Journal of Network and Computer Applications 36 (2013) 908–926

Contents lists available at SciVerse ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

A hybrid approach to minimize state space explosion problem for the solution of two stage tandem queues

Enver Ever[a,*], Orhan Gemikonakli[a], Altan Kocyigit[b], Eser Gemikonakli[a]

[a] Computer Communications Department, School of Engineering & Information Sciences, Middlesex University, UK
[b] Informatics Institute, Middle East Technical University, Turkey

ARTICLE INFO

Article history:
Received 26 March 2012
Received in revised form
20 August 2012
Accepted 13 October 2012
Available online 8 November 2012

Keywords:
Computer network performance
Fault tolerant systems
Queuing analysis
Performability modeling
Two stage open networks
State space explosion

ABSTRACT

Two stage open queuing networks are used for modeling the subsystem-behaviour in computers and communication networks, mass storage devices, memory servers, and queuing analysis of wireless mobile cellular networks. The queuing analysis of wireless systems is essential in order to quantify the impact of different factors on quality of service (QoS); performance measures so that wireless protocols can be designed and/or tuned in an optimal manner. In that sense two stage open queuing systems are particularly important to model handoff phenomena, especially for the integration of two different systems such as cellular and wireless local area networks (WLANs). Analytical solutions for two-dimensional Markov processes suffer from the state space explosion problem. The numerical difficulties caused by large state spaces, make it difficult to handle multiple servers at the second stage of a tandem queuing system together with server failures and repairs. This study presents a new approach to analytical modeling of open networks offering improvements in alleviating this problem. The proposed solution is a hybrid version, which combines well known spectral expansion, and hierarchical Markov reward rate approaches. Using this approach, two-stage open networks with multiple servers, break-downs, and repairs at the second stage and feedback can be modeled as three-dimensional Markov processes and solved for performability measures. Comparative results show that the new algorithm used for solution, provides a high degree of accuracy, and it is computationally more efficient than the existing approaches. The proposed model is capable of solving other three-dimensional Markov processes.

1. Introduction

Multi-stage service systems are commonly used to model communication networks since they provide the service in series, where customers may depart after receiving the service or join another system for further service. There has been a growing interest in the development of analytical methods for the analysis of queuing network models. Several studies have been published concerning the methodology as well as the application of queuing networks to computer systems modeling. Markov-modulated queues have a wide range of applications in modeling such queuing systems (Carrasco, 2003).

In this paper, spectral expansion technique (Chakka, 1998; Mitrani, 2005) is used together with well-known Markov reward models (Trivedi, 2002) for modeling three-dimensional Markov processes. The solution offered by spectral expansion is used to compute reward rates. Therefore, the computational efficiency of Markov reward models is harnessed with the exact solution offered by spectral expansion for minimizing the state space explosion problem while providing results close to exact solutions. In this sense, the approach can be considered as a hybrid version of both of these well known approaches. The results obtained from the new approach are presented in comparison to the well-known spectral expansion method, iterative method in Gemikonakli et al. (2009) and simulation results. Findings show that the new algorithm presented works for larger state spaces providing high degree of accuracy, and computationally, it is significantly more efficient than the existing approaches. The paper is organized as follows: The next section presents the application areas of two stage systems. Section 3 provides the details of state explosion problem and Section 4 is about the model considered in this study. Section 5 introduces the hybrid solution approach, and Section 6 presents numerical results for the performability measures in order to address the accuracy and the efficacy of the hybrid approach and illustrate the contribution of the new approach. Relevant conclusions are drawn in the final section.

* Corresponding author. Tel.: +44 208411 2696.
E-mail addresses: e.ever@mdx.ac.uk (E. Ever), o.gemikonakli@mdx.ac.uk (O. Gemikonakli), Kocyigit@metu.edu.tr (A. Kocyigit), e.gemikonakli@mdx.ac.uk (E. Gemikonakli).

1084-8045/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jnca.2012.10.006

2. Related work

Analysis of two-stage queuing networks has been of interest to network engineers for many years. In Konheim and Reiser (1976) an analysis of the two-stage tandem networks with Poisson arrivals and exponentially distributed service times in each stage are considered to present models for software of terminal-oriented computer systems, I/O device CPU interaction, Mass-Storage Devices, and Concentrator-Processor Combinations. Queue length distributions are provided for both stages, potential server failures and multi-server systems at the second stage could not be considered since this would introduce a large number of states. An analytical model is presented in Shah et al. (2009) for a system with two queues in tandem in which the first stage has one server and the second stage has multiple servers. The system under study is modeled as M/M/1 and M/M/c queues in tandem and solved by utilizing the matrix-geometric solution approach. However, server failures are not considered for either of the stages. Instead, pure performance measures are presented. In Chakka (1998), and Gemikonakli et al. (2006) two-dimensional Markov chains are used to model similar systems. Each dimension is used for random variables which represent the number of jobs in each stage. In Gemikonakli et al. (2006) such a model is used in order to represent a local area network and a single server system which is connected to this network. However, server failures and multi-server systems at the second stage could not be considered since this would introduce a large number of states. It is possible to divide one of the random variables in order to present the number of jobs in the system for operative and non-operative states of the server, but such an approach rapidly increases the number of states required for each two-dimensional representation. As the number of operative states increases, the existing solution methods start to suffer numerical inconvenience and ill conditioned matrices. It is desirable to develop an approach for two stage tandem systems in order to ease the numerical difficulties caused by large number of states. Such an approach can be used to handle potential failures at the second stage.

As a result of the recent technological breakthrough in the field of wireless systems, it is desirable to maintain the guaranteed QoS to users within the homogeneous and across heterogeneous radio access technologies. The performance and reliability of wireless systems can be modeled by single or multi-dimensional Markov chains (Carrasco, 2003; Trivedi et al., 2002). Two stage queuing systems are used for modeling wireless communication systems as well. They are particularly useful to represent the vertical handoff phenomena of heterogeneous networks. In Song et al. (2008) open network models are used together with multi rate Erlang loss systems to derive the performance metrics of inter-working systems which consist of wireless cellular systems and WLANs. Handover and blocking probabilities are presented without considering potential channel failures. In Weiwei and Lianfeng (2009), cellular network–WLAN interaction is modeled to validate the performance of two new vertical handoff schemes. Multi-dimensional Markov chains are employed for this purpose. Cellular networks and WLANs are modeled independently by using two and three-dimensional Markov chains for corresponding scenarios. Iterative algorithms are used to solve the sets of nonlinear equations introduced. The complexity of the method considered does not allow the iterative solution to incorporate potential channel failures. The state space representation of a two stage network with Markov modulated queues is employed to evaluate the performance of an integrated system consisting of a cellular network and a WLAN (Kirsal et al., 2010). The two stage systems are used to model the cellular network-WLAN interaction. The spectral expansion solution approach is used to solve the system for state probabilities. Since the state space representation of two independent systems is used, horizontal as well as vertical handoffs could be considered together with characteristics such as user mobility, queue capacity and cell structure of cellular network and WLAN considered. Channel failures could not be considered since the exact spectral expansion method employed can suffer from the well known state space explosion problem when such complicated scenarios are considered.

The high expectations of performance and availability for wireless mobile systems present challenges especially for the modeling and design of fault tolerant wireless systems. The channels in a cellular network are subject to failures. In practice, these failures may occur when the base station controller and base repeaters have problematic communication links, or in case a critical function unit fails (Trivedi et al., 2003). In order to analyze the performance degradation caused by failures, performability modeling is the proper methodology (Rupe and Kuo, 1998; Trivedi et al., 2003). Ignoring failures and recovery delays, and considering only the resource contention for wireless systems would cause the overestimation of the systems ability to perform (Reibman, 1990; Rupe and Kuo, 1998; Trivedi et al., 2003). Analytical modeling of cellular systems with hard handoff is presented in Trivedi et al. (2002). Failure and repair of channels are allowed and hence performability analysis of cellular systems is provided. A recursive computation and fixed point iteration method is used to calculate the performance measures for the handoff phenomena. Markov reward modeling approach is employed in order to incorporate channel availability. Similarly in Trivedi et al. (2003) reward rates are specified as the loss rates for Markov reward rate approach and Erlang B loss formula is extended in order to obtain a continuous time Markov chain to cover handoffs. The traditional two-level performability model presented combines these reward rates with an availability model and provides corresponding loss formulas. The closed form solutions obtained in Trivedi et al. (2002) and Trivedi et al. (2003) are useful only for homogeneous systems since the vertical handover phenomenon is not considered. A performability model is presented for heterogeneous wireless communication systems in Gowrishankar et al. (2008). Markov chains are presented for performance and availability of the systems, Stochastic Activity Network and stochastic petri nets are employed and performability measures such as blocking and dropping probabilities are presented. The interaction between two stages, such as the amount of calls interchanged between two systems are not taken into account since two systems have not been considered as separate queuing systems in interaction.

3. State space explosion

Markov-modulated queuing systems can be described as two-dimensional processes where transitions are only possible between adjacent states of a given model. A subset of these Markov processes are the Quasi-Birth-and-Death processes. The popularity of two-dimensional processes resulted in the development of various numerical procedures for their steady-state analysis (Akar and Sohraby, 1997; Bini and Mein, 1996; Chakka, 1998; Ciardo and Smirni, 1999; Gemikonakli et al., 2009; Mitrani, 2005; Trivedi, 2002). These approaches have received some attention and the major ones have been reviewed in Haverkort and Ost (1997), and Mitrani and Chakka (1995), where Haverkort, and Mitrani compare matrix-geometric and spectral expansion approaches. Also in Ost (2001) matrix-geometric solution methods are considered as well as transform methods, non-skip free QBDs and extensions on QBDs. Review efforts highlight strengths as well as weaknesses of numerical techniques used. The functional equations arising in the analysis of such processes usually