Theoretical Computer Science 669 (2017) 1–21

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Symbolic optimal expected time reachability computation and controller synthesis for probabilistic timed automata

Aleksandra Jovanović [a], Marta Kwiatkowska [a,*], Gethin Norman [b], Quentin Peyras [c]

[a] *Department of Computer Science, University of Oxford, Oxford, UK*
[b] *School of Computing Science, University of Glasgow, Glasgow, UK*
[c] *Département Informatique de l'ENS Cachan, Université Paris-Saclay, France*

ARTICLE INFO

Article history:
Received 21 April 2016
Received in revised form 21 December 2016
Accepted 16 January 2017
Available online 7 February 2017
Communicated by P.A. Abdulla

Keywords:
Probabilistic timed automata
Controller synthesis
Probabilistic verification
Symbolic model checking

ABSTRACT

In this paper we consider the problem of computing the optimal (minimum or maximum) expected time to reach a target and the synthesis of an optimal controller for a probabilistic timed automaton (PTA). Although this problem admits solutions that employ the digital clocks abstraction or statistical model checking, symbolic methods based on zones and priced zones fail due to the difficulty of incorporating probabilistic branching in the context of dense time. We work in a generalisation of the setting introduced by Asarin and Maler for the corresponding problem for timed automata, where *simple* and *nice* functions are introduced to ensure finiteness of the dense-time representation. We find restrictions sufficient for value iteration to converge to the optimal expected time on the uncountable Markov decision process representing the semantics of a PTA. We formulate Bellman operators on the backwards zone graph of a PTA and prove that value iteration using these operators equals that computed over the PTA's semantics. This enables us to extract an ε-optimal controller from value iteration in the standard way.

1. Introduction

Systems which exhibit real-time, probabilistic and nondeterministic behaviour are widespread and ubiquitous in many areas such as medicine, telecommunications, robotics and transport. Timing constraints are often vital to the correctness of embedded devices and stochasticity is needed due to unreliable channels, randomisations and component failure. Finally, nondeterminism is an important concept which allows us to model and analyse systems operating in a distributed environment and/or exhibiting concurrency. A natural model for such systems, *probabilistic timed automata* (PTAs), a probabilistic extension of timed automata (TAs) [1], was proposed in [2–4]. They are finite-state automata equipped with real-valued clocks which measure the passage of time and whose transitions are probabilistic. More specifically, transitions are expressed as discrete probability distributions over the set of edges, each such edge specifying a successor location and a set of clocks to reset.

An important class of properties on PTAs are *probabilistic reachability* properties. They allow us to check statements such as: "with probability 0.05 or less the system aborts" or "the data packet will be delivered within 1 second with minimum 0.95 probability". Model checking algorithms for these properties are well studied. Forwards reachability [3] yields only

[*] Corresponding author.
E-mail address: marta.kwiatkowska@cs.ox.ac.uk (M. Kwiatkowska).

http://dx.doi.org/10.1016/j.tcs.2017.01.015
0304-3975/© 2017 Elsevier B.V. All rights reserved.

approximate probability values (upper bounds on maximum reachability probabilities). An abstraction refinement method, based on stochastic games, has subsequently been proposed in [5] for the computation of exact values and implemented in PRISM [6]. An alternative method is backward reachability [7], also giving exact values. These are all symbolic algorithms based on *zones*, a structure that represents in a concise way sets of the automaton states with equivalent behaviour.

Another important class of properties, which is the focus of this paper, is *expected reachability*. They can express statements such as “the expected number of packets sent before failure is at least 100” or “the expected time until a message is delivered is at most 20 ms”. These properties turned out to be more difficult to verify on PTAs and currently no symbolic approach exists. Even for TAs, the research first concentrated on checking whether there exist system behaviours that satisfy a certain property (for example, reaching the target set of states). In many situations this is not sufficient, as we often want to distinguish between behaviours that reach target states in 10 or 1000 seconds. In [8], a backward fixed-point algorithm was proposed for controller synthesis for TAs, which generates a controller that reaches the target in minimum time. The analogous problem for priced timed automata, a model comprising more general reward (or cost) structures, was also considered. The minimum reward reachability for this model has been solved using the region graph method [9], and later extended for more efficient *priced zones* [10] and implemented in Uppaal [11].

Contributions. We propose the first zone-based algorithms to compute the optimal expected time to reach a target set in a PTA and synthesise an ε-optimal controller. The semantics of a PTA is an uncountable Markov decision process (MDP). Under suitable restrictions, we are able to prove that value iteration converges to the optimal expected time on this MDP. We formulate Bellman operators on the backwards zone graph of a PTA and show that value iteration using these operators yields the same values as those computed on the MDP. This enables us to extract an ε-optimal controller from value iteration in the standard way. This problem has been open for several years, with previous symbolic zone-based methods, including priced zones, being unsuitable for computing expected values since accumulated rewards are *unbounded*. In order to represent the value functions we introduce rational simple and rational nice functions, a generalisation of Asarin and Maler’s classes of simple and nice functions [8].

Related work. Expected reachability properties of PTAs can be verified using the *digital clocks* method [12], which assumes an integral model of time as opposed to a dense model of time. Although this method suffers from state-space explosion, it has been shown useful in practice for the analysis of a number of real-world protocols, see for example [12,13]. In addition, this approach has recently been extended to allow the analysis of partially observable PTAs against expected reachability properties [14]. In [15], the minimum expected reward for priced timed games has been solved using *statistical model checking* and Uppaal-SMC [16]. This is orthogonal to numerical model checking, and is based on simulation and hypothesis testing, thus giving only approximate results which are not guaranteed to be correct.

In [17] the authors consider priced probabilistic timed automata and study reward-bounded probabilistic reachability, which determines whether the maximal probability to reach a set of target locations, within given bounds on the accumulated reward and elapsed time, exceeds a threshold. Although this problem is shown to be undecidable [18], a semi-decidable backwards algorithm using priced zones, which terminates if the problem is affirmative, is implemented in Fortuna [19].

Outline. In Section 2 we define MDPs and give existing results concerning optimal reward computation for uncountable MDPs. Section 3 defines PTAs and introduces the assumptions needed for the adoption of the results of Section 2. In Section 4, we present our algorithms for computing optimal expected time reachability and synthesis of an ε-optimal controller using the backwards zone graph of a PTA. Section 4 also introduces a representation of the value functions that generalise the simple and nice functions of [8] and gives an example demonstrating the approach. We conclude with Section 5.

A preliminary conference version of this paper was published as [20], where only minimum expected time was considered.

2. Background

Let $\mathbb{R}$ be the set of reals, $\mathbb{R}_+$ the set of non-negative reals, $\mathbb{N}$ the natural numbers (including 0), $\mathbb{Q}$ the rationals and $\mathbb{Q}_+$ the non-negative rationals. A discrete probability distribution over a (possibly uncountable) set S is a function $\mu : S{\rightarrow}[0,1]$ such that $\sum_{s\in S}\mu(s)=1$ and the set $\{s\in S \mid \mu(s){>}0\}$ is finite. We denote by $\mathsf{Dist}(S)$ the set of distributions over S. A distribution $\mu\in\mathsf{Dist}(S)$ is a point distribution if there exists $s\in S$ such that $\mu(s){=}1$.

Markov Decision Processes (MDPs) is a widely used formalism for modelling systems which exhibit both nondeterministic and probabilistic behaviour.

Definition 1. An MDP is a tuple $\mathcal{M}{=}(S, s_0, A, P_{\mathcal{M}}, R_{\mathcal{M}})$, where:

- S is a (possibly uncountable) set of states;
- $s_0\in S$ is an initial state;
- A is a (possibly uncountable) set of actions;