IMPROVING THE EFFICIENCY OF THE ANALYSIS OF DSPN MODELS

M. Ajmone Marsan
Dipartimento di Scienze della Informazione
Università di Milano – Italy

G. Chiola
Dipartimento di Informatica
Università di Torino – Italy

A. Fumagalli
Dipartimento di Elettronica
Politecnico di Torino – Italy

Abstract

The applicability of DSPN models has been limited by the computational complexity of the algorithm for the evaluation of the steady state probability distribution over reachable markings, so that it was often necessary to resort to simulation rather than analysis. Two techniques for the improvement of the efficiency of the analysis of DSPN are outlined in this paper, using a previously published model of a high speed local area network as an example of the application of the proposed techniques, and as a benchmark for the assessment of their efficiency.

KEY WORDS - Stochastic Petri nets, Performance evaluation, Markov chains.

CONTENTS

1 INTRODUCTION

DSPN [1,2] are timed transition Petri nets (TTPN) in which three classes of transitions are defined. *Exponential transitions* are associated with an exponentially distributed random firing delay, like in basic stochastic Petri nets (SPN) [3-6]. *Immediate transitions* fire as soon as they become enabled, with priority over transitions of other classes, like in generalized SPN (GSPN) [7,8]. *Deterministic transitions* fire after a constant enabling time.

Markings enabling immediate transitions are named *vanishing*, since the model spends no time in them, and they can be removed in the analysis process. The other markings are called *tangible*, and they enable only timed (exponential or deterministic) transitions.

For a model to be classified as a DSPN, and hence be solvable using the approach described in [1,2], it is necessary that no more than one deterministic transition be enabled in any reachable tangible marking.

DSPN were proposed, like other classes of TTPN [6,9-11], for the purpose of providing the user with a greater flexibility in the model specification than it is possible with SPN and GSPN. Indeed, these two classes of TTPN limit the user to the specification of delays that are exponentially distributed random variables. In some cases this may be a reasonable assumption, but it often happens that some temporal characteristics of a system should be better modeled using either random variables with different distributions, or constant delays. The increased modeling power of DSPN and similar models is however paid with a higher complexity in the computation of the model solution, which often becomes prohibitive, thus forcing users to simulate rather than analyze their models.

In fact, the complexity issue is a problem that already plagues simpler classes of models, such as SPN and GSPN. In those cases the complexity derives from the necessity of solving large systems of linear equations in order to obtain the steady-state probability distribution over relevant markings. For SPN and GSPN, however, the generation of the equations is fairly simple, and it can be achieved with little extra cost during the generation of the reachability graph, due to the isomorphism existing between those models and continuous-time Markov chains. In the case of DSPN, the solution is based on the more complex theory of semi-Markov processes, and the generation of the equations needed for the solution often requires a large number of matrix exponentiations. As a conclusion, only toy examples can be handled with the standard approach to the DSPN solution.

The goal of this paper is to show, mainly through an example, that there are cases in which the complexity of the analysis of a DSPN model can be enormously reduced. The key tool on which the reduction of the complexity of the solution is based is the structural analysis of the model.

Two approaches are considered in the paper. The first one is based on a folding of the model before performing the analysis, while the second one is based on the subdivision of the model in smaller subnets that can be analyzed in isolation, and whose results can then be combined to obtain the global result.

The example that we use in the paper is taken from [12], and models a high speed local area network (LAN). For the sake of conciseness, we do not describe the system behaviour in much detail, referring the reader to [12,13], and to the references therein for a deeper discussion.

2 THE DSPN MODEL

The DSPN model that we use as an example in this paper is shown in Figure 1. Immediate transitions are depicted as thin bars, whereas exponential transitions are depicted as thick white bars, and deterministic transitions as thick black bars.

The DSPN in Figure 1 models a LAN where three stations generate data packets, which can be orderly transmitted on a broadcast communication channel according to the EXPRESSNET protocol [13]. This is a virtual token multiple access protocol, i.e., a protocol in which the recognition of special events observed on the channel provides stations with transmission permits (virtual tokens). Stations are orderly visited by the virtual token, and they are allowed to transmit one packet at a time, if they have something ready for transmission. After all stations have been visited by the virtual token, the observation of the channel allows the start of a new transmission cycle after a fixed delay.

The model of the LAN operating according to this protocol is constructed by avoiding the detailed description of the individual station operations, as defined by the EXPRESSNET specification. Rather, the sequences of idle and transmission times on the channel are represented. This approach allows the reduction of the graphical complexity of the model, and greatly reduces the number of reachable markings. The model that we obtain is adequate for the computation of the LAN performance, but it cannot be used for other purposes, such as the protocol validation.

The three identical subnets containing places Si_on, and Si_ready, and transition Si_msg with $i = 1, 2, 3$, model active stations. Transition Si_msg describes the generation of a packet to be transmitted on the