An Analysis of Application of An Aggregation-Disaggregation Algorithm for Nearly Completely Decomposable Markov Chains to Hypercube Queueing Model

L. Luque[1], S.V. Carvalho[1]
[1]Laboratory for Computing and Applied Mathematics - LAC
Brazilian National Institute for Space Research - INPE
C. Postal 515 – 12245-970 – São José dos Campos - SP
BRAZIL
E-mail: leandro@lac.inpe.br, solon@lac.inpe.br

Keywords: Markov chain, hypercube queueing model, decomposition, iterative aggregation-disaggregation, nearly completely decomposable.

Markov chain models are used in a variety of applications, including, among others, queueing systems, inventory problems and maintenance models. A Markov model is especially useful in determining the steady state probabilities (long-run behavior) for the modeled system. The standard procedure for obtaining these steady state probabilities involves the solution of a set of linear equations. Unfortunately, when modeling real systems, the state space and hence the number of equations is often very large and, therefore, the computation of the stationary probabilities through traditional methods becomes impractical in terms of computer memory and computation time.

An example of a Markov chain-based model which may present these computational restrictions is the hypercube queueing model, proposed by Larson (1974) and studied by several authors (Brandeau & Larson, 1986; Burwell et al., 1993; Chelst & Jarvis, 1979; Chelst & Barlach, 1981; Gau & Larson, 1988; Halpern, 1977; Jarvis, 1985; Larson, 1975; Larson & Odoni, 1975; Mendonça & Morabito, 2000; Swersey, 1994). The hypercube model is a valuable tool for planning server-to-customer service systems. The computational restrictions due to the fact that the identity of individual service units is preserved on the state space of the model to calculate some performance measures and to allow complex dispatch policies.

For moderately sized models the steady state probabilities can be obtained in a number of ways, such as through Gaussian elimination method or Gauss-Seidel iterative method, but for large systems, the solution becomes impractical through the traditional direct and iterative methods. For systems with homogenous service units (with the same service time), Larson developed an approximate procedure based on a system of N non-linear equations which is computationally much less time and memory-consuming than exact methods, but with low accuracy for unbalanced systems (Luque & Carvalho, 2006).

As in many practical applications (mainly in emergency service systems of big cities), the service units are non-homogeneous and appear in great number and given the lack of efficient methods for the solution of the model, the study of alternative methods of solution for hypercube queueing model becomes important.

In large models, often the structure of the Markov chain can be exploited, in an approximate or exact solution procedure, to reduce the computational load and memory requirements. Some examples of special structures that can be exploited are exactly and weakly lumpable, single input or exit state decomposable, mandatory sets and nearly completely decomposable (Kim & Smith, 1991). The structure of interest in this paper is the near completely decomposable-NCD (another commonly used terminology is *nearly uncoupled* or *nearly separable*). This class of Markov chains is characterized by a clustering of the states into groups, with strong interactions between states in the same group, and weak interactions between states in different groups (Courtois, 1977; Meyer, 1989; Simon & Ando, 1961).

In this paper we will study the application of Koury, McAllister and Stewart-KMS (Stewart, 1994) aggregation-disaggregation algorithm which exploits the near completely decomposable structure of Markov chains to hypercube queueing model. This study will be carried out in two parts. Firstly, the KMS algorithm will be applied to the original hypercube Markov chain. Secondly, a partitioning algorithm (Dayar, 1998; Dayar & Stewart, 2000; Sezer & Šiljak, 1986) will be applied to the original Markov chain before the application of KMS algorithm. The models that will be used in this study will be generated randomly with variations in some parameters, such as number of servers, traffic intensity etc.

REFERENCES

Brandeau, M.L. and Larson, R.C. (1986), *Extending and applying the hypercube model to deploy ambulances in Boston*, in Swersey, A.J. and Ignall, E.J. (Eds) "Delivery of urban service: with a view towards applications in management science and operations research", TIMS studies in the management sciences, 22, North-Holland, Amsterdam, pp. 121–153.

Burwell, T.H., Jarvis, J.P. and Mcknew, M. A. (1993), *Modeling co-located servers and dispatch ties in the hypercube model*, Computers & Operations Research, 20 (2), 113-119.

Chelst, K.R. and Jarvis, J.P. (1979), *Estimating the distribution of travel times in urban emergency service systems*, Operations Research, 27 (1), 199-204.

Chelst, K.R. and Barlach, Z. (1981), *Multiple unit dispatches in emergency services: models to estimate system performance*, Management Science, 27 (12), 1390–1409.

Courtois, P.J. (1977), *Decomposability: queueing and computer system applications*, Academic Press, New York.

Dayar, T. (1998), *Permuting Markov chains to nearly completely decomposable form*, Technical Report BU-CEIS-9808, Department of Computer Engineering and Information Science, Bilkent University, Ankara, Turkey.

Dayar, T. and Stewart, W. J. (2000), *Comparison of Partitioning Techniques for Two-Level Iterative Solvers on Large, Sparse Markov Chains*, SIAM Journal on Scientific Computing, 21 (5), 1691-1705.

Gau, S. and Larson, R. C. (1988), *Hypercube model with multiple-unit dispatches and police patrol-initiated activities*, MIT Operations Research Center Working Paper #OR 188-88.

Halpern, J. (1977), *The accuracy of estimates for the performance criteria in certain emergency service queueing systems*, Transportation Science, 11 (3), 223–242.

Jarvis, J. P. (1985), *Approximating the equilibrium behavior of multi-server loss systems*, Management Science, 31 (2), 235–239.

Kim, D.S. and Smith, R.L. (1991), *An exact aggregation-disaggregation algorithm for mandatory set decomposable Markov chain*, in Numerical Solution of Markov Chains, New York: Marcel Dekker, 89-103.

Larson, R.C. (1974), *A hypercube queuing model for facility location and redistricting in urban emergency services*, Computers & Operations Research, 1 (1), 67-95.

Larson, R.C. (1975), *Approximating the performance of urban emergency service systems*, Operations Research, 23 (5), 845-868.

Larson, R.C. and Odoni, A.R. (1975), *Urban Operations Research*, Prentice-Hall, New Jersey.

Luque, L. and Carvalho, S.V. (2006), *Modelo hipercubo de filas: análise e resultados para o método aproximado de solução*, in XIII Congreso Latino-Iberoamericano de Investigación de Operaciones y Sistemas - CLAIO, Montevidéu, aceito para publicação nos anais do congresso.

Mendonça, F.C. and Morabito, R. (2000), *Aplicação do modelo hipercubo para análise de um sistema médico-emergencial em rodovia*, Gestão & Produção, 7 (1), 73–91.

Meyer, C.D. (1989), *Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems*, SIAM Review, 31 (2), 240-272.

Sezer, M.E. and Šiljak, D.D. (1986), *Nested e-decompositions and clustering of complex systems*, Automatica, 22 (3), 321-331.

Simon, H.A. and Ando, A. (1961), *Aggregation of variables in dynamic systems*, Econometrica, 29 (2), 111-137.

Stewart, W.J. (1994), *Introduction to the numerical solution of Markov chains*, Princeton University Press.