Bounding the partition function of BCMP multiclass queueing networks

Jonatha Anselmi[1] and Paolo Cremonesi[1]

[1]Dipartimento of Elettronica e Informazione
Politecnico di Milano

BWWQT 2009, Minsk

Outline

1. Motivation
   - BCMP Queueing Networks
   - Difficulties of BCMP Queueing Networks

2. Our Result
   - Hölder's Inequality
   - Bounding the Partition Function
   - Load-Dependent Stations

Outline

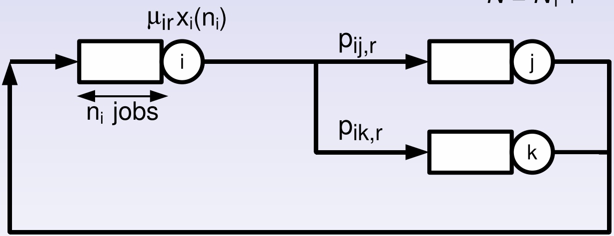
- Closed BCMP (**B**asket, **C**handy, **M**untz, **P**alacios, 1975) queueing networks have been widely adopted to analytically evaluate the performance of computer and communication systems,

- Extension of Gordon-Newell (1967) networks
  1. Jobs belong to multiple classes,
  2. more types of service disciplines,
  3. more general service times probability distributions.

Notation of a Closed BCMP Network

An Example

R: # job classes
M: # stations
N_r: # jobs of class r
$N = N_1 + \dots + N_R$



$\rho_{ir} = e_{ir} / \mu_{ir}$: mean loading (service demand) of class-r jobs in station i

Main assumptions of Closed BCMP Networks

	Per-class service time distribution
FCFS	Exponential (with mean μ_{ir}^{-1}, and $\mu_1 = \cdots = \mu_R$)
LCFS,PS,IS	Coxian (with mean μ_{ir}^{-1})

Routing probabilities $p_{ij,r}$ are constants.

Outline

Stationary Distribution of Closed BCMP Networks

The complexity of the model

BCMP models are CTMC, and the stationary distribution is

$$\pi(n_{11},\dots,n_{MR}) = \frac{1}{G(\mathbf{N})}\prod_{i=1}^{M}\left(\sum_{r=1}^{R} n_{ir}\right)!\prod_{r=1}^{R}\frac{D_{ir}^{n_{ir}}}{n_{ir}!} \tag{1}$$

$$G(\mathbf{N}) = \sum_{\mathbf{S}(\mathbf{N})}\prod_{i=1}^{M}\left(\sum_{r=1}^{R} n_{ir}\right)!\prod_{r=1}^{R}\frac{D_{ir}^{n_{ir}}}{n_{ir}!} \tag{2}$$

$$\mathbf{S}(\mathbf{N}) = \left\{n_{11},\dots,n_{MR} : \sum_{i=1}^{M} n_{ir} = N_r, 1 \le r \le R\right\}. \tag{3}$$

Then, the main problem is the computation of $G(\mathbf{N})$. In fact,

$$|\mathbf{S}(\mathbf{N})| = \prod_{r=1}^{R}\binom{N_r + M - 1}{M - 1} \tag{4}$$

Stationary Distribution of Closed BCMP Networks

The complexity of the model

BCMP models are CTMC, and the stationary distribution is

$$\pi(n_{11},\ldots,n_{MR}) = \frac{1}{G(\mathbf{N})} \prod_{i=1}^{M} \left(\sum_{r=1}^{R} n_{ir}\right)! \prod_{r=1}^{R} \frac{D_{ir}^{n_{ir}}}{n_{ir}!} \quad (1)$$

$$G(\mathbf{N}) = \sum_{\mathbf{S}(\mathbf{N})} \prod_{i=1}^{M} \left(\sum_{r=1}^{R} n_{ir}\right)! \prod_{r=1}^{R} \frac{D_{ir}^{n_{ir}}}{n_{ir}!} \quad (2)$$

$$\mathbf{S}(\mathbf{N}) = \left\{ n_{11},\ldots,n_{MR} : \sum_{i=1}^{M} n_{ir} = N_r, 1 \le r \le R \right\}. \quad (3)$$

Then, the main problem is the computation of $G(\mathbf{N})$. In fact,

$$|\mathbf{S}(\mathbf{N})| = \prod_{r=1}^{R} \binom{N_r + M - 1}{M - 1} \quad (4)$$

What are the solutions?

A number of alternative analyses

Given that

- no (exact) polynomial algorithm is known for the solution of closed BCMP models,

alternative solutions are available:

1. Bounding Analysis
2. Asymptotic Analysis
3. Approximate Analysis
4. Bottleneck Analysis

Outline

Hölder's Inequality in two dimensions

- Let p and q be positive real numbers such that $1/p + 1/q = 1$.
- If $\mathcal{S}$ is a measurable subset on $\Re^n$, and f_1 and f_2 are measurable real-valued functions on $\mathcal{S}$, then

$$\int_{\mathcal{S}} \left| f_1(x) f_2(x) \right| \mathrm{d}x \leq \left(\int_{\mathcal{S}} |f_1(x)|^p \, \mathrm{d}x \right)^{1/p} \left(\int_{\mathcal{S}} |f_2(x)|^q \, \mathrm{d}x \right)^{1/q} . \quad (5)$$

Outline

Our Objective

To apply the Hölder's Inequality (in the general case)

$$\int_{\mathcal{S}} \left| \prod_{r=1}^{R} f_r(x) \right| \mathrm{d}x \leq \prod_{r=1}^{R} \left(\int_{\mathcal{S}} |f_r(x)|^{p_r} \, \mathrm{d}x \right)^{1/p_r} \tag{6}$$

with $\sum_{r=1}^{R} 1/p_r = 1$, to the partition function of closed BCMP queueing networks

$$G(\mathbf{N}) = \sum_{\mathbf{S}(\mathbf{N})} \prod_{i=1}^{M} \left(\sum_{r=1}^{R} n_{ir} \right)! \prod_{r=1}^{R} \frac{D_{ir}^{n_{ir}}}{n_{ir}!} \tag{7}$$

to obtain a computationally efficient bound.

Integral representation of $G(\mathbf{N})$

- It is possible to show that

$$G(\mathbf{N}) = \frac{1}{\prod_{r=1}^{R} N_r!} \int_{\Re^{+M}} \prod_{r=1}^{R} H(r,\mathbf{u})^{N_r} e^{-(u_1+\ldots+u_M)} d\mathbf{u} \tag{8}$$

where $H(r,\mathbf{u}) = \rho_{0r} + \rho_{1r} u_1 + \cdots + \rho_{Mr} u_M$.

- (8) can be rewritten as

$$G(\mathbf{N}) = \frac{1}{\prod_{r=1}^{R} N_r!} \int_{\Re^{+M}} \prod_{r=1}^{R} \left[H(r,\mathbf{u})^{N} e^{-(u_1+\ldots+u_M)} \right]^{\beta_r} d\mathbf{u} \tag{9}$$

where $\beta_r = N_r / N$.

Integral representation of $G(\mathbf{N})$

- It is possible to show that

$$G(\mathbf{N}) = \frac{1}{\prod_{r=1}^{R} N_r!} \int_{\Re^{+M}} \prod_{r=1}^{R} H(r,\mathbf{u})^{N_r} e^{-(u_1+\ldots+u_M)} d\mathbf{u} \quad (8)$$

where $H(r,\mathbf{u}) = \rho_{0r} + \rho_{1r}u_1 + \cdots + \rho_{Mr}u_M$.

- (8) can be rewritten as

$$G(\mathbf{N}) = \frac{1}{\prod_{r=1}^{R} N_r!} \int_{\Re^{+M}} \prod_{r=1}^{R} \left[H(r,\mathbf{u})^{N} e^{-(u_1+\ldots+u_M)}\right]^{\beta_r} d\mathbf{u} \quad (9)$$

where $\beta_r = N_r/N$.

Applying Hölder's result

- Therefore, Hölder's inequality yields

$$G(\mathbf{N}) \leq \prod_{r=1}^{R} \frac{1}{N_r!} \left[\int_{\Re^{+M}} H(r,\mathbf{u})^N e^{-(u_1+\ldots+u_M)} d\mathbf{u} \right]^{\beta_r} \quad (10)$$

- Now, multiplying and dividing by $N!$, we obtain

$$G(\mathbf{N}) \leq N! \prod_{r=1}^{R} \frac{1}{N_r!} \left[\frac{1}{N!} \int_{\Re^{+M}} H(r,\mathbf{u})^N e^{-(u_1+\ldots+u_M)} d\mathbf{u} \right]^{\beta_r} \quad (11)$$

- $\Rightarrow$ the expression in the brackets can be interpreted as the integral representation of the partition function of a singleclass network populated by N class-r jobs only.

Applying Hölder's result

- Therefore, Hölder's inequality yields

$$G(\mathbf{N}) \leq \prod_{r=1}^{R} \frac{1}{N_r!} \left[\int_{\Re^{+M}} H(r,\mathbf{u})^N e^{-(u_1+\ldots+u_M)} d\mathbf{u} \right]^{\beta_r} \quad (10)$$

- Now, multiplying and dividing by $N!$, we obtain

$$G(\mathbf{N}) \leq N! \prod_{r=1}^{R} \frac{1}{N_r!} \left[\frac{1}{N!} \int_{\Re^{+M}} H(r,\mathbf{u})^N e^{-(u_1+\ldots+u_M)} d\mathbf{u} \right]^{\beta_r} \quad (11)$$

- $\Rightarrow$ the expression in the brackets can be interpreted as the integral representation of the partition function of a singleclass network populated by N class-r jobs only.

Applying Hölder's result

- Therefore, Hölder's inequality yields

$$G(\mathbf{N}) \leq \prod_{r=1}^{R} \frac{1}{N_r!} \left[\int_{\Re^{+M}} H(r,\mathbf{u})^N e^{-(u_1+\ldots+u_M)} d\mathbf{u} \right]^{\beta_r} \quad (10)$$

- Now, multiplying and dividing by $N!$, we obtain

$$G(\mathbf{N}) \leq N! \prod_{r=1}^{R} \frac{1}{N_r!} \left[\frac{1}{N!} \int_{\Re^{+M}} H(r,\mathbf{u})^N e^{-(u_1+\ldots+u_M)} d\mathbf{u} \right]^{\beta_r} \quad (11)$$

- $\Rightarrow$ the expression in the brackets can be interpreted as the integral representation of the partition function of a singleclass network populated by N class-r jobs only.

Main Result

- Hence,

$$G(\mathbf{N}) \leq \prod_{r=1}^{R} \frac{[N!G(N\mathbf{e}_r)]^{\beta_r}}{N_r!} = \binom{N}{N_1, \ldots, N_R} \prod_{r=1}^{R} G(N\mathbf{e}_r)^{\beta_r} \quad (12)$$

  where $\mathbf{e}_r$ is the unit vector in direction r.
- $\Rightarrow$ the upper bound on the partition function of a closed, multiclass BCMP queueing network is provided.

Notes on $G(N\mathbf{e}_r)$

$G(N\mathbf{e}_r)$ refers to a single-class queueing network. Therefore,

- it can be computed efficiently
  1. The computational complexity of the MVA is $O(MN)$
  2. The computational complexity of Koenigsberg's formula is $O(M^2)$)
- Koenigsberg's formula provides the asymptotic behavior

$$G(N\mathbf{e}_r) \approx \frac{(\max_{i\geq 1}\rho_{ir})^{N+M-1}}{\prod_{j=1, j\neq \mathrm{argmax}_{i\geq 1}\rho_{ir}}^{M} (\max_{i\geq 1}\rho_{ir} - \rho_{jr})} \quad (13)$$

i.e., the order of magnitude of $G(N\mathbf{e}_r)$.

Notes on $G(N\mathbf{e}_r)$

$G(N\mathbf{e}_r)$ refers to a single-class queueing network. Therefore,

- it can be computed efficiently
  1. The computational complexity of the MVA is $O(MN)$
  2. The computational complexity of Koenigsberg's formula is $O(M^2)$)
- Koenigsberg's formula provides the asymptotic behavior

$$G(N\mathbf{e}_r) \approx \frac{(\max_{i\geq 1} \rho_{ir})^{N+M-1}}{\prod_{j=1, j\neq \text{argmax}_{i\geq 1}\rho_{ir}}^{M} (\max_{i\geq 1} \rho_{ir} - \rho_{jr})} \quad (13)$$

i.e., the order of magnitude of $G(N\mathbf{e}_r)$.

Outline

Extension to LD Stations

- Load-Dependent (LS) station: service rates μ_{ir} are multiplied by queue-lenght functions, i.e., $x_i(n)$ or $y_{ir}(n)$.
- Making the following replacements

$$\rho_{ir} \leftarrow \begin{cases} \rho_{ir} / \inf_n x_i(n) \\ \rho_{ir} / \inf_n y_{ir}(n) \end{cases} \tag{14}$$

  the resulting queueing network is composed of only LI stations and the new value of the partition function bounds from above the original one, and
- Bound (12) can be applied again.

university-logo

Extension to LD Stations

- Load-Dependent (LS) station: service rates μ_{ir} are multiplied by queue-lenght functions, i.e., $x_i(n)$ or $y_{ir}(n)$.
- Making the following replacements

$$\rho_{ir} \leftarrow \begin{cases} \rho_{ir} / \inf_n x_i(n) \\ \rho_{ir} / \inf_n y_{ir}(n) \end{cases} \tag{14}$$

  the resulting queueing network is composed of only LI stations and the new value of the partition function bounds from above the original one, and
- Bound (12) can be applied again.

Summary

- We proposed a new inequality upper bounding the partition function of multiclass, closed BCMP queueing networks in terms of R partition functions related to single-class, closed BCMP queueing networks,
- The upper bound can be computed efficiently,
- Beyond the theoretical interest, it can provide an estimate of the minimum amount of memory that exact solution algorithms implementations should allocate to avoid numerical instabilities.

Summary

- We proposed a new inequality upper bounding the partition function of multiclass, closed BCMP queueing networks in terms of R partition functions related to single-class, closed BCMP queueing networks,
- The upper bound can be computed efficiently,
- Beyond the theoretical interest, it can provide an estimate of the minimum amount of memory that exact solution algorithms implementations should allocate to avoid numerical instabilities.

Summary

- We proposed a new inequality upper bounding the partition function of multiclass, closed BCMP queueing networks in terms of R partition functions related to single-class, closed BCMP queueing networks,
- The upper bound can be computed efficiently,
- Beyond the theoretical interest, it can provide an estimate of the minimum amount of memory that exact solution algorithms implementations should allocate to avoid numerical instabilities.

For Further Reading I

Jonatha Anselmi.
New Analyses in BCMP Queueing Networks Theory.
PhD Thesis, March 2009.
Politecnico di Milano, Italy