Reachability Probabilities in Markovian Timed Automata

Taolue Chen, Tingting Han, Joost-Pieter Katoen, and Alexandru Mereacre

Abstract—We propose a novel stochastic extension of timed automata, i.e., Markovian Timed Automata. We study the problem of optimizing time-bounded reachability probabilities in this model, i.e., the maximum likelihood to hit a set of goal locations within a given deadline. We propose Bellman equations to characterize the probability, and provide two approaches to solve the Bellman equations, namely, a discretization and a reduction to Hamilton-Jacobi-Bellman equations.

I. INTRODUCTION

This paper introduces Markovian Timed Automata (MTA), a novel extension of timed automata [1] with exponentially distributed location residence times. To give some motivation, let’s look at the following example: A robot moves on a $3\times3$-grid (Fig. 1), starting from $A$ and trying to reach $B$ in $T_2$ units of time. At each cell, it can move up ($u$), down ($d$), left ($l$) and right ($r$) (when applicable). Cells are associated with rates, which intuitively represent the speed of the robot. As soon as it moves to a cell, the robot decides a direction to move to the next cell, and then waits in the current cell for a certain amount of time, which is governed by an exponential distribution with a given rate $\lambda$, i.e., the probability of leaving the cell within time $t$ is $1-e^{-\lambda t}$. The robot is allowed to stay in consecutive dark cells for at most $T_1$ units of time, while there is no time constraint for the bright cells. At each cell, the robot has a probability $0.1$ to break down, apart from moving to the neighboring one. Since there are different ways to reach $B$, each of which is associated with some probability, one natural question is: What’s the maximum probability to reach $B$ from $A$ within $T_2$ units of time? This problem can be readily formulated as a controller synthesis problem for MTA depicted in Fig. 2, where each location in $\{l_0, \ldots, l_8\}$ corresponds to a cell and $x, y$ are clocks to specify the time constraints. (N.B. the failure location and all the transitions leading to it are omitted for the clearer illustration. For the same reason the representation of the MTA here is a bit different from the one in Fig. 3.)

This work is partially supported by the DFG research training group 1295 AlgoSyn, the ERC Advanced Grant VERIWARE, and the FP7 Project MoVeS.

T. Chen, T. Han and A. Mereacre are with the Department of Computer Science, Oxford University. Email: \{tchen, than, mereacre\}@cs.ox.ac.uk.

J.-P. Katoen is with the Software Modelling and Verification Group, RWTH Aachen University. Email: katoen@cs.rwth-aachen.de.

1This seems to be restrictive. However, we note that in practice, one can approximate any distribution by phase-type distributions, resulting in series-parallel combinations of exponential distributions [13].

2The minimal one can be treated in a completely dual fashion. We omitted it for the sake of simplicity.
undecidable in general (for 2^\frac{1}{2}-player games), while the qualitative question “= 0” or “= 1” can be solved in PTIME for 1^\frac{1}{2}-player games with a single clock. MTA are essentially 1\frac{1}{2}-player stochastic timed games. However, our focus is on quantitative analysis rather than on qualitative analysis or decidability issues. In [6], a game extension of semi-Markov decision process (SMDP) was considered and the winning objective was specified by a deterministic timed automaton. Again, only the qualitative question was addressed.

As a by-product of our work we obtain two procedures to compute maximum time-bounded reachability probabilities in locally uniform CTMDPs. This problem has also been treated in [2][5][15], where [2][5] mainly addressed time-abstract schedulers which are not necessarily optimal, while here time-dependent schedulers are considered, as well as [15]. Comparing to [15], we discretize both the time and the state space. Moreover a system of PDEs is derived in order to characterize the maximum reachability probability. The error bound that we obtain improves the error bound given in [15] yielding a substantial reduction in the required number of iterations.

Due to space restriction, all proofs and some further explanation are omitted here. We refer the reader to the technical report [9] for the full details.

II. MARKOVIAN TIMED AUTOMATA

Given a set H, let Pr : F(H) → [0, 1] be a probability measure on the measurable space (H, F(H)), where F(H) is a σ-algebra over H. Let Dist(H) denote the set of probability measures on this measurable space.

A. Markov Decision Processes

Definition 1: [MDP] A (continuous-state) Markov decision process is a tuple \( D = (Act, S, s_0, P) \) where

- Act is a denumerable set of actions;
- S is a set of states;
- \( s_0 \in S \) is the initial state;
- \( P : S \times Act \times F(S) \rightarrow [0, 1] \) is the transition probability function, where \( P(s, \alpha, \cdot) \) is a probability measure over \( F(S) \) for any \( s \in S \) and \( \alpha \in Act \), such that \( P(\cdot, \cdot, A) \) is measurable for any \( A \in F(S) \).

The measure \( P(s, \alpha, A) \) is the one-step transition probability from state \( s \in S \) to the set of states \( A \in F(S) \) by taking action \( \alpha \in Act \). Notice that in general one can extend the MDP model to uncountably many actions (see [16]). In this paper we will consider only MDPs which have finitely many actions, i.e., finitely-branching MDPs.

B. Markovian Timed Automata

Let \( X = \{ x_1, ..., x_n \} \) be a set of nonnegative variables in \( \mathbb{R} \), called clocks. A clock-vlaue is a function \( \eta : X \rightarrow \mathbb{R}_{\geq 0} \) assigning to each variable \( x \) a value \( \eta(x) \). Let \( \mathcal{V}(X) \) denote the set of all clock-valuations over \( X \). A clock constraint on \( X \), denoted by \( \triangleright \), is a conjunction of expressions of the form \( x \triangleright c \) for clock \( x \in X \), comparison operator \( \triangleright \in \{ <, \leq, >, \geq \} \) and \( c \in \mathbb{N} \). Let \( B(X) \) denote the set of clock constraints over \( X \). A clock valuation \( \eta \) satisfies constraint \( x \triangleright c \), denoted \( \eta \models x \triangleright c \), if and only if \( \eta(x) \triangleright c \); it satisfies a conjunction of such expressions if and only if \( \eta \) satisfies all of them. Let \( \partial \) denote the valuation that assigns 0 to all clocks. For a subset \( X \subseteq \mathcal{X} \), the reset of \( X \), denoted \( \eta[X := 0] \), is the valuation \( \eta' \) such that \( \forall x \in X \), \( \eta'(x) := 0 \) and \( \forall x \notin X \), \( \eta'(x) := \eta(x) \). For \( \delta \in \mathbb{R}_{\geq 0} \) and \( X \)-valuation \( \eta \), \( \eta + \delta \) is the \( X \)-valuation \( \eta'' \) such that \( \forall x \in X \), \( \eta''(x) := \eta(x) + \delta \), which implies that all clocks proceed at the same speed.

Definition 2: [MTA] A Markovian timed automaton is a tuple \( M = (Act, X, Loc, \ell_0, E, \leadsto) \), where

- Act is a finite set of actions;
- \( X \) is a finite set of clocks;
- Loc is a finite set of locations;
- \( \ell_0 \in Loc \) is the initial location;
- \( E : Loc \rightarrow \mathbb{R}_{>0} \) is the exit rate function and
- \( \leadsto \subseteq Loc \times Act \times B(X) \times Dist(2^X \times Loc) \) is the edge relation.

For simplicity we abbreviate \( (\ell, \alpha, \eta, \zeta) \leadsto \) by \( \ell \xrightarrow{\alpha, \eta} \zeta \), where \( \zeta \) is a probability distribution over \( 2^X \times Loc \). Here we don’t include location invariants, as in [1], and we don’t require edge relation \( \leadsto \) to be total, e.g., there might be some clock constraints \( \eta \) for which \( \leadsto \) is not defined.

Example 1: An example MTA is shown in Fig. 3, where there are 6 locations with \( \ell_0 \), the initial location. In \( \ell_0 \) (resp. \( \ell_2 \)), there is a decision to be made between actions \( \alpha_1 \) and \( \alpha_2 \) (resp. \( \gamma_1 \) and \( \gamma_2 \)) when the clock valuation is \( \eta(x) \in [0, 1] \) (resp. \( \eta(x) \in (1, 2) \)). \( \ell_0 \) is the goal location, which will be used later. For edge \( \ell_0 \xrightarrow{\alpha_1} \zeta_0 \), \( \zeta_0(\emptyset, \ell_2) = 0.2 \) and \( \zeta_0(\emptyset, \ell_5) = 0.8 \).

![Fig. 3. An example MTA](image)

a) Semantics: Intuitively, an MTA behaves as follows. Consider location \( \ell_1 \) in Fig. 3. As soon as location \( \ell_1 \) is entered with clock valuation \( \eta \), action \( \beta \) is chosen and a waiting time \( \tau \) in \( \ell_1 \) is sampled from the probability distribution \( E(\ell_1)e^{-E(\ell_1)\tau} \). If \( \eta + \tau \models x < 1 \), there will be a jump to location \( \ell_0 \) with probability 0.5 or to location \( \ell_4 \) with probability 0.5, otherwise no jump occurs and MTA remains in location \( \ell_1 \). For instance the next location \( \ell_0 \) is entered with clock valuation \( \eta' := \eta + \tau \) (in this example), action \( \alpha_1 \) or \( \alpha_2 \) is chosen and the waiting time \( \tau' \) in \( \ell_0 \) is sampled from the probability distribution \( E(\ell_0)e^{-E(\ell_0)\tau'} \). Suppose \( \alpha_2 \) is picked and \( \eta' + \tau' \models x \geq 0 \) (the guard of action \( \alpha_2 \)), the MTA jumps to the next location according to the probability distribution associated with \( \alpha_2 \). The MTA follows the same behavior continuously.

The semantics of an MTA with clock set \( X \) is given as a continuous-state MDP, where states \( s \) are of the form \( (\ell, \eta) \) where \( \ell \in Loc \) and \( \eta \in \mathcal{V}(X) \) is a clock valuation.
Definition 3: [Semantics] Let $\mathcal{M} = (Act, X, Loc, \ell_0, E, \sim)$ be an MTA. The MDP associated with $\mathcal{M}$ is $D(\mathcal{M}) = (Act, S, \pi_0, P)$ where $S = Loc \times V(X); \pi_0=(\ell_0, 0)$; and

- for each edge $\ell, \eta, A \in M$ with $\eta(X, \ell, \tau) = p > 0$,

$$\mathcal{P}(\ell, \eta, A) := \int_{\eta > 0} E(\ell)e^{-E(\ell)\tau} \cdot 1_g(\eta + \tau) \cdot p \, d\tau,$$

where $A = \{\ell, \eta, p\} | \exists \tau \in R_\geq 0$. $\eta = (\eta + \tau)[X := 0]$ and $\eta + \tau = g$, and $1_g(\cdot)$ is the characteristic function, i.e., $1_g(\eta + \tau) = 1$ if $\eta + \tau = g$; otherwise, 0.

We emphasize that MTA is a Markovian model with decisions (so it is nondeterministic) instead of a pure stochastic model like deterministic MTA (DMTA) studied in [8]. To state it alternatively, any DMTA coincides with an MTA with $Act = \{\alpha\}$. Moreover, for DMTA, the edge relation is defined as $\sim \subseteq Loc \times B(X) \times R^{+} \times Dist(Loc)$, while the MTA model allows for each set of transitions to reset their clocks differently. This has also been used in probabilistic timed automata (PTA, [12]). In this sense, our model can be considered as a continuous-time extension of PTA, due to the presence of exponential distributions. Any MTA where the exit rate of any location is zero is a PTA. Locally uniform continuous-time MDPs (CTMDPs) [14] (the exit rate of each location does not depend on actions) with finite state space are zero-clock MMDPs (i.e., $X = \varnothing$). We note that as in MDPs, it is assumed that action labels from any location in MTA are pairwise different.

Finite paths in MTA $\mathcal{M}$ are of the form $\ell_0 \sim_0 \alpha_0.\ell_1 \sim_1 \alpha_1.\ell_2 \sim_2 \ldots \sim_{n-1} \alpha_{n-1}.\ell_n \sim_n \ell_n$, where for each edge $\ell \sim \zeta_i$ of $\mathcal{M}$ with $\zeta(X, \ell, \tau) > 0$ ($\zeta \in Loc, \alpha \in Act, \tau \in R_\geq 0$, $X \subseteq X$ and $0 \leq i < n$), we have that $\eta_i$ is a valid clock valuation on entering location $\ell_i$ satisfying $\eta_0 = 0$, $(\eta_i + \tau) = g_i$, and $\eta_{i+1} = \eta_{i+1} \cdot X_i := 0$. Let Paths($\mathcal{M}$) (resp. Paths$_{\ell, \eta, \theta}(\mathcal{M})$) denote the set of finite paths (resp. starting in location $\ell$ with initial clock valuation $\eta$) in $\mathcal{M}$. Given a set of locations $G \subseteq Loc$, we write $RPaths_{\ell, \eta}(G) \subseteq$ Paths$_{\ell, \eta}(\mathcal{M})$ as the set of paths reaching $G$ from location $\ell$ and clock-valuation $\eta$.

b) Schedulers: The decision of which action to choose in an MTA is resolved by schedulers. A scheduler must have enough “knowledge” to make such a decision which might be the current location and clock valuation (memoryless/position independent), or the path from the initial to the current location (history dependent schedulers). We use $I(\ell) \subseteq Act$ to denote the set of actions enabled in location $\ell$.

Definition 4: [Schedulers] Let $\mathcal{M} = (Act, X, Loc, \ell_0, E, \sim)$ be an MTA. A scheduler for $\mathcal{M}$ is a measurable function $\theta : Paths(\mathcal{M}) \to Act$ such that for $n \in N$,

$$\theta(\ell_0 \sim_0 \alpha_0.\ell_1 \sim_1 \alpha_1.\ell_2 \sim_2 \ldots \sim_{n-1} \alpha_{n-1}.\ell_n) \in I(\ell_n).$$

In the definition above we assume that the scheduler makes the decision as soon as a location is entered. These are called early schedulers [14]. In contrast, a late scheduler will decide which action to take upon leaving a location, i.e., besides the history it will consider also the waiting time. In this paper we will mainly consider early schedulers, although the theory can be adapted to late schedulers easily (see Remark 1)

Given any initial location $\ell_0$, the initial clock valuation $\eta$, and scheduler $\theta$, one obtains a probability measure $Pr_{\ell_0, \eta, \theta}$ over a set of paths in a standard way (see [9]).

c) Maximum reachability: We are mainly interested in computing the maximum probability to reach a set of goal locations $G \subseteq Loc$ from the initial location $\ell_0$.

Definition 5: Let $\mathcal{M} = (Act, X, Loc, \ell_0, E, \sim)$ be an MTA and $Sched$ the set of all schedulers. The maximum probability to reach a set of goal locations $G \subseteq Loc$ from a location $\ell$ and clock valuation $\eta$ is the function $p_{\max}^{G, \ell} : Loc \times V(X) \to [0, 1]$ defined as

$$p_{\max}^{G, \ell}(\ell, \eta) = \sup_{\theta \in Sched} Pr_{\ell_0, \eta, \theta}(RPaths_{\ell, \eta}(G)).$$

In words, the maximum probability is the maximal one among all the schedulers.

The next theorem says that for MTA and maximum reachability probabilities, it suffices to consider positional schedulers instead of history dependent schedulers defined in Def. 4, i.e., the decision depends only on the current location and clock valuation.

Theorem 1: [Reachability in MTA] Let $\mathcal{M} = (Act, X, Loc, \ell_0, \sim)$ be an MTA and $G \subseteq Loc$ a set of goal locations. The maximum reachability probability $p_{\max}^{\mathcal{M}, G}$ is the least fixpoint of the integral operator $F : (Loc \times V(X) \to [0, 1]) \to (Loc \times V(X) \to [0, 1])$, where for the given function $Pr : Loc \times V(X) \to [0, 1]$, location $\ell \in Loc$ and clock valuation $\eta$, $F(Pr)(\ell, \eta) = 1$ if $\ell \in G$, and $F(Pr)(\ell, \eta) = 0$ if $RPaths_{\ell, \eta}(G) = \varnothing$. For all $\ell \notin G$ we have $F(Pr)(\ell, \eta) = \max_{\alpha \in I(\ell)} \left\{ \int_{0}^{\infty} E(\ell)e^{-E(\ell)\tau} \cdot 1_g(\eta + \tau) \cdot p \cdot Pr(\ell', \eta') \cdot d\tau \right\} = \ell \cdot \alpha.p.X \cdot f_{\alpha,p,X} \cdot \ell$, where transition $\ell \rightarrow \ell'$ is defined by transition $\ell \rightarrow \ell'$ in $\mathcal{M}$, $\zeta(X, \ell') = p$ and $\eta' = (\eta + \tau)[X := 0]$.

Remark 1: One can transform Thm. 1 to deal with the class of late schedulers. This is obtained by moving the “max” term inside the integral of (3), i.e., $F(Pr)(\ell, \eta) = \max_{\alpha \in I(\ell)} \left\{ \int_{0}^{\infty} E(\ell)e^{-E(\ell)\tau} \cdot \sum_{\alpha \in I(\ell)} 1_g(\eta + \tau) \cdot p \cdot Pr(\ell', \eta') \cdot d\tau \right\} \cdot \ell \cdot \alpha.p.X \cdot f_{\alpha,p,X} \cdot \ell$.

C. Region construction for MTA

A main step in computing maximum reachability probability or the least fixpoint of the operator defined in Thm. 1 is to apply the region construction [11] in a similar way as for standard TA. Formally, a region is an equivalence class under $\equiv$, an equivalence relation on clock valuations, which can be characterized by a specific form of a clock constraint. Let $c_x$, be the largest constant with which $x \in X$ is compared in some guard in the MTA. Clock evaluations $\eta, \eta' \in V(X)$ are clock-equivalent, denoted $\eta \equiv \eta'$, if and only if either

3In control engineering, one tends to use another terminology, i.e., controllers, while in CS community, schedulers, adversaries, policies, strategies, etc are more common. We do not distinguish them in the current paper.
1) for any \( x \in \mathcal{X} \) it holds: \( \eta(x) > c_x \) and \( \eta'(x) > c_x \), or
2) for any \( x_i, x_j \in \mathcal{X} \) with \( \eta(x_i), \eta(x_j) \leq c_x \) and \( \eta(x_j) \leq c_{x_j} \), \( \eta'(x_j) \leq c_{x_j} \) it holds: \( \eta(x_j) \leq \eta'(x_j) \) iff \( \eta(x_j) = \{\eta'(x_i)\} \) and \( \eta(x_j) \leq \{\eta'(x_i)\} \), where \( \{d\} \) is the integral (fractional) part of \( d \in \mathbb{R} \).

This clock equivalence is coarser than the traditional definition [1] by merging the “non-delayable” regions (those with point constraints like \( "x = 0" \)) into the “delayable” regions (those only with interval constraints like \( "0 < y < 1" \)). For instance, in \( \mathcal{X} = \{x_1, x_2\} \), the non-delayable regions \( (x_1 = 0, x_2 = 0) \), \( (0 < x_1 < 1, x_2 = 0) \) and \( (x_1 = 0, 0 < x_2 < 1) \) are merged with the delayable region \( (0 < x_1 < 1, 0 < x_2 < 1) \) yielding \( 0 < x_1 < 1 \). The reason for this slight change will become clear later.

We define the boundary of a region \( \Theta \) as \( \partial \Theta = \overline{\Theta} \setminus \Theta \), where \( \overline{\Theta} \) is the closure and \( \Theta \) is the interior of \( \Theta \), respectively. For instance, a region \( \Theta = (x_1 \leq 1, x_2 > 0) \) has its closure \( \overline{\Theta} = (x_1 \leq 1, x_2 \geq 0) \), its interior \( \Theta = (x_1 < 1, x_2 > 0) \) and its boundary \( \partial \Theta = (x_1 = 1, x_2 = 2) \). Here \( \Theta \) is viewed as a set of elements from \( \mathcal{V}(\mathcal{X}) \).

Let \( \mathcal{R}(\mathcal{X}) \) be the set of regions over the set \( \mathcal{X} \) of clocks. For \( \Theta, \Theta' \in \mathcal{R}(\mathcal{X}) \), \( \Theta' \) is the successor region of \( \Theta \) if for all \( \eta \models \Theta \) there exists \( \delta \in \mathcal{R}_{>0} \) such that \( \eta + \delta \models \Theta' \) and \( \forall \delta' < \delta \; \eta + \delta' \models \Theta \vee \Theta' \). The region \( \Theta \) satisfies the guard \( g \), denoted \( \Theta \models g \), iff \( \forall \eta \models \Theta \; \eta \models g \). The reset operation on region \( \Theta \) is defined as \( \Theta[X := 0] := \{\eta[X := 0] \mid \eta \models \Theta\} \).

**Notation:** Given a tuple \( v = (v_1, \ldots, v_n) \) with \( n \) components, by \( v_k \) we denote the \( k \)-th component of \( v \). In particular, for a node \( v = (\ell, \Theta) \), \( v_1 \) returns location \( \ell \), while \( v_2 \) returns the associated region \( \Theta \).

**Definition 6:** [Region graph of MTA] The region graph of MTA \( M = (\mathcal{A}, \mathcal{X}, \text{Loc}, \ell_0, E, \leadsto) \) with the set of goal locations \( G \subseteq \text{Loc} \) is \( \mathcal{G}(M) = (\mathcal{V}, \mathcal{E}, \Lambda, \widetilde{\delta}) \), where

- \( \mathcal{V} = \text{Loc} \times \mathcal{R}(\mathcal{X}) \) is a finite set of vertices with initial vertex \( v_0 = (\ell_0, \Theta_0) \), where \( \Theta_0 \) is the initial region such that \( 0 \models \Theta_0 \);
- \( \Lambda : \mathcal{V} \to \mathbb{R}_{>0} \) is the exit rate function where:
  \[
  \Lambda(v) = \begin{cases} 
  E(v_1) & \text{if } v \overset{\alpha,p,X}{\rightarrow} v' \text{ for some } v' \in \mathcal{V} \\
  0 & \text{otherwise}.
  \end{cases}
  \]
- \( \delta \subseteq \mathcal{V} \times (\mathcal{A} \times \{0,1\} \times 2^\mathcal{U} \cup \{\delta\}) \times \mathcal{V} \) is the transition (edge) relation, such that:
  - \( v \overset{\delta}{\rightarrow} v' \) if \( v_1 = v'_1 \), and \( v_2 \) is the successor region of \( v_2' \);
  - \( v \overset{\alpha,p,X}{\rightarrow} v' \) if \( v_1 \overset{\alpha,g}{\rightarrow} v'_1 \) with \( v_2 \models g \), and \( v_2' \models g \).
We also define \( \mathcal{G} = \{v \in \mathcal{V} \mid v_1 \in G\} \) to be the set of goal vertices in \( \mathcal{G}(M) \).

Any vertex in the region graph is a pair consisting of a location and a region. For a vertex \( v \in \mathcal{V} \) and clock valuation \( \eta \in \mathcal{V}^\mathcal{X} \) we define the boundary function \( b(v, \eta) = \inf\{\delta \mid \eta + \delta \in \partial v_2\} \), which is the minimum time (if it exists) to “hit” the boundary of the region corresponding to vertex \( v \) starting from a clock valuation \( \eta \). Edges of the form \( v \overset{\delta}{\rightarrow} v' \) are called delay edges (jump), whereas those of the form \( v \overset{\alpha,p,X}{\rightarrow} v' \) are called Markovian edges (jump). Note that Markovian edges emanating from a vertex corresponding to a non-delayable region do not contribute to the reachability probability. The waiting time in such vertex is always zero. Therefore, we can safely remove all the Markovian edges emanating from vertices with non-delayable regions and combine each such non-delayable region with its unique delayable (direct) successor. In the sequel, by slight abuse of notation, we refer to this simplified region graph as \( \mathcal{G}(M) \). Note that then \( v_2[X := 0] \subseteq v'_2 \) in the last item of Def. 6.

An example region graph is shown in Fig. 4.

**Example 2:** For the MTA \( M \) in Fig. 3, the reachable part (forward reachable from the initial vertex and backward reachable from the accepting vertices) of the region graph \( \mathcal{G}(M) \) is shown in Fig. 4.

Now we define a system of integral equations on the region graph \( \mathcal{G}(M) \) which will help compute the maximum reachability probability from Thm. 1.

**Definition 7:** Given the region graph \( \mathcal{G}(M) = (\mathcal{A}, \mathcal{V}, v_0, \Lambda, \leadsto) \) of the MTA \( M = (\mathcal{A}, \text{Loc}, \mathcal{X}, \ell_0, E, \leadsto) \) and the set of goal vertices \( G \), for the function \( \text{Prob}_v(\eta) : \mathcal{V} \times \mathcal{V}(\mathcal{X}) \to [0,1] \) let the operator \( \overline{F} : (\mathcal{V} \times \mathcal{V}(\mathcal{X}) \to [0,1]) \to (\mathcal{V} \times \mathcal{V}(\mathcal{X}) \to [0,1]) \) be defined as \( \overline{F}(\text{Prob}_v(\eta)) = 1 \) if \( v \in G \); \( \overline{F}(\text{Prob}_v(\eta)) = 0 \) if \( v \notin G \) and \( G \) cannot be reached from \( v \); and for \( v \notin G \) we have

\[
\overline{F}(\text{Prob}_v(\eta)) = \text{Prob}_v,\delta(\eta) + \max_{\alpha \in \mathcal{A}(v_1)} \text{Prob}_v,\alpha(\eta),
\]

\[
\text{Prob}_v,\alpha(\eta) = \int_0^{\lambda(v,\eta)} \Lambda(v)e^{-\Lambda(v)\tau}\sum_{v' \in \mathcal{E}(v,\eta)} p_{v,v'}((\eta + \tau)[X := 0]) \, d\tau,
\]

\[
\text{Prob}_v,\delta(\eta) = e^{-\Lambda(v,\delta)(\eta \cup \delta)} \cdot \text{Prob}_v,\eta \cup \delta(\eta),
\]

where \( \text{Prob}_v,\alpha(\eta) \) denotes the probability to reach \( G \) by taking a Markovian jump and \( \text{Prob}_v,\delta(\eta) \) the probability to reach \( G \) through vertex \( v' \) by taking the delay jump \( v \overset{\delta}{\rightarrow} v' \).

**Theorem 2:** Let \( M = (\mathcal{A}, \text{Loc}, \mathcal{X}, \ell_0, E, \leadsto) \) be an MTA with the set of goal locations \( G \) and \( \text{Prob}_v(\eta) \) be the least fixpoint of the operator \( \overline{F} \), then for \( v_1 = \ell \) we have

\[
\text{Prob}_v(\eta) = \text{Prob}_{\max}^M(\ell, \eta).
\]
Based on this theorem, we will focus on the efficient computation of maximum time-bounded reachability probabilities in region graphs (instead of in MTAs) in the following section.

III. TIME-BOUNDED REACHABILITY

In this section, we concentrate on maximizing time-bounded reachability probabilities in a region graph $G(M) = (Act, V, v_0, \Lambda, \rightarrow)$, i.e., given a set $G$ of goal vertices and time bound $T \in \mathbb{N}$, we are interested in maximizing the probability to reach $G$ within $T$ time units. To this end, we introduce a fresh clock $t$, denoting the global time which is initialized to zero and never reset. To distinguish the role of $t$, we write the state of $G(M)$ as $(v, \eta, t)$. As in this case time bounds to reach $G$ have to be considered, $t \leq T$ should be added to each vertex of $G(M)$. Formally, by notation abuse we overload $\beta(v, \eta, t)$ to be $\min\{\beta(v, \eta), T - t\}$, i.e. the minimum time to hit the boundary $\delta v_{1|2}$ at time $t$. For instance, let $\partial v_{1|2}$ be $x = 1 \land y = 2$, $T = 100$, and suppose $\eta(x) = 0.5, \eta(y) = 1.7$. If $t = 2.5$, then $\beta(v, \eta, t) = \min\{1 - 0.5, 2 - 1.7, 100 - 2.5\} = 0.3$; if $t = 99.9$, then $\beta(v, \eta, t) = \min\{1 - 0.5, 2 - 1.7, 100 - 99.9\} = 0.1$.

The following Bellman (dynamic programming) equations derived from Def. 7 play an essential role in solving the time-bounded reachability problem. Let $P(v, \eta, t)$ be the maximum probability at time $t$ for state $(v, \eta)$ to reach $G$ within time bound $T$. $P(v, \eta, t) = 1$ if $v \in G$ and $t \leq T$; 0 if $t > T$ or $G$ is not reachable from $v$; and otherwise $P(v, \eta, t) =$

$$\max_{\alpha \in \mathcal{I}(v)} \left\{ \sum_{v', \eta', t'} P(v', \eta', t') \int_{0}^{\beta(v, \eta, t)} \Lambda(v) e^{-\Lambda(v) \tau} P(v', \eta', t+\tau) d\tau \right\},$$

where $\mathcal{I}(v)$ is the set of actions enabled in $v$, $\eta' = (\eta + \tau) | X := 0$ and $v \xrightarrow{\delta} v''$, where $v''$ is the time successor of $v$. Term (I) represents the maximum reachability probability (among all enabled actions) by taking a Markovian jump $v \xrightarrow{\alpha, p, X} v'$ and (II) represents the probability of taking the boundary jump $v \xrightarrow{\delta} v''$. Note that $(\ast)$ is the density function of taking $v \xrightarrow{\alpha, p, X} v'$ at time $\tau$ and $(\ast\ast)$ is the probability not to leave location $v$ within $\beta(v, \eta, t)$ time units.

We will now provide two ways to solve (4): one by discretization (4) and the other based on the Hamilton-Jaciob-Bellman equation [11].

A. Discretization

Our first approach is to discretize the continuous variables in the Bellman equation. Using a discretization step $h = \frac{1}{N}$ $(N \in \mathbb{N}_{>0})$, the aim is to obtain a finite-state MDP $D(M)$ from $G(M)$. For this MDP, a similar Bellman equation can be derived and solved efficiently e.g. by value iteration [3]. Intuitively, $h$ is the length of time in which a single Markovian jump takes place from a given location. By $h$, each Markovian jump in $G(M)$ can be approximated by a Markovian jump which only takes place at time points $\{0, h, \ldots, NT\}$. This gives rise to an MDP:

**Definition 8:** Given $G(M) = (Act, V, v_0, \Lambda, \rightarrow)$ and discretization step $h = \frac{1}{N}$ $(N \in \mathbb{N}_{>0})$, the MDP $D_h(M) = (Act \cup \{\perp\}, S, s_0, \mathcal{P})$ is defined as follows:

- $S = \{(v, \eta, t) \mid v \in V \land \eta \in \nu_{2} \land t \leq T\}$;
- $s_0 = (v_0, 0, 0)$;
- $\perp$ is a fresh action encoding the “delay” in $G(M)$;

For each $(v, \eta, t) \in S$ we distinguish three cases:

(i) If $h < \beta(v, \eta, t)$ and $v \xrightarrow{\alpha, p, X} v'$ then

$$\mathcal{P}(\{(v, \eta, t), \alpha, (v', \eta[X := 0])\}) = p \cdot (1 - e^{-\Lambda(v)h});$$

(ii) If $h < \beta(v, \eta, t)$ and Act($v$) = $\perp$ then

$$\mathcal{P}(\{(v, \eta, t), \perp, (v, \eta + h, t + h)\}) = e^{-\Lambda(v)h};$$

(iii) If $h \geq \beta(v, \eta, t)$ and $v \xrightarrow{\delta} v'$ then

$$\mathcal{P}(\{(v, \eta, t), \perp, (v', \eta + \beta(v, \eta, t), t + h(v, \eta, t))\}) = e^{-\Lambda(v)h(v, \eta, t)}.$$
where $\tilde{y} = \eta [X := 0]$ and $v \xrightarrow{\delta} v''$.

By the discretization step $h$, the regions of $G(\mathcal{M})$ contain finitely many points (one point is a state in the MDP). To be more exact, each region has maximally $h$ points for each clock. Together with the fact that there are only finitely many regions of interest in $G(\mathcal{M})$, the number of states in the MDP is finite.

Theorem 3: [Error bound] For any state $(v, \eta)$, time bound $T$, discretization step $h = \frac{1}{N}$, and $\lambda = \max_{v \in V} \{\Lambda(v)\}$ which is the maximum rate of all exponential distributions appearing in the region graph:
\[
\sup_{t \in [0, T]} |P(v, \eta, t) - Y(v, \eta, t)| \leq (1 - e^{-\lambda h})(1 - e^{-\lambda T}).
\]

B. Hamilton-Jacobi-Bellman Equations

As in traditional control theory [3], the dynamic programming principles lead to a first-order integro-differential equation, which is the Hamilton-Jacobi-Bellman (HJB) partial differential equation (PDE).

Given the region graph $G(\mathcal{M}) = (Act, V, v_0, \Lambda, \rightarrow)$ let $f(v, \eta, t) := P(v, \eta, t)$ be the maximum time-bounded reachability probability at time $t$ in $G(\mathcal{M})$. For every $v \in V$ and $\eta \in \mathcal{V}(\mathcal{X})$ and $t \leq T$ let $f(v, \eta, t)$ be given as follows:
\[
\frac{\partial f(v, \eta, t)}{\partial t} + \sum_{i=1}^{[\lambda]} \frac{\partial f(v, \eta, t)}{\partial \eta^{(i)}} = \max_{\alpha \in \mathcal{T}(v)} \left\{ \Lambda(v) \cdot \sum_{\alpha \xrightarrow{X \rightarrow v'} \eta'} p(f(v, \eta, t) - f(v', \eta, [X := 0], t)) \right\},
\]

where $\eta^{(i)}$ is the $i$'th clock variable. The initial conditions of the above PDE are $f(v, \eta, T) = 1_G(v, \eta)$ for any $v \in V$ and $\eta \in v|_2$. Moreover, for every $\eta \in \partial v|_2$ and transition $v \xrightarrow{\delta} v'$, the boundary conditions take the form $f(v, \eta, t) = f(v', \eta, t)$.

Several methods can be used to solve the above HJB equation, e.g., the finite volume method [17] or the time and space discretization technique [7]. Note that for zero-clock MTA, i.e., CTMDPs, we obtain a system of ODEs instead of PDEs; for details see [9].

IV. Conclusion

We have defined an extension of timed automata with exponentially distributed durations on locations. We constructed region graphs of such automata based on which the time-bounded reachability problems were investigated. We proposed Bellman equations to characterize the probability, and presented two approaches to solve these equations, namely, by discretization and a reduction to HJB PDEs.

Another interesting question is the time-unbounded reachability problem, i.e., the deadline of reaching the goal locations is absent. We refer the readers to [10], [9] for the solution of this question. In a nutshell, we show that, for single-clock MTA, they can be characterized as the solution of a system of linear equations whose coefficients are maximum reachability probabilities in CTMDPs, i.e., zero-clock MTA. Moreover, for general MTA, Bellman equations, which are a variant of Eq. (4), are proposed to characterize the probability. We remark that in this paper only the locally uniform model (i.e., the exit rate of each location solely depends on the location instead of actions) was addressed, however the general case can be treated without any difficulty.

Many future works remain to be done. For example, we plan to extend the MTA model to rewards; also one can consider more general variables than clocks, resulting in Markovian hybrid automata.

References