MAMSolver: A matrix analytic methods tool*

Alma Riska and Evgenia Smirni

Department of Computer Science
College of William and Mary
Williamsburg, VA 23187-8795, USA
{riska,esmirni}@cs.wm.edu

Abstract. MAMsolver is a software tool for the solution of M/G/1-type, GI/M/1-type, and QBD processes. The collection of solution algorithms implemented by MAMsolver are known as *matrix-analytic methods* and are used to compute stationary measures of interest such as the probability vector, the queue length distribution, the waiting time, the system queue length, and any higher moments of the queue length. The tool also provides probabilistic measures that describe the stability of the queueing system such as the caudal characteristic.

1 Introduction

Matrix analytic techniques, pioneered by Marcel Neuts [7, 8], provide a framework that is widely used for the exact analysis of a general and frequently encountered class of queueing models. In these models, the embedded Markov chains are two-dimensional generalizations of elementary GI/M/1 and M/G/1 queues and their intersection, i.e., quasi-birth-death (QBD) processes. Such processes are often referred to as "processes with repetitive structure", because their state space $\mathcal{S}$ can be partitioned into subsets $\mathcal{S}^{(i)}$ for $i \geq 0$, whose states have identical interactions, except maybe the states in $\mathcal{S}^{(0)}$ known as the "boundary portion of the process" (see Section 2 for more details). Matrix analytic techniques have been successfully used for modeling and analysis of complex computer systems. Representative classes of Markov chains that can be solved using matrix analytic techniques are BMAP/PH/1 and MAP/PH/1 queues which can be used for modeling queueing systems that experience both correlation and high variability, salient characteristics of Internet traffic. MAMSolver is a collection of the most efficient solution methodologies for M/G/1-type, GI/M/1-type, and QBD processes. In contrast to existing tools such as MAGIC [12] and MGMTool [4], which provide solutions only for QBD processes, MAMSolver provides implementations for the classical and the most efficient algorithms that solve M/G/1-type, GI/M/1-type, and QBD processes. The solution provided by MAMSolver consists of the stationary probability vector for the processes under study, the computation of simple measures of interest, like queue length or response time of the queueing system, as well as probabilistic indicators for the processes such as the caudal characteristic [5].

* This work was partially supported by National Science Foundation under grants EIA-9974992, EIA-77030, CCR-0098278, and ACI-0090221.

2 MAMSolver framework

MAMSolver deals with the solution of processes with repetitive structure, namely M/G/1-type, GI/M/1-type, and QBD processes. MAMSolver provides solution for both continuous and discrete time Markov processes but for the sake of brevity we refer only to continuous time Markov processes throughout this exposition. The formal representation of processes under study is given by a block partitioned infinitesimal generator as follows:

$$\underset{M/G/1}{\mathbf{Q}} = \begin{bmatrix} \widehat{\mathbf{L}} & \widehat{\mathbf{F}}^{(1)} & \widehat{\mathbf{F}}^{(2)} & \widehat{\mathbf{F}}^{(3)} & \cdots \\ \widehat{\mathbf{B}} & \mathbf{L} & \mathbf{F}^{(1)} & \mathbf{F}^{(2)} & \cdots \\ \mathbf{0} & \mathbf{B} & \mathbf{L} & \mathbf{F}^{(1)} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{B} & \mathbf{L} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \underset{GI/M/1}{\mathbf{Q}} = \begin{bmatrix} \widehat{\mathbf{L}} & \widehat{\mathbf{F}} & \mathbf{0} & \mathbf{0} & \cdots \\ \widehat{\mathbf{B}}^{(1)} & \mathbf{L} & \mathbf{F} & \mathbf{0} & \cdots \\ \widehat{\mathbf{B}}^{(2)} & \mathbf{B}^{(1)} & \mathbf{L} & \mathbf{F} & \cdots \\ \widehat{\mathbf{B}}^{(3)} & \mathbf{B}^{(2)} & \mathbf{B}^{(1)} & \mathbf{L} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \underset{QBD}{\mathbf{Q}} = \begin{bmatrix} \widehat{\mathbf{L}} & \widehat{\mathbf{F}} & \mathbf{0} & \mathbf{0} & \cdots \\ \widehat{\mathbf{B}} & \mathbf{L} & \mathbf{F} & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{B} & \mathbf{L} & \mathbf{F} & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{B} & \mathbf{L} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

We use the letters "L", "B", and "F" to describe submatrices corresponding to "local", "backward", and "forward" transitions respectively, in relation to states in $\mathcal{S}^{(i)}$ for $i \geq 1$, and we use "$\widehat{\ }$" to describe those matrices related to states in $\mathcal{S}^{(0)}$, i.e., the boundary portion of the process. This block partitioning of the infinitesimal generator defines a partition of the corresponding stationary probability vector in sub-vectors $\boldsymbol{\pi}^{(i)}$ for $i \geq 0$.

2.1 M/G/1-type Processes

For the class of M/G/1-type continuous Markov chains, the infinitesimal generator is of the $\mathbf{Q}_{M/G/1}$ form. Key to all algorithms that solve M/G/1-type processes is the computation of the matrix $\mathbf{G}$, which is given by the solution of the following matrix equation:

$$\mathbf{B} + \mathbf{L}\mathbf{G} + \sum_{j=1}^{\infty} \mathbf{F}^{(j)} \mathbf{G}^{j+1} = \mathbf{0}. \tag{1}$$

Matrix $\mathbf{G}$ has an important probabilistic interpretation: an entry (k, l) in $\mathbf{G}$ expresses the conditional probability of the process first entering $\mathcal{S}^{(j-1)}$ through state l, given that it starts from state k of $\mathcal{S}^{(j)}$[8, page 81]. Recent advances show that the computation of $\mathbf{G}$ is more efficient when displacement structures are used based on the representation of M/G/1-type processes by means of QBD processes [1,5]. The most efficient algorithm for the computation of $\mathbf{G}$ is cyclic-reduction [1]. The calculation of the stationary probability vector is based on the stable recursive Ramaswami's formula [9], which defines the following recursive relation among stationary probability vectors $\boldsymbol{\pi}^{(i)}$ for $i \geq 0$:

$$\boldsymbol{\pi}^{(j)} = -\left(\boldsymbol{\pi}^{(0)} \widehat{\mathbf{S}}^{(j)} + \sum_{k=1}^{j-1} \boldsymbol{\pi}^{(k)} \mathbf{S}^{(j-k)} \right) {\mathbf{S}^{(0)}}^{-1} \quad \forall j \geq 1. \tag{2}$$

Matrices $\widehat{\mathbf{S}}^{(j)}$ and $\mathbf{S}^{(j)}$ are defined as follows:

$$\widehat{\mathbf{S}}^{(j)} = \sum_{l=j}^{\infty} \widehat{\mathbf{F}}^{(l)} \mathbf{G}^{l-j}, \ j \geq 1, \qquad \mathbf{S}^{(j)} = \sum_{l=j}^{\infty} \mathbf{F}^{(l)} \mathbf{G}^{l-j}, \ j \geq 0 \ \text{(letting } \mathbf{F}^{(0)} \equiv \mathbf{L}),$$

Given the above definition of $\boldsymbol{\pi}^{(j)}$ and the normalization condition, a unique vector $\boldsymbol{\pi}^{(0)}$ can be obtained by solving the following system of linear equations:

$$\boldsymbol{\pi}^{(0)} \left[\left(\widehat{\mathbf{L}} - \widehat{\mathbf{S}}^{(1)} {\mathbf{S}^{(0)}}^{-1} \widehat{\mathbf{B}} \right)^{\diamond} \;\middle|\; \mathbf{1}^T - \left(\sum_{j=1}^{\infty} \widehat{\mathbf{S}}^{(j)} \right) \left(\sum_{j=0}^{\infty} \mathbf{S}^{(j)} \right)^{-1} \mathbf{1}^T \right] = [\mathbf{0} \mid 1].$$

where the symbol "$\diamond$" indicates that we discard one (any) column of the corresponding matrix, since we added a column representing the normalization condition[1]. Once $\boldsymbol{\pi}^{(0)}$ is known, one can then iteratively compute $\boldsymbol{\pi}^{(i)}$ for $i \geq 1$, stopping when the accumulated probability mass is close to one. After this point, measures of interest can be computed. [6] gives an improved version of Ramaswami's formula, where the stationary probability vector is computed using matrix-generating functions associated with triangular Toeplitz matrices. These matrix-generating functions are computed efficiently by using fast Fourier transforms (FFT).

2.2 GI/M/1-type Processes

The continuous GI/M/1-type Markov processes, with infinitesimal generator of the $\mathbf{Q}_{GI/M/1}$ form, admit the very-elegant matrix-geometric solutions proposed by Neuts [7]. In matrix-geometric solutions the relation between the stationary probability vectors $\boldsymbol{\pi}^{(i)}$ is:

$$\boldsymbol{\pi}^{(i)} = \boldsymbol{\pi}^{(1)} \cdot \mathbf{R}^{i-1}, \qquad \forall i \geq 1 \tag{3}$$

where $\mathbf{R}$ is the solution of the matrix equation

$$\mathbf{F} + \mathbf{R} \cdot \mathbf{L} + \sum_{k=1}^{\infty} \mathbf{R}^{k+1} \cdot \mathbf{B}^{(k)} = \mathbf{0}. \tag{4}$$

Matrix $\mathbf{R}$ has an important probabilistic interpretation: $\mathbf{R}$ records the expected number of visits to each state in $\mathcal{S}^{(i)}$, starting from each state in $\mathcal{S}^{(i-1)}$, before reentering $\mathcal{S}^{(i-1)}$ [5]. $\mathbf{R}$ can be obtained using iterative methods. Eq.(4) together with the normalization condition are then used to obtain $\boldsymbol{\pi}^{(0)}$ and $\boldsymbol{\pi}^{(1)}$ by solving the following system of linear equations:

$$[\boldsymbol{\pi}^{(0)},\ \boldsymbol{\pi}^{(1)}] \cdot \left[\begin{array}{c|c|c} \widehat{\mathbf{L}}^{\diamond} & \widehat{\mathbf{F}}^{(1)} & \mathbf{1}^T \\ \left(\sum_{k=1}^{\infty} \mathbf{R}^{k-1} \cdot \widehat{\mathbf{B}}^{(k)}\right)^{\diamond} & \mathbf{L} + \sum_{k=1}^{\infty} \mathbf{R}^k \cdot \mathbf{B}^{(k)} & (\mathbf{I} - \mathbf{R})^{-1} \cdot \mathbf{1}^T \end{array} \right] = [\mathbf{0} \mid 1].$$

For $i > 1$, $\boldsymbol{\pi}^{(i)}$ can be obtained numerically from Eq.(3). More importantly, many useful performance metrics, such as the expected number of jobs in the service node, can be computed exactly in explicit form using $\boldsymbol{\pi}^{(0)}$, $\boldsymbol{\pi}^{(1)}$, and $\mathbf{R}$ alone based on the formula $\boldsymbol{\pi}^{(1)}(\mathbf{I} - \mathbf{R})^{-2}$.

[1] $\mathbf{1}^T$ denotes a column vector of ones with the appropriate dimension, while $[\mathbf{0} \mid 1]$ denotes a row vector of zeros except its last element which is one.

2.3 QBD Processes

The intersection of GI/M/1 and M/G/1 type processes results on quasi birth-death processes, whose infinitesimal generator is of the $\mathbf{Q}_{QBD}$ form. QBD processes are most commonly associated with GI/M/1-type processes because they can be solved using the very elegant matrix-geometric approach [7]. In the case of QBD processes, Eq.(4) reduces to the matrix quadratic equation

$$\mathbf{F} + \mathbf{R} \cdot \mathbf{L} + \mathbf{R}^2 \cdot \mathbf{B} = \mathbf{0}. \tag{5}$$

Since QBD is a special case of an M/G/1 process, matrix $\mathbf{G}$ of Eq.(1) is defined and widely used on generating efficient solutions for QBDs, The most efficient solution of Eq.(5) is provided by the logarithmic reduction algorithm [5], which first computes $\mathbf{G}$, and then derives $\mathbf{R}$ based on the relation $\mathbf{R} = -\mathbf{F}(\mathbf{L}+\mathbf{F}\dot{\mathbf{G}})^{-1}$. Vectors $\boldsymbol{\pi}^{(0)}$ and $\boldsymbol{\pi}^{(1)}$ are obtained by solving the following system of linear equations:

$$[\boldsymbol{\pi}^{(0)},\ \boldsymbol{\pi}^{(1)}] \cdot \left[\begin{array}{c|c|c} \widehat{\mathbf{L}}^{\diamond} & \widehat{\mathbf{F}} & \mathbf{1}^T \\ \hline \widehat{\mathbf{B}}^{\diamond} & \mathbf{L} + \mathbf{R} \cdot \mathbf{B} & (\mathbf{I} - \mathbf{R})^{-1} \cdot \mathbf{1}^T \end{array}\right] = [\mathbf{0} \mid 1].$$

Similarly to GI/M/1-type processes, several measures of interest have closed form formulas depending only on $\boldsymbol{\pi}^{(0)}$, $\boldsymbol{\pi}^{(1)}$, and $\mathbf{R}$.

2.4 ETAQA Methodology

[11, 2, 3] propose a methodology to compute only *aggregate* probabilities instead of the whole stationary probability vector for M/G/1 and QBD processes. This aggregate methodology, known as ETAQA is exact, numerically stable and most importantly very efficient with respect to both its time and space complexity. For an ergodic CTMC, it requires only the knowledge of matrix $\mathbf{G}$. The aggregate solution of an M/G/1 process is obtained as the solution of the following system of linear equations:

$$[\boldsymbol{\pi}^{(0)},\ \boldsymbol{\pi}^{(1)},\ \boldsymbol{\pi}^{(*)}] \cdot \left[\begin{array}{c|c|c} \mathbf{1}^T & \widehat{\mathbf{L}} & \widehat{\mathbf{F}}^{(1)} - \sum_{i=3}^{\infty} \widehat{\mathbf{S}}^{(i)} \cdot \mathbf{G} \\ \hline \mathbf{1}^T & \widehat{\mathbf{B}} & \mathbf{L} - \sum_{i=2}^{\infty} \mathbf{S}^{(i)} \cdot \mathbf{G} \\ \hline \mathbf{1}^T & \mathbf{0} & \mathbf{B} - \sum_{i=1}^{\infty} \mathbf{S}^{(i)} \cdot \mathbf{G} \end{array} \begin{array}{c} (\sum_{i=2}^{\infty} \widehat{\mathbf{F}}^{(i)} + \sum_{i=3}^{\infty} \widehat{\mathbf{S}}^{(i)} \cdot \mathbf{G})^{\diamond} \\ (\sum_{i=1}^{\infty} \mathbf{F}^{(i)} + \sum_{i=2}^{\infty} \mathbf{S}^{(i)} \cdot \mathbf{G})^{\diamond} \\ (\sum_{i=1}^{\infty} \mathbf{F}^{(i)} + \mathbf{L} + \sum_{i=1}^{\infty} \mathbf{S}^{(i)} \cdot \mathbf{G})^{\diamond} \end{array}\right] = [1, \mathbf{0}],$$

where $\boldsymbol{\pi}^{(*)}$ is the aggregate probability vector defined as $\boldsymbol{\pi}^{(*)} = \sum_{i=2}^{\infty} \boldsymbol{\pi}^{(i)}$.

Since QBDs are a special case of M/G/1-type processes, they can be solved using this aggregate approach with the following system of linear equations:

$$[\boldsymbol{\pi}^{(0)},\ \boldsymbol{\pi}^{(1)},\ \boldsymbol{\pi}^{(*)}] \cdot \left[\begin{array}{c|c|c} \mathbf{1}^T & \widehat{\mathbf{L}} & \widehat{\mathbf{F}} \\ \hline \mathbf{1}^T & \widehat{\mathbf{B}} & \mathbf{L} \\ \hline \mathbf{1}^T & \mathbf{0} & \mathbf{B} - \mathbf{F} \cdot \mathbf{G} \end{array} \begin{array}{c} \mathbf{0}^{\diamond} \\ \mathbf{F}^{\diamond} \\ (\mathbf{L} + \mathbf{F} + \mathbf{F} \cdot \mathbf{G})^{\diamond} \end{array}\right] = [1, \mathbf{0}],$$

[11, 2, 3] provide algorithms for the computation of measures of interest based on the computed aggregate probability $[\boldsymbol{\pi}^{(0)},\ \boldsymbol{\pi}^{(1)},\ \boldsymbol{\pi}^{(*)}]$.

3 MAMSolver

MAMSolver is a tool that provides implementation of the algorithms highlighted in Section 2. The tool provides solutions for both DTMCs and CTMCs. The matrix-analytic algorithms that we take into consideration are defined in terms of matrices, making matrix manipulations and operations the basic elements of the tool. The input to MAMSolver, in the form of a structured text file, is the finite set of matrices that accurately describe the process to be solved. Since there are different algorithms that provide solution for the same process, the user specifies the method to be used. However several tests are performed within the tool to insure that special cases are treated separately and therefore more efficiently. MAMSolver is implemented in C++, and classes define the basic components of the type of processes under consideration.

Matrix is the module that implements all the basic matrix operations such as matrix assignments, additions, subtractions, multiplications, and inversions. To address possible numerical problems that may arise during matrix operations, we use well known and heavily tested routines provided by the Lapack and BLAS packages[2]. Since solution of system of linear equations is very common for matrix-analytic algorithms, MAMSolver provides several methods depending on the size of the problem, i.e., the size of the coefficient matrices. For small-size problems exact methods such as LU-decomposition are used, otherwise the Lapack implementation of numerical methods like GMRES and BiCGSTAB, are chosen.

Matrix-analytic modules handle both CTMC and DTMC processes. First these modules serve as storage for the input matrices. In addition, these modules provide all the routines necessary for the implementation of the algorithms outlined in Section 2. Both data structures and routines of the matrix-analytic modules are based on the data-structures and routines provided by the matrix module.

The solution of **QBD** processes, requires computation of $\mathbf{R}$ (and sometimes of $\mathbf{G}$ depending on the solution algorithm). Matrix $\mathbf{R}$ is computed using the logarithmic reduction algorithm [5], however for completeness we provide also the classical numerical algorithm. There are cases when $\mathbf{G}$ (and $\mathbf{R}$) can be computed explicitly [10]. We check if the conditions for the explicit computation hold in order to simplify and speedup the solution. The available solution methods for QBD processes are matrix-geometric and ETAQA.

GI/M/1 processes require the computation of matrix $\mathbf{R}$. The classic matrix geometric solution is implemented to solve this type of processes. The algorithm goes first through a classic iterative procedure to compute $\mathbf{R}$ (to our knowledge, there is no alternative more efficient one). Then, it computes the boundary part of the stationary probability vector. Since there exists a geometric relations between vectors $\boldsymbol{\pi}^{(i)}$ for $i \geq 1$, there is no need to compute the whole stationary probability vector.

M/G/1 processes require the computation of matrix $\mathbf{G}$. More effort has been placed on efficient solution of M/G/1 processes. MAMSolver provides the

[2] Available from http://www.netlib.org.

classic iterative algorithm, the cyclic-reduction algorithm, and the explicit one for special cases. The stationary probability vector is computed recursively using either Ramaswami's formula or its fast FFT version. ETAQA is the other alternative for the solution of M/G/1 processes.

MAMsolver computes and stores the stationary probability vector, allowing the user to compute any measures of interest. MAMSolver computes a customized first moment of the form:

$$r = \boldsymbol{\pi}^{(0)} \cdot \boldsymbol{\rho}^{(0)} + \boldsymbol{\pi}^{(1)} \cdot \boldsymbol{\rho}^{(1)} + \sum_{i=2}^{infty} \boldsymbol{\pi}^{(i)} (\mathbf{a}^{[0]} + \mathbf{a}^{[1]} \cdot i).$$

All the coefficient vectors $\boldsymbol{\rho}^{(0)}$, $\boldsymbol{\rho}^{(1)}$, $\mathbf{a}^{[0]}$, $\mathbf{a}^{[1]}$ are to be read from a text file with default values: $\boldsymbol{\rho}^{(0)} = \mathbf{0}$, $\boldsymbol{\rho}^{(1)} = \mathbf{1}$, $\mathbf{a}^{[0]} = \mathbf{0}$, $\mathbf{a}^{[1]} = \mathbf{1}$. Computation of higher moments can be achieved by introducing additional coefficient vectors $\mathbf{a}^{[j]}$, in the above formula.

For lack of space we do not present here any MAMSolver input or output examples. Although at the moment there is no graphical interface for the tool, the input can be constructed easily in text format. The output files are stored in text format as well. For a set of input and output examples and the source code of MAMSolver, we point the interested reader to the tool's website `http://www.cs.wm.edu/MAMSolver/`.

Acknowledgments

We thank Beatrice Meini for providing us with her Fortran implementation of the cyclic reduction algorithm for computation of matrix $\mathbf{G}$ and the fast FFT Ramaswami's formula. We also thank Daniela Puiu for her contribution in earlier versions of the tool.

References

1. D.A. Bini and B. Meini. Using displacement structure for solving non-skip-free M/G/1 type Markov chains. In A. Alfa and S. Chakravarthy Eds. *Advances in Matrix Analytic Methods for Stochastic Models*, pages 17-37, Notable Publications Inc, NJ, 1998.
2. G. Ciardo, A. Riska, and E. Smirni. An aggregation-based solution method for M/G/1-type processes. In B. Plateau, W. J. Stewart, and M. Silva, editors, *Numerical Solution of Markov Chains*, pages 21–40. Prensas Universitarias de Zaragoza, Zaragoza, Spain, Sept. 1999.
3. G. Ciardo and E. Smirni. ETAQA: an efficient technique for the analysis of QBD-processes by aggregation. *Performance Evaluation*, Vol.(36-37), pages 71–93, 1999.
4. B.R. Haverkort, A.P.A. Van Moorsel and A. Dijkstra. MGMtool: A Performance Analysis Tool Based on Matrix Geometric Methods. In R. Pooley, and J. Hillston, editors *Modelling Techniques and Tools*, pages 312–316, Edinburgh University Press, 1993.
5. G. Latouche and V. Ramaswami. *Introduction to Matrix Geometric Methods in Stochastic Modeling*. ASA-SIAM Series on Statistics and Applied Probability, SIAM Press, Philadelphia, PA, 1999.
6. B. Meini. An improved FFT-based version of Ramaswami's formula. *Comm. Statist. Stochastic Models*, Vol. 13 pages 223-238, 1997.
7. M. F. Neuts. *Matrix-geometric solutions in stochastic models*. Johns Hopkins University Press, Baltimore, MD, 1981.
8. M. F. Neuts. *Structured stochastic matrices of M/G/1 type and their applications*. Marcel Dekker, New York, NY, 1989.
9. V. Ramaswami. A stable recursion for the steady state vector in Markov chains of M/G/1 type. *Commun. Statist.- Stochastic Models*, Vol. 4 pages 183-263, 1988.
10. V. Ramaswami and G. Latouche. A general class of Markov processes with explicit matrix-geometric solutions. *OR Spektrum*, 8:209–218, Aug. 1986.
11. A. Riska, and E. Smirni. An exact aggregation approach for M/G/1-type Markov chains. to appear at ACM SIGMETRICS'02 Conference, Marina Del Rey, CA, June 2002.
12. Mark.S.Squillante. MAGIC: A computer performance modeling tool based on matrix-geometric techniques. In Proc. *Fifth International Conference on Modelling Techniques and Tools for Computer Performance Evaluation*, 1991.