ELSEVIER

Computer Networks and ISDN Systems 29 (1998) 1999–2019

COMPUTER NETWORKS and ISDN SYSTEMS

Fast time-dependent evaluation of multi-service networks [1]

Ibrahim Matta [a,*], A. Udaya Shankar [b]

[a] College of Computer Science, Northeastern University, Boston, MA 02115, USA
[b] Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA

Abstract

We present a numerical-analytical method to evaluate multi-service networks with adaptive routing, scheduling and admission controls. We apply our method to connection-oriented networks supporting different types of real-time connections. The network dynamics is described by difference equations which can be solved for both transient and steady-state performances. Results indicate that our method is computationally much cheaper than discrete-event simulation, and yields accurate performance measures. Connection routing algorithms are usually evaluated individually in terms of steady-state performance measures. In this paper, we also use our time-dependent evaluation method to compare several connection routing schemes in terms of *instantaneous* measures. Our results show that a routing scheme which defines the cost of a path as the sum of measured link utilizations yields more stable behavior and lower connection blocking probability over a wide range of workload parameters and network configurations than other traditional schemes. © 1998 Elsevier Science B.V.

Keywords: Multi-service networks; Transient performance; Dynamic flow models; Queueing models; Resource allocation algorithms; Routing

1. Introduction

Multi-service packet-switched networks, such as ATM (Asynchronous Transfer Mode) networks [49], are expected to support a wide variety of applications (e.g., multimedia, voice, mail) with heterogeneous quality-of-service (QoS) requirements. To meet these requirements, new algorithms have been proposed for controlling routing, admission, and scheduling. *Routing* provides a selection of routes, based on cost functions associated with the transmission links. *Admission* defines the criteria used to accept or reject a new incoming application, based on the service requested and the resources available. *Scheduling* defines how link resources (bandwidth, buffers, etc.) are allocated among the different services.

The overall end-to-end performance of the network hinges on the algorithms used in the routing, admission, and scheduling components. The algorithms are often *adaptive*, with parameters being

* Corresponding author. Fax: +1-617-373-5121; e-mail: matta@ccs.neu.edu.

[1] This work was supported in part by NSF CAREER grant NCR-9701988, NEU grant RSDF-377090, and by ARPA under contract DABT63-96-C-0075 administered by Army ARO to Computer Science Department, University of Maryland. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of NSF, ARPA, ARO, or the U.S. Government.

PII S0169-7552(97)00117-7

varied dynamically according to service class and current or delayed system state information. Arrival and service statistics are often time-dependent. As a result, there is significant interaction among the three components.

The accurate and fast evaluation of such time-dependent systems is critical to their cost-effective design. Existing evaluation methods for these systems are inadequate. *Analytical methods are typically too coarse*. They usually assume steady-state conditions and do not account for adaptive policies and the effect of delayed feedback. Incorporating adaptive time-dependent behavior makes them analytically intractable and computationally expensive to solve numerically due to the large state space. *Simulation approaches are often too expensive*. They can handle realistic detail and dynamic situations, but they are invariably computationally prohibitive, especially for evaluating high-speed networks where the number of scheduled events (packets, connections, etc.) is usually enormous.

1.1. Our contribution

In this paper, we present a *numerical-analytical* method that yields the time-evolution of *instantaneous* performance measures. Our method takes into account the interaction and time-dependent nature of the control algorithms, and is computationally inexpensive. The numerical foundation of our method provides a modeling power close to that of simulation at a fraction of the computation expense, typically less expensive than simulation by many orders.

Our method integrates techniques from several areas, namely: standard queueing theory steady-state techniques [37]; the link decomposition technique widely used for packet delay analysis in packet-switched data networks [36] as well as for call blocking probability analysis in circuit-switched telephone networks [21]; the dynamic flow technique used for approximating system dynamics and nonlinearity [18]; and the technique of repeated substitutions used in numerical analysis to solve nonlinear equations [35]. This integrated evaluation approach allows for handling the *transient behavior of general multi-class systems*. Few studies (e.g., [52]) have considered only non-blocking multi-class systems, others (e.g., [19]) have considered blocking single-class systems, and many studies have considered only steady-state behaviors. A detailed discussion of related work is in Section 9.

We apply our method to a model that permits the evaluation of a connection-oriented packet-switched network (e.g., ATM) that supports real-time communication (voice, video, etc.) by making use of various adaptive routing, scheduling and admission policies. Thus this model can be applied to achieve more comprehensive evaluation of existing strategies and to propose more effective network control schemes.

Among the main performance measures of interest are the *instantaneous end-to-end connection blocking probabilities*. To calculate them, we use the link decomposition technique [36,21] to approximate the multi-link network as a collection of single-link networks. For each link, we approximate the relationship between the instantaneous *local* (or link-level) connection blocking probabilities and the instantaneous average numbers of established connections by the relationship at steady-state. The latter is available, usually in implicit form, from standard queueing theory [37]. We solve these instantaneous relationships iteratively [35]. After all single-link models have converged, we compute the instantaneous end-to-end connection blocking probabilities by invoking the link independence assumption.

To obtain the time behavior of the instantaneous end-to-end connection blocking probabilities, we introduce difference equations in the average numbers of established connections. These difference equations relate the instantaneous flow rates of departure and admission of connections. They can be solved iteratively in conjunction with the previous solution (in previous paragraph).

This allows the investigation of both transient and steady-state performances of various control schemes. We point out that our iterative procedure differs from iterations commonly used in steady-state analysis (e.g., [31,33,38,14,50,10,43,23]), which only solve for steady-state measures. Our results indicate that our method is computationally much cheaper than discrete-event simulation, which requires the averaging of a large number of independent simulation runs (replications) to obtain reliable performance estimates. Furthermore, the performance measures it yields are very close to the exact values obtained by simulation.