methods. In *Proc. of NSMC Workshop*, 2010.

[2] P.G. Harrison. Turning back time in Markovian process algebra. *Theor. Comp. Sci.*. 290(3):1947–1986, 2003.

*The work of G. Casale was supported by an Imperial College Junior Research Fellowship. The work of P. G. Harrison and M. G. Vigliotti was supported in part by the Engineering and Physical Sciences Research Council of the United Kingdom, research grant number EP/D061717/1.