A continuous-time model-based approach for activity recognition in pervasive environments

Marco Biagi, Laura Carnevali, Marco Paolieri, Fulvio Patara, and Enrico Vicario, *IEEE member*

***Abstract*—We present a model-based approach to Activity Recognition (AR) in Ambient Assisted Living (AAL). The approach leverages an a-priori stochastic model termed Continuous Time Hidden Semi-Markov Model (CT-HSMM), capturing the continuous-time durations of activities and inter-event times. The model is enhanced according to the observed statistics, associating the events with an occurrence probability, and the sojourn time and the inter-event time in each activity with a continuous-time Probability Density Function (PDF), allowing effective fitting of observed durations through non-Markovian distributions. The model is updated at run-time according to a sequence of time-stamped observations, exploiting the method of stochastic state classes to perform transient analysis and derive a measure of likelihood that an activity is currently performed. The approach supports both online AR, predicting the activity performed at time t using only the events observed until that time, and offline AR, applying a Forward-Backward procedure that exploits all the events observed before and after time t. The approach is experimented on a real data set of the literature, providing performance measures that can be compared with those of offline Hidden Markov Models (HMMs) and offline Hidden Semi-Markov Models (HSMMs).**

***Index Terms*—Activity Recognition (AR), Ambient Assisted Living (AAL), continuous-time stochastic model, Semi-Markov Process (SMP), transient analysis, model@run.time.**

I. INTRODUCTION

SMART homes support monitoring of high level human activities through typed and timestamped low-level observations collected by physical sensors, enabling services for healthcare, safety, security, and comfort, with notable applications in Ambient Assisted Living (AAL) [1]. To this end, Activity Recognition (AR) comprises a key challenge which has been widely addressed in pervasive computing systems [2] through data-driven and knowledge-based models [3], with various time abstractions enabling different exploitation of information carried by observed durations.

Data-driven approaches for AR have largely developed on Hidden Markov Models (HMMs) [4] and Conditional Random Fields (CRFs) [5]. In the HMM approach, activities are represented as hidden states of a *discrete* time Markov chain and an observation depending only on the current visited state is generated at each transition. Stochastic parameters of the HMM can be derived from a ground truth, through supervised learning or in closed form when model states are identified with predefined activities, and a maximum likelihood sequence of hidden activities corresponding to an observation sequence can be identified through efficient algorithms. CRFs relax the assumption of independent observations by directly modeling the conditional probability of a sequence of activities given a sequence of observations. HMMs and CRFs are applied and compared in [6] in the recognition of predefined Activities of Daily Living (ADLs) from sensor readings acquired in a smart home setting. In [7], activities are discovered by clustering emerging patterns through an unsupervised approach, an HMM is constructed for each cluster, and classifications are combined through a voting mechanism.

The *sojourn duration* in each hidden state of an HMM is implicitly associated with a geometric distribution, which is memoryless and fully identified by a single parameter, thus limiting exploitation of timing information that may be significant in various contexts [8]. The duration of activities is considered in [9] using Hidden Semi-Markov Models (HSMM) [10] and Semi-Markov Conditional Random Fields (SMCRFs) [11], which extend HMMs and CRFs, respectively, by explicitly modeling the distribution of the time segment during which the hidden state sojourns in a logical location. Interval State Hidden Semi Markov Models (IS-HSMMs) [12] extend HSMMs with additional states representing no-observation intervals between consecutive events. Similarly, Interval Length Probability Hidden Semi-Markov Models (ILP-HSMMs) model the discrete-time distribution of such intervals [12]. In [13], HSMMs are extended to consider long-range dependencies between observations.

As a common trait, HMMs, CRFs, and their semi-Markov extensions all develop on a *discrete* abstraction of time, which in AR is usually realized by sampling the continuous underlying process at equidistant time points through the assumption of a time-tick. A few models were proposed to preserve a *continuous time* representation so as to support model compositionality, or to avoid the trade-off between accuracy and complexity that affects discretization, or more generally to fit the native characteristics of the observation process. Continuous-time HMMs were proposed in [14] in the context of performance evaluation to enable data-driven identification of a Continuous Time Markov Chain (CTMC). The problem can be reduced to the identification of a (discrete time) HMM which is then enhanced with continuous sojourn times, which however are necessarily distributed as memoryless exponential random variables. In Hidden non-Markovian Models (Hn-

This work was supported in part by the REMIND Project, funded by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skodowska-Curie grant agreement n° 734355.

M. Biagi, L. Carnevali, F. Patara and E. Vicario are with the Information Engineering Department, University of Florence, via di Santa Marta 3, 50139 Firenze, Italy. E-mail: {marco.biagi, laura.carnevali, fulvio.patara, enrico.vicario}@unifi.it

M. Paolieri is with the Computer Science Department, University of Southern California, 941 Bloom Walk, Los Angeles, CA 90089, USA. E-mail: paolieri@usc.edu

MMs) [15], sojourn times in hidden states are associated with non-Markovian (i.e. non exponential) continuous distributions under the assumption that every change of the hidden state is made observable by the emission of an event, and approximate transient probabilities are derived by discretization through the Proxel simulation method [16] and used as a measure of likelihood for the identification of the hidden state. Non-Markovian sojourn times are considered also in Generalized Hidden Semi Markov Models (GHSMMs) [17], for which efficient algorithms are provided for training and inference under the assumption of a one-to-one mapping between hidden states and observation symbols.

Knowledge-based approaches reduce dependency on the volume of training data through pre-defined models capturing a-priori knowledge about hidden states. Online AR is addressed in [18], using ontologies for activity modeling and semantic reasoning for activity inferencing, supporting AR at coarse-grained and fine-grained levels [19] and adaptation of ADL models over time [20]. In [21], a top-level ontology links knowledge implicit in elementary information to higher-level activity descriptions. In *process mining* [22], knowledge is extracted from event logs and used to improve processes [23]. Good results have been achieved in reconstruction of administrative workflows [24], while applicability appears more difficult for (less structured) healthcare pathways [25]. Techniques for *process enhancement* improve models with information extracted from observations, e.g., they enrich untimed models with continuous time stochastic parameters [26].

In this paper, we largely extend preliminary results of [27], presenting a novel approach to AR that exploits the discrimination capability of continuous-time durations of activities and inter-occurrence times of observed events. The approach relies on a novel generative stochastic model termed Continuous Time Hidden Semi-Markov Model (CT-HSMM), which captures the statistics of sojourn time in hidden activities and of the frequency of typed events observed in each of them, with notable extensions with respect to HnMMs and GHSMMs. In particular, the model accepts a randomized number of events observed during the sojourn in each activity and represents durations with non-Markovian distributions, making the likelihood of activities be dependent on the time elapsed between observations. Transient probabilities of future behavior until the next events conditioned on the actual sequence of past observations are efficiently computed at runtime [28] using the method of stochastic state classes [29], [30], and they are used as a measure of likelihood for the identification of the current ongoing hidden activity. The proposed approach enables both online and offline classification, yielding performance measures that can be compared with those of state-of-the-art offline approaches on a real data set of the AAL literature [6].

In the following: we formulate the problem and present the computed statistics (Section II); we define CT-HSMMs and perform model enhancement extending [31] to fit sample data (Section III); we present a solution technique for online and offline classification, tailoring the method of stochastic state classes [29], [30] to derive the likelihood of activities (Section IV); we discuss experimental results, (Section V), and, we draw our conclusions (Section VI). Technical details on stochastic state classes and on the proposed fitting approach are reported in the Supplemental Material with theorem proofs.

II. Problem formulation: AR in AAL

A. An abstraction for activities and events

AR in AAL is a problem of partial observability where high-level *activities* are inferred from low-level *events* about human-object interactions. Events are typically captured by sensors (e.g., reed switches [6]) as typed and time-stamped observations (e.g., *Opening fridge* at 4:20pm). Similarly, activities performed by a human subject are characterized by a *type* in the index of Activities of Daily Living (ADLs) [32] (e.g., *Preparing dinner*) and by the *time period* (identified by a starting and an ending time-stamp) during which a stream of events is generated while accomplishing the activity goal. A special type of activity termed *Idling* identifies the time intervals during which no specified activity is performed.

The process of performed activities and observed events is uncertain, with different activities sharing the same event type (e.g., *Opening/Closing fridge* is a common event of *Preparing dinner* and *Preparing a beverage*) and different event types occurring in the same activity (e.g., *Opening/Closing fridge* and *Starting/Ending dishwasher* for *Preparing dinner*). Moreover, human activities are typically performed in complex and loosely structured manners, with limited repetitiveness. In this conceptual framework, AR consists in evaluating a measure of likelihood that an activity is currently performed, starting from a historical data set containing a stream of typed and time-stamped events with ground truth data about tagged activities.

B. A statistical characterization of AR data sets

The volume of available data is a key factor in the selection of statistics that can be affordably exploited. The proposed approach addresses recognition of sequential activities from events generated by switch sensors, leveraging five statistics on the frequency of events, the continuous duration of each activity, and the continuous time between consecutive events. As shown in Fig. 1, let: $\mathsf{E} = \{\mathsf{e}_1, ..., \mathsf{e}_H\}$ and $\mathsf{A} = \{\mathsf{a}_1, ..., \mathsf{a}_K\}$ be the sets of event types and activity types, respectively; a_* be the *Idling* type; $\mathsf{Ev} = \{\mathsf{ev}_n = \langle \mathsf{e}_{x(n)}, t_n \rangle\}_{n=1}^{N}$ be the list of observed events in dual change-point representation (distinguishing activation and de-activation events), where $\mathsf{e}_{x(n)}$ and t_n are the type and time of the n-th event, respectively; and, $\mathsf{Act} = \{\mathsf{act}_m = \langle \mathsf{a}_{y(m)}, t_m^{\text{start}}, t_m^{\text{end}} \rangle\}_{m=1}^{M}$ be the list of tagged activities, where $\mathsf{a}_{y(m)}$ is the type of the m-th activity, and t_m^{start} and t_m^{end} are its start and end times, respectively.

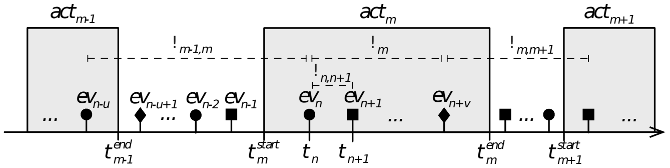


Fig. 1. A stream of observed events with sequential annotated activities. Circle, square and diamond pins represent punctual events with different types.

1) $\mathsf{F}_\mathrm{d}(\mathsf{a}_k)$ is the duration histogram of activity type a_k: if $\mathsf{a}_k \neq \mathsf{a}_*$, duration is the time δ_m from the first to the last event of activity act_m of type a_k; if $\mathsf{a}_k = \mathsf{a}_*$, it is the time $\delta_{m-1,m}$ from the last event of activity act_{m-1} to the first event of the consecutive activity act_m. Tagged durations are not considered to avoid dependency on annotated data sets (not always available) and tagging errors (start and end times are often delayed/anticipated).
2) $\mathsf{F}_\mathrm{i}(\mathsf{a}_k)$ is the inter-event time histogram of activity type a_k: if $\mathsf{a}_k \neq \mathsf{a}_*$, inter-event time is the time $\delta_{n,n+1}$ between consecutive events ev_n and ev_{n+1} in activity of type a_k; if $\mathsf{a}_k = \mathsf{a}_*$, it is the time between consecutive events ev_n, ev_{n+1} s.t. ev_n occurs in *Idling*, and ev_{n+1} occurs in *Idling* or is the first event of a tagged activity (e.g., $[\mathsf{ev}_{n-2}, \mathsf{ev}_{n-1}]$ and $[\mathsf{ev}_{n-1}, \mathsf{ev}_n]$ in Fig. 1).
3) $\mathsf{P}_\mathrm{x}(\mathsf{e}_h, \mathsf{a}_k)$ is the frequency of an event of type e_h in an activity of type a_k, i.e., $\mathsf{P}_\mathrm{x}(\mathsf{e}_h, \mathsf{a}_k) := \frac{\mathcal{E}(\mathsf{e}_h, \mathsf{a}_k)}{\mathcal{E}(\mathsf{a}_k)}$, where $\mathcal{E}(\mathsf{e}_h, \mathsf{a}_k)$ is the number of events of type e_h observed in activities of type a_k and $\mathcal{E}(\mathsf{a}_k)$ is the number of events of any type emitted in activities of type a_k.
4) $\mathsf{P}_\mathrm{y}(\mathsf{e}_h)$ is the frequency of an event of type e_h to occur during *Idling* or as the starting event of a tagged activity, i.e., $\mathsf{P}_\mathrm{y}(\mathsf{e}_h) := \frac{\mathcal{E}_\mathrm{si}(\mathsf{e}_h)}{\mathcal{E}_\mathrm{si}}$, where $\mathcal{E}_\mathrm{si}(\mathsf{e}_h)$ is the number of events of type e_h occurring during *Idling* or as the starting event of a tagged activity, and $\mathcal{E}_\mathrm{si}$ is the number of events (of any type) occurring during *Idling* or as the starting event of a tagged activity.
5) $\mathsf{P}_\mathrm{z}(\mathsf{e}_h, \mathsf{a}_k)$ is the frequency of event of type e_h to be the starting event of activity of type a_k, i.e., $\mathsf{P}_\mathrm{z}(\mathsf{e}_h, \mathsf{a}_k) := \frac{\mathcal{E}_\mathrm{s}(\mathsf{e}_h, \mathsf{a}_k)}{\mathcal{E}_\mathrm{si}(\mathsf{e}_h)}$, where $\mathcal{E}_\mathrm{s}(\mathsf{e}_h, \mathsf{a}_k)$ is the number of events of type e_h that occur as the first event of activity of type a_k.

III. Continuous Time Hidden Semi-Markov Model

Information carried by observation data is captured into a generative model which we term Continuous Time Hidden Semi Markov Model (CT-HSMM). The state of the model consists of the current activity (either tagged or *Idling*), the remaining sojourn time, and the remaining time to the next observable event. An observable event may maintain the model in the current activity or may bring the model from *Idling* to an annotated activity; conversely, expiration of the remaining sojourn time in an annotated activity and in *Idling* is an unobservable event that brings the model to *Idling* and to a deadlock location, respectively. The model is enhanced through the computed statistics, associating observable events with a probability, and the sojourn time and inter-event time in each activity with a Probability Density Function (PDF).

A. Model definition

Syntax. A CT-HSMM is a tuple $M := \langle A, a_*, E, P_\mathrm{x}, P_\mathrm{y}, P_\mathrm{z}, T_\mathrm{o}, T_\mathrm{u}, F_\mathrm{d}, F_\mathrm{i} \rangle$ [1] where:

- $A = \{a_1, \ldots, a_K\}$ is the set of *annotated activities*, a_* is the *Idling* activity, and $A_* = A \cup \{a_*\}$.
- $E = \{e_1, \ldots, e_H\}$ is the set of *observable events*.
- $P_\mathrm{x} : E \times A \to [0,1]$ is the probability that $e \in E$ occurs in $a \in A$, i.e., $P_\mathrm{x}(e,a) := P\{\epsilon(t) = e \,|\, \alpha(t) = \alpha(t_\mathrm{pr}) = a\}$, $\epsilon(t)$ being the event observed at t, $\alpha(t)$ the activity performed at t, and t_pr the last observation time before t.
- $P_\mathrm{y} : E \to [0,1]$ is the probability that $e \in E$ is the first event in a tagged activity or occurs in *Idling*, i.e., $P_\mathrm{y}(e) := P\{\epsilon(t) = e | \alpha(t) = a_* \vee \epsilon(t)\ 1^\mathrm{st}\ \text{ev.}\ \alpha(t) \in A\}$.
- $P_\mathrm{z} : E \times A \to [0,1]$ is the probability that an event is the first one in $a \in A$ given that it is $e \in E$ and occurs in *Idling* or as the first event of a tagged activity, i.e., $P_\mathrm{z}(e,a) := P\{\epsilon(t)\ 1^\mathrm{st}\ \text{ev.}\ \alpha(t) = a \,|\, \epsilon(t) = e \wedge (\alpha(t) = a_* \vee \epsilon(t)\ 1^\mathrm{st}\ \text{ev.}\ \alpha(t) \in A)\}$.
- $T_\mathrm{o} \subseteq A_* \times E \times A_*$ is the set of *observable transitions* from an activity (*Idling* or tagged) to itself and from *Idling* to a tagged activity, i.e., $(a, e, a') \in T_\mathrm{o}$ if $a = a' \in A \wedge P_\mathrm{x}(e,a) \neq 0$, or $a = a' = a_* \wedge P_\mathrm{y}(e) \neq 0 \wedge \sum_{j \in A} P_\mathrm{z}(e,j) \neq 1$, or $a = a_* \wedge a' \in A \wedge P_\mathrm{z}(e,a) \neq 0$.
- $T_\mathrm{u} = (A \times \{a_*\}) \cup (a_*, \phi)$ is the set of *unobservable transitions* from an annotated activity to *Idling*, and from *Idling* to a deadlock location ϕ.
- $F_\mathrm{d} : A_* \to [0,\infty)^{[F^\mathrm{min}_{\mathrm{d},a}, F^\mathrm{max}_{\mathrm{d},a}]}$ provides the duration PDF $F_\mathrm{d}(a)$ of $a \in A_*$, i.e., $F_\mathrm{d}(a) : [F^\mathrm{min}_{\mathrm{d},a}, F^\mathrm{max}_{\mathrm{d},a}] \to [0,\infty)$, with $F^\mathrm{min}_{\mathrm{d},a} \in \mathbb{Q}_{\geq 0}$ and $F^\mathrm{max}_{\mathrm{d},a} \in \mathbb{Q}_{\geq 0} \cup \{\infty\}$.
- $F_\mathrm{i} : A_* \to [0,\infty)^{[F^\mathrm{min}_{\mathrm{i},a}, F^\mathrm{max}_{\mathrm{i},a}]}$ provides the inter-event time PDF $F_\mathrm{i}(a)$ in $a \in A_*$, i.e., $F_\mathrm{i}(a) : [F^\mathrm{min}_{\mathrm{i},a}, F^\mathrm{max}_{\mathrm{i},a}] \to [0,\infty)$, with $F^\mathrm{min}_{\mathrm{i},a} \in \mathbb{Q}_{\geq 0}$, $F^\mathrm{max}_{\mathrm{i},a} \in \mathbb{Q}_{\geq 0} \cup \{\infty\}$.

Semantics. The state of the model is a triple $s = \langle \psi, \theta, \iota \rangle$ where $\psi \in A_*$ is the current activity, $\theta \in \mathbb{R}_{\geq 0}$ is the remaining sojourn time, and $\iota \in \mathbb{R}_{\geq 0}$ is the remaining time to the next observable event. System evolution from a state $s = \langle \psi, \theta, \iota \rangle$ depends on whether the current location is a_* or an annotated activity, and θ is larger or smaller than ι.

- Case $\psi = a_*$.
  Event observed. If $\theta > \iota$, an event $e \in E$ is observed with probability $P_\mathrm{y}(e)$, and a new state $s' = \langle \psi', \theta', \iota' \rangle$ is selected randomly: with probability $P_\mathrm{z}(e,a)$, an observable transition (a_*, e, a) from a_* to $a \in A$ occurs, in which case $\psi' = a$, while θ' and ι' are sampled from $F_\mathrm{d}(a)$ and $F_\mathrm{i}(a)$, respectively; with probability $1 - \sum_{a \in A} P_\mathrm{z}(e,a)$, an observable transition (a_*, e, a_*) occurs, so that $\psi' = a_*$, θ' is derived reducing θ by the time ι elapsed in s, and ι' is sampled from $F_\mathrm{i}(a_*)$.
  Duration expired. If $\theta \leq \iota$, the unobservable transition (a_*, ϕ) occurs and the deadlock state $s' = \langle \phi, \cdot, \cdot \rangle$ is reached, from which no other transition is possible.
- Case $\psi = a \in A$.
  Event observed. If $\theta > \iota$, an observable transition (a, e, a) occurs from a to itself with probability $P_\mathrm{x}(e,a)$, yielding $s' = \langle a, \theta', \iota' \rangle$ where θ' is θ minus the time ι elapsed in s, and ι' is sampled from $F_\mathrm{i}(a)$.
  Duration expired. If $\theta \leq \iota$, an unobservable transition (a, a_*) occurs, yielding $s' = \langle a_*, \theta', \iota' \rangle$ where θ' and ι' are sampled from $F_\mathrm{d}(a_*)$ and $F_\mathrm{i}(a_*)$, respectively.

At the initial time, the system is in state $s_0 = \langle \psi_0, \theta_0, \iota_0 \rangle$ where $\psi_0 = a_*$, and θ_0 and ι_0 are sampled from $F_\mathrm{d}(a_*)$ and $F_\mathrm{i}(a_*)$, respectively. Any annotated activity might be

[1] To simplify presentation, in the CT-HSMM model we refer to activity types and event types as activities and events, respectively.

selected as the initial location as well, without impacting on the complexity of the subsequent treatment.

Running example. The model in Fig. 2 has 2 annotated activities a_1 and a_2, and 4 observable events $e_1, \dots, e_4$, with activities depicted as circles, and transitions as solid or dashed arrows if they are observable or unobservable, respectively, with tick or thin line if they restart both the remaining inter-event time ι and the remaining sojourn time θ, or ι only, respectively ($\forall\ a \in A$, $\forall\ e \in E$, transitions (a, e, a), (a_*, e, a_*), (a_*, e, a), and (a, a_*) are depicted as solid thin arrows labeled with $\langle e, P_{\mathrm{x}}(e, a)\rangle$, solid thin arrows labeled with $\langle e, P_{\mathrm{y}}(e), 1 - \sum_{a\in A} P_{\mathrm{z}}(e, a)\rangle$, solid tick arrows labeled with $\langle e, P_{\mathrm{y}}(e), P_{\mathrm{z}}(e, a)\rangle$, and dashed tick arrows, respectively). In *Idling* a_*, events e_1, e_2, e_3 occur with probability 0.35, 0.15, and 0.5, respectively: at e_1 or e_2, *Idling* a_* continues with probability 1; at e_3, *Idling* a_* continues with probability 0.25, or a_1 or a_2 are started with probability 0.30 and 0.45, respectively. During a_1, events e_1, e_2, e_4 occur with probability 0.30, 0.60, and 0.10, respectively; besides, e_2, e_3, e_4 occur during a_2 with probability 0.35, 0.15, and 0.50, respectively.

Idling a_* has hyper-exponential sojourn time with PDF $f(x) = \sum_{i=1}^{2} p_i\, \lambda_i\, e^{-\lambda_i\, x}$ (not shown in Fig. 2) where $\lambda_1 = 0.01$, $\lambda_2 = 0.02$, $p_1 = 0.2$, $p_2 = 0.8$, and exponential inter-event time with PDF $f(x) = \lambda\, e^{-\lambda x}$ where $\lambda = 0.05$; a_1 has hyper-exponential sojourn time with $\lambda_1 = 0.001$, $\lambda_2 = 0.005$, $p_1 = 0.3$, $p_2 = 0.7$, and exponential inter-event time with $\lambda = 0.03$; and, a_2 has hypo-exponential sojourn time with PDF $f(x) = \frac{\lambda_1 \lambda_2}{\lambda_1 - \lambda_2}(e^{-\lambda_2\, x} - e^{-\lambda_1\, x})$ where $\lambda_1 = 0.05$, $\lambda_2 = 0.1$, and exponential inter-event time with $\lambda = 0.2$.

Underlying stochastic process. In any state, the remaining time to the end of the current activity (θ) and to the next observable event (ι) are concurrent and (possibly) generally distributed random variables, and thus the stochastic process $\mathbb{X} := \{s(t),\ t \in \mathbb{R}_{\geq 0}\}$, with $s(t)$ denoting the state of the model at time t, is not a Markov chain. At the observation of any event, ι is restarted but θ continues; but, always, with probability 1, an activity will eventually terminate and both θ and ι will be resampled; at that point, the future behavior depends on the past only through the current state (Markov condition), i.e., the process $\mathbb{X}$ *regenerates*. According to this, $\mathbb{X}$ is a Markov Regenerative Process (MRP) [33], [34]. Note that any number of events may be observed during an activity, thus the number of transitions between consecutive regenerations of $\mathbb{X}$ is unbounded. However, the number of transitions between observable events is not larger than 2: in any tagged activity, an unobservable transition can reach only *Idling*; in turn, in *Idling*, an unobservable transition can lead only to a deadlock.

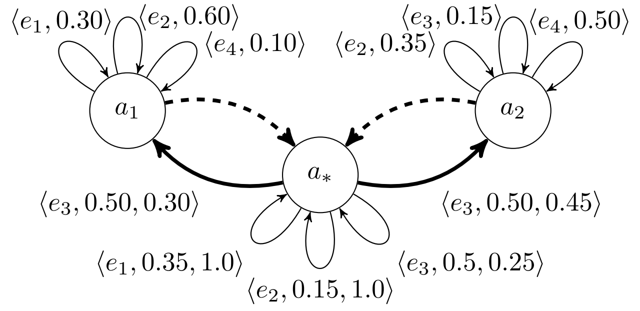

$\langle e_1, 0.30\rangle$
$\langle e_2, 0.60\rangle$
$\langle e_4, 0.10\rangle$
$\langle e_3, 0.15\rangle$
$\langle e_4, 0.50\rangle$
$\langle e_2, 0.35\rangle$
a_1
a_2
a_*
$\langle e_3, 0.50, 0.30\rangle$
$\langle e_3, 0.50, 0.45\rangle$
$\langle e_1, 0.35, 1.0\rangle$
$\langle e_3, 0.5, 0.25\rangle$
$\langle e_2, 0.15, 1.0\rangle$

Fig. 2. A model with 2 annotated activities and 4 observable events. Sojourn and inter-event times, and the unobservable transition (a_*, ϕ), are not shown.

While $\mathbb{X}$ is an MRP, the way it is used in combination with observations permits to manipulate it as an instance of the more restricted class of (continuous-time) Semi-Markov Processes (SMP) [33], [34]. In fact, at each observed event, the remaining time θ to the end of the current activity continues, but the event timestamp associates the time elapsed since the previous state with a *determined* value τ. It is thus possible to associate θ with a new distribution that makes the future behavior of $\mathbb{X}$ be dependent on the past only through the current state (Markov condition): if θ_p is the random variable at the previous event and f_{θ_p} is its PDF, θ is $(\theta_p - \tau | \theta_p > \tau)$ and its PDF is thus $f_\theta(x) = f_{\theta_p}(x + \tau)/(\int_\tau^\infty f_{\theta_p}(y)dy)$. For this reason, following the notation of [14], the model is termed Continuous Time Hidden Semi-Markov Model (CT-HSMM).

B. Model enhancement

The model is derived in closed form from the statistical abstraction of Section II-B, which can also be regarded as a case of model enhancement in process mining [26]. Specifically, A and E are equal to the sets A and E, respectively, while P_{x}, P_{y}, and P_{z} are equal to the statistics P_{x}, P_{y}, and P_{z}, respectively. Then, T_{o} is derived from A, E, P_{x}, P_{y}, and P_{z} as the set that contains: an observable transition $(a, e, a)\ \forall\, a \in A$ and $\forall\, e \in E$ such that $\mathcal{P}_{\mathrm{x}}(e, a) \neq 0$; an observable transition $(a_*, e, a_*)\ \forall\, e \in E$ such that $\mathcal{P}_{\mathrm{y}}(e) \neq 0 \wedge \sum_{a\in A} \mathcal{P}_{\mathrm{z}}(e, q)$; and, an observable transition $(a_*, e, a)\ \forall\, a \in A$ and $\forall\, e \in E$ such that $\mathcal{P}_{\mathrm{z}}(e, a) \neq 0$. Similarly, T_{u} is derived from A as the set that contains an unobservable transition $(a, a_*)\ \forall\, a \in A$.

Finally, the PDFs $F_{\mathrm{d}}(a)$ and $F_{\mathrm{i}}(a)$ of duration and inter-event time of each activity $a \in A_*$ are defined so as to fit expected value μ and coefficient of variation CV of corresponding histograms $\mathsf{F}_{\mathrm{d}}(a)$ and $\mathsf{F}_{\mathrm{i}}(a)$, respectively. To this end, we extend [31]: for histograms with CV $< \frac{1}{\sqrt{2}}$, we replace the Shifted Exponential random variable (r.v.), which would produce a false negative for any value under a firm threshold, with the sum of an Exponential r.v. and an Erlang r.v. with sufficient phases to attain CV lower than samples. Specifically, if CV < 1, the approximating PDF $f(x)$ with $x \in [0, \infty)$ is the PDF of the sum of an Erlang r.v. with shape n and rate λ_1, and an Exponential r.v. with rate λ_2, where $n = \lceil 1/\mathrm{CV}^2 \rceil - 1$, λ_1 is solution of $\mathrm{CV}^2 \mu^2 = \frac{n}{\lambda_1^2} + (\mu - \frac{n}{\lambda_1})^2$, and $\lambda_2 = 1/(\mu - n/\lambda_1)$ (see Section II of Supplemental Material). If CV ≈ 1, $f(x) = \lambda\, e^{-\lambda\, x}$ with $x \in [0, \infty)$ is the PDF of an Exponential r.v. with rate $\lambda = \frac{1}{\mu}$ [31]. If CV > 1, $f(x) = \sum_{i=1}^{2} p_i\, \lambda_i\, e^{-\lambda_i\, x}$ with $x \in [0, \infty)$ is the PDF of a hyper-exponential r.v. with $\lambda_i = \frac{2 p_i}{\mu}$ and $p_i = \frac{1}{2}(1 \pm \sqrt{(\mathrm{CV}^2 - 1)/(\mathrm{CV}^2 + 1)})$ [31].

Note that, in the special case $\frac{1}{\sqrt{2}} < \mathrm{CV} < 1$, the proposed approximating PDF takes the same form as that of [31], and that the sequential composition of an Exponential r.v. and an Erlang r.v. conceptually follows the approach proposed in [35] to fit higher order moments. Finally, note that this class of functions is used here without loss of generality, as all the steps of the method and its implementation can be applied for any general distribution in the class of expolynomial functions.

IV. Classification technique

We define a diagnosis technique for online AR (Section IV-B) and offline AR (Section IV-C). To this end, *transient* analysis of the model of Section III is performed to derive the probability that an activity is performed at a given time conditioned on a sequence of time-stamped observations. The analysis can be efficiently performed through a tailoring of the method of stochastic state classes [29] (Section IV-A).

A. Preliminaries on stochastic state classes

A *stochastic state class* (class for short) represents the continuous set of states that can be reached after a sequence of (observable or unobservable) transitions. Specifically, a class is a tuple $\Sigma = \langle \psi, D_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle}, f_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle}\rangle$, where: $\psi \in A_*$ is an activity; $D_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle} \subseteq \mathbb{R}^2_{\geq 0} \times \mathbb{R}_{\leq 0}$ is the support of the random vector $\langle\theta,\iota,\tau_{\mathrm{age}}\rangle$ made of the remaining sojourn time θ in ψ, the remaining time ι to the next observable event in ψ, and the absolute time τ_{age}; and, $f_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle} : D_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle} \to [0,\infty)$ is the PDF of $\langle\theta,\iota,\tau_{\mathrm{age}}\rangle$ (as discussed in the Supplemental Material, τ_{age} is the opposite of the elapsed time).

A succession relation $\Sigma \overset{t,\mu}{\Rightarrow} \Sigma'$ is defined between classes, meaning that Σ' is the successor of Σ through transition t with probability μ. Enumeration of the succession relation from an initial class Σ_0 yields a *transient tree*, where nodes are classes and edges are labeled with transitions and their occurrence probability. The *reaching probability* $\eta(\Sigma_k)$ of a class Σ_k reached from Σ_0 through $\Sigma_0 \overset{t_1,\mu_1}{\Longrightarrow} \Sigma_1 \ldots \overset{t_k,\mu_k}{\Longrightarrow} \Sigma_k$ is the product of transition probabilities, i.e., $\eta(\Sigma_k) := \prod_{i=1}^{k} \mu_i$. Technical details on the calculus of successor classes are deferred to the Supplemental Material.

We consider an initial class $\Sigma_0 = \langle a_*, D^0_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle}, f^0_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle}\rangle$ where the logical location is the *Idling* activity a_* and the random variables θ, ι, and τ_{age} are independently distributed. Specifically, θ and ι are distributed according to $F_{\mathrm{d}}(a_*)$ and $F_{\mathrm{i}}(a_*)$, respectively, while τ_{age} (measuring the elapsed time) is associated with the Dirac delta function centered at 0, i.e., $D^0_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle} := [F^{\min}_{\mathrm{d},a_*}, F^{\max}_{\mathrm{d},a_*}] \times [F^{\min}_{\mathrm{i},a_*}, F^{\max}_{\mathrm{i},a_*}] \times [0,0]$ and $f^0_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle}(x_\theta, x_\iota, x_{\mathrm{age}}) := F_{\mathrm{d}}(a_*)(x_\theta)\, F_{\mathrm{i}}(a_*)(x_\iota)\, \delta(x_{\mathrm{age}})$.

B. Online classification

At any time t, each activity $a \in A_*$ has a *measure of likelihood* $L^{\mathrm{on}}(a, t, \Sigma_0, \overrightarrow{\omega}_{1,n})$ defined as the probability that a is performed at t conditioned on the initial class Σ_0 and on a sequence of observations $\overrightarrow{\omega}_{1,n}= \langle\omega_1, \ldots, \omega_n\rangle$ taken within $[0,t]$, with $\omega_i = \langle e_i, t_i\rangle \;\forall\, i \in [1,n]$ and $t_1 \leq t_2 \leq \ldots \leq t_n \leq t$: [2]

$$L^{\mathrm{on}}(a, t, \Sigma_0, \overrightarrow{\omega}_{1,n}) := P\{\alpha(t) = a \,|\, \Sigma_0, \omega_1, \ldots, \omega_n\}. \quad (1)$$

The activity with the largest likelihood is the predicted class $\rho^{\mathrm{on}}(t)$, compared with the actual class $\xi(t)$ tagged in the ground truth, i.e., $\rho^{\mathrm{on}}(t) := \arg\max_{a\in A_*} L^{\mathrm{on}}(a, t, \Sigma_0, \overrightarrow{\omega}_{1,n})$.

At time t_0, no event of $\overrightarrow{\omega}_{1,n}$ has been observed yet, and the initial class Σ_0 collects the plausible states (see Section IV-A). According to this, *Idling* has likelihood 1, while any tagged activity has likelihood 0, i.e., $L^{\mathrm{on}}(a_*, t_0, \Sigma_0, \overrightarrow{\omega}_{1,n}) = 1$ and $L^{\mathrm{on}}(a, t_0, \Sigma_0, \overrightarrow{\omega}_{1,n}) = 0 \;\forall\, a \in A$. At any time $t > t_0$, the likelihood of an activity is derived in a different manner depending on whether t is the time of an observation (Section IV-B1) or a time between consecutive observations (Section IV-B2).

1) Classification right after an observation: The likelihood of $a \in A_*$ right after the n-th observation $\omega_n = \langle e_n, t_n\rangle$ is evaluated as the sum of the likelihoods of the classes with logical location a that can be reached from the initial class Σ_0 conditioned on the observation sequence $\overrightarrow{\omega}_{1,n}$:

$$L^{\mathrm{on}}(t_n, a, \Sigma_0, \overrightarrow{\omega}_{1,n}) = \sum_{\Sigma\in\Gamma_n \,|\, \Sigma=\langle a,\cdot,\cdot\rangle} l^{\mathrm{on}}(\Sigma^m_n) \quad (2)$$

where $l^{\mathrm{on}}(\Sigma^m_n) := P\{\Sigma^m_n \,|\, \Sigma_0, \overrightarrow{\omega}_{1,n}\}$ is the probability that Σ^m_n is entered at time t_n conditioned on Σ_0 and $\overrightarrow{\omega}_{1,n}$, and Γ_n is the set of classes reached from Σ_0 through $\overrightarrow{\omega}_{1,n}$. According to the definition of conditional probability and to the law of total probability, $l^{\mathrm{on}}(\Sigma^m_n)$ can be expressed as:

$$l^{\mathrm{on}}(\Sigma^m_n) = \frac{\chi(\Sigma_0, \overrightarrow{\omega}_{1,n}, \Sigma^m_n)}{\displaystyle\sum_{\Sigma^i_n\in\Gamma_n} \chi(\Sigma_0, \overrightarrow{\omega}_{1,n}, \Sigma^i_n)} \quad (3)$$

where $\chi(\Sigma_0, \overrightarrow{\omega}_{1,n}, \Sigma^m_n) := P\{\Sigma_0, \overrightarrow{\omega}_{1,n}, \Sigma^m_n\}$ is a *forward variable* representing the probability that Σ^m_n is reached at t_n through $\overrightarrow{\omega}_{1,n}$ starting from Σ_0. If $\Gamma_n = \emptyset$, classification is restarted from Σ_0, i.e., $\Gamma_n = \{\Sigma_0\}$ and $l^{\mathrm{on}}(\Sigma_0) = 1$.

In the following, we first derive the set of classes Γ_n and then we evaluate $\chi(\Sigma_0, \overrightarrow{\omega}_{1,n}, \Sigma^m_n)$ for each class $\Sigma^m_n \in \Gamma_n$.

1.1) Derivation of set Γ_n of classes after n observations. Γ_0 consists of Σ_0, i.e., $\Gamma_0 = \{\Sigma_0\}$. When $\omega_1 = \langle e_1, t_1\rangle$ occurs, Σ_0 is conditioned on the occurrence of e_1 at t_1, yielding $\Gamma_1 = \{\Sigma^1_1, \Sigma^2_1, \ldots, \Sigma^{M_1}_1\}$, e.g., in Fig. 3, $\Gamma_1 = \{\Sigma^1_1, \Sigma^2_1, \Sigma^3_1\}$. In so doing, when $\omega_n = \langle e_n, t_n\rangle$ is observed, each class in $\Gamma_{n-1} = \{\Sigma^1_{n-1}, \Sigma^2_{n-1}, \ldots, \Sigma^{M_{n-1}}_{n-1}\}$ is conditioned on the occurrence of e_n at t_n, yielding $\Gamma_n = \{\Sigma^1_n, \Sigma^2_n, \ldots, \Sigma^{M_n}_n\}$.

As shown in Fig. 3, it may be the case that: *i)* a class in Γ_{n-1} yields two or more classes in Γ_n, which may occur, for instance, when the class in Γ_{n-1} has the *Idling* location and the occurrence of e_n may represent the continuation of *Idling* or start of multiple annotated activities, (e.g., Σ_0 yields Σ^1_1, Σ^2_1, and Σ^3_1); *ii)* two or more classes in Γ_{n-1} yield the same class in Γ_n, which occurs, for instance, when those classes in Γ_{n-1} have different activity and the occurrence of e_n represents the continuation of *Idling* after the completion of the activity, (e.g., Σ^1_2 and Σ^2_2 yield Σ^2_3); *iii)* a class in Γ_{n-1} yields no class in Σ_n, which occurs when e_n at t_n is not possible in any behavior allowed by that class of Γ_{n-1} (e.g., class Σ^3_1).

Hence, conditioning each class $\Sigma^m_{n-1} \in \Gamma_{n-1}$, with $m \in [1, M_{n-1}]$, on the occurrence of e_n at time t_n yields a (possibly empty) set of classes Γ^{m,e_n,t_n}_{n-1}, so that $\Gamma_n = \cup_{m=1}^{M_{n-1}} \Gamma^{m,e_n,t_n}_{n-1}$. The conditioning is developed in two steps:

- **Conditioning on the observed event e_n.** Each class $\Sigma^m_{n-1} \in \Gamma_{n-1}$, with $m \in [1, M_{n-1}]$, is conditioned

[2] To simplify notation, the type of the i-th event is indicated as e_i instead of $e_{x(i)}$ as in Section II, avoiding the use of an indexing function x.

on the observation of e_n, yielding the set Γ_{n-1}^{m,e_n} of classes that: *i*) are leaves of the *transient tree* enumerated from Σ_{n-1}^{m} up to the next observable events (i.e., enumeration is halted if an observable transition or the unobservable transition (a_*, ϕ) occurs), and *ii*) are reached by an observable transition $(\cdot, e_n, \cdot)$. Note that the tree is *finite*, given that at most two unobservable transitions may occur before any observable event, as remarked in Section III-A. In Fig. 4, the observation of $\omega_2 = \langle e_4, 400\,\mathrm{s}\rangle$ makes plausible only the hypothesis that a_2 is continued, leading to $\Gamma_1^{2,e_4} = \{\Sigma_1^{2\mathrm{d}}\}$.

- **Conditioning on the observed time** t_n**.** In turn, each class $\Sigma_{n-1}^{m,e_n,k} \in \Gamma_{n-1}^{m,e_n}$, with $k \in [1, K_{n-1}^{m,e_n}]$, is conditioned on the observation time t_n by imposing that the time elapsed between the observations ω_{n-1} and ω_n is equal to $t_n - t_{n-1}$ (i.e., $-\tau_{\mathrm{age}} = t_n - t_{n-1}$), yielding a class $\bar{\Sigma}_{n-1}^{m,e_n,k}$. In each class $\bar{\Sigma}_{n-1}^{m,e_n,k}$, θ and ι are independently distributed, so that the support and the PDF of τ_{age} can be reset to $[0,0]$ and $\delta(x_{\mathrm{age}})$, respectively, as proved by the following Theorem. The set of classes Γ_{n-1}^{m,e_n,t_n} is finally derived as $\Gamma_{n-1}^{m,e_n,t_n} = \cup_{k=1}^{K_{n-1}^{m,e_n}} \{\bar{\Sigma}_{n-1}^{m,e_n,k}\}$. For instance, conditioning class $\Sigma_1^{2\mathrm{d}}$ in Fig. 4 on the time of observation $\omega_2 = \langle e_4, 400\,\mathrm{s}\rangle$ yields class Σ_2^2 in Fig. 3.

Theorem 1. *$\forall$ class $\Sigma = \langle \psi, D_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle}, f_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle}\rangle \in \Gamma_n$, with $n > 0$, the random variables θ, ι, and τ_{age} are independently distributed, and the random vector $\langle\theta,\iota,\tau_{\mathrm{age}}\rangle$ has domain $D_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle} = D_\theta \times [F_{\mathrm{i},a_*}^{\min}, F_{\mathrm{i},a_*}^{\max}] \times [0,0]$ and PDF $f_{\langle\theta,\iota,\tau_{\mathrm{age}}\rangle}(x_\theta, x_\iota, x_{\mathrm{age}}) = f_\theta(x_\theta)\, F_{\mathrm{i}}(a_*)(x_\iota)\, \delta(x_{\mathrm{age}})$, where D_θ and f_θ are the marginal domain and the marginal PDF of θ, respectively. If entering class Σ accounts for starting activity ψ, then $D_\theta = [F_{\mathrm{d},\psi}^{\min}, F_{\mathrm{d},\psi}^{\max}]$ and $f_\theta(x_\theta) = F_{\mathrm{d}}(\psi)$. Otherwise (i.e., if entering Σ represents a continuation of ψ), $D_\theta = \{x_\theta \mid x_\theta + \Delta_n \in D_{\theta_{\mathrm{p}}} \cap \{x_{\theta_{\mathrm{p}}} \geq \Delta_n\}\}$ and $f_\theta(x_\theta) = f_{\theta_{\mathrm{p}}}(x_\theta + \Delta_n)/\int_{\Delta_n}^{\infty} f_{\theta_{\mathrm{p}}}(y)dy$, where $\Delta_n = t_n - t_{n-1}$, and $D_{\theta_{\mathrm{p}}}$ and $f_{\theta_{\mathrm{p}}}$ are the marginal domain and the marginal PDF, respectively, of the remaining sojourn time θ_{p} in the class $\Sigma_{\mathrm{p}} \in \Gamma_{n-1}$ (at the previous event) that yields Σ.*

For instance, in Fig. 3, the PDF of θ in Σ_1^1 is $F_{\mathrm{d}}(a_1)(x_\theta)$, while in Σ_2^1 it is $F_{\mathrm{d}}(a_1)(x_\theta + 220)/\int_{220}^{\infty} F_{\mathrm{d}}(a_1)(y)dy$. Theorem 1 is instrumental to prove that Γ_n is *finite* $\forall\, n > 0$.

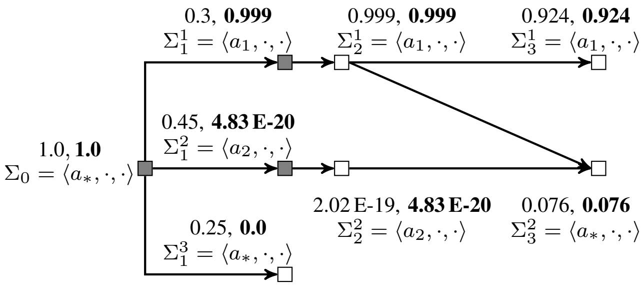

0.3, **0.999**
$\Sigma_1^1 = \langle a_1, \cdot, \cdot\rangle$
0.999, **0.999**
$\Sigma_2^1 = \langle a_1, \cdot, \cdot\rangle$
0.924, **0.924**
$\Sigma_3^1 = \langle a_1, \cdot, \cdot\rangle$
0.45, **4.83 E-20**
$\Sigma_1^2 = \langle a_2, \cdot, \cdot\rangle$
1.0, **1.0**
$\Sigma_0 = \langle a_*, \cdot, \cdot\rangle$
2.02 E-19, **4.83 E-20**
$\Sigma_2^2 = \langle a_2, \cdot, \cdot\rangle$
0.076, **0.076**
$\Sigma_3^2 = \langle a_*, \cdot, \cdot\rangle$
0.25, **0.0**
$\Sigma_1^3 = \langle a_*, \cdot, \cdot\rangle$

Fig. 3. The sets $\Gamma_1 = \{\Sigma_1^1, \Sigma_1^2, \Sigma_1^3\}$, $\Gamma_2 = \{\Sigma_2^1, \Sigma_2^2\}$, $\Gamma_3 = \{\Sigma_3^1, \Sigma_3^2\}$ of plausible classes for the model of Fig. 2 after observations $\omega_1 = \langle e_3, 180\,\mathrm{s}\rangle$, $\omega_2 = \langle e_4, 400\,\mathrm{s}\rangle$, $\omega_3 = \langle e_1, 420\,\mathrm{s}\rangle$, respectively. Each class Σ is labeled with its *online likelihood* $l^{\mathrm{on}}(\Sigma)$ (in plain text) and *offline likelihood* $l^{\mathrm{off}}(\Sigma)$ (in bold), and filled with *gray* or *white* if the corresponding activity is started or continued when the class is entered, respectively.

Theorem 2. *$\forall\, n > 0$, the number of classes in Γ_n is bounded by $Q_{\max} = (|A|+1)\cdot 2^C$, where $|A|$ is the number of annotated activities, and C is the maximum number of consecutive events that may represent the start or the continuation of an activity.*

1.2) Derivation of the forward variable $\chi(\Sigma_0, \vec{\omega}_{1,n}, \Sigma_n^m)$.

Theorem 3. *$\forall\, n > 0$ and $\forall\, \Sigma_n^m \in \Gamma_n > 0$, it holds that:*

$$\chi(\Sigma_0, \vec{\omega}_{1,n}, \Sigma_n^m) = \sum_{\Sigma_{n-1}^i \in \Gamma_{n,m}^{\mathrm{parent}}} \chi(\Sigma_0, \vec{\omega}_{1,n-1}, \Sigma_{n-1}^i) \cdot P\{\omega_n, \Sigma_n^m \mid \Sigma_{n-1}^i\} \tag{4}$$

where: $\Gamma_{n,m}^{\mathrm{parent}} \subseteq \Gamma_{n-1}$ is the set of classes that, conditioned on $\omega_{n-1} = \langle e_{n-1}, t_{n-1}\rangle$, yield Σ_{n-1}^m; $P\{\omega_n, \Sigma_n^m \mid \Sigma_{n-1}^i\}$ is the probability that ω_n is observed and class Σ_n^m is reached at time t_n given that class Σ_{n-1}^i was entered at time t_{n-1}; and, $\chi(\Sigma_0, \vec{\omega}_{1,0}, \Sigma_0) = 1$ is the probability of the initial class.

In turn, $P\{\omega_n, \Sigma_n^m \mid \Sigma_{n-1}^i\}$ can be derived as:

$$P\{\omega_n, \Sigma_n^m \mid \Sigma_{n-1}^i\} = \eta(\Sigma_{n-1}^{i,\mathrm{leaf}}) \int_{I_n} f_{\langle\tau_{\mathrm{age}}\rangle}^{i,\mathrm{leaf}}(x_{\mathrm{age}})dx_{\mathrm{age}}. \tag{5}$$

The first factor is the reaching probability of the class $\Sigma_{n-1}^{i,\mathrm{leaf}}$ that: *i*) is a leaf in the tree enumerated from $\Sigma_{n-1}^i \in \Gamma_{n,m}^{\mathrm{parent}}$ up to any observable event, *ii*) is reached by event e_n, and *iii*) yields Σ_n^m if conditioned on the observation time t_n. The second factor is the integral of the marginal PDF of τ_{age} in $\Sigma_{n-1}^{i,\mathrm{leaf}}$ over $I_n = [\Delta_n - \frac{\delta}{2}, \Delta_n + \frac{\delta}{2}]$, where $\Delta_n = t_n - t_{n-1}$ and δ is the minimum time between consecutive events. Hence, the time of observations is taken into account, e.g., though in the model of Fig. 2 the probability to observe e_4 is larger during a_2 than during a_1, in Fig. 3 the online likelihood of Σ_2^2 is lower than that of Σ_2^1 due to the fact that the probability to observe e_4 in a neighborhood of time 220 s (since the activity start) is significantly larger in a_1 than in a_2.

2) Classification between consecutive observations: In the interval (t_n, t_{n+1}) between observations ω_n and ω_{n+1}, transient analysis [29] is performed from each class $\Sigma_n^m \in \Gamma_n$ up to the next observable events with time limit Δ_n, yielding $\forall\, a \in A_*$ the probability $p(a, t, \Sigma_n^m) := P\{\alpha(t) = a \mid \Sigma_n^m\}$ that a is performed at time $t \in (t_n, t_{n+1})$ conditioned on the

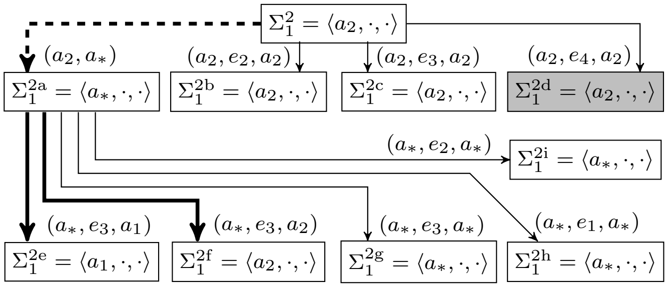

$\Sigma_1^2 = \langle a_2, \cdot, \cdot\rangle$
(a_2, a_*)
(a_2, e_2, a_2)
(a_2, e_3, a_2)
(a_2, e_4, a_2)
$\Sigma_1^{2\mathrm{a}} = \langle a_*, \cdot, \cdot\rangle$
$\Sigma_1^{2\mathrm{b}} = \langle a_2, \cdot, \cdot\rangle$
$\Sigma_1^{2\mathrm{c}} = \langle a_2, \cdot, \cdot\rangle$
$\Sigma_1^{2\mathrm{d}} = \langle a_2, \cdot, \cdot\rangle$
(a_*, e_2, a_*)
$\Sigma_1^{2\mathrm{i}} = \langle a_*, \cdot, \cdot\rangle$
(a_*, e_3, a_1)
(a_*, e_3, a_2)
(a_*, e_3, a_*)
(a_*, e_1, a_*)
$\Sigma_1^{2\mathrm{e}} = \langle a_1, \cdot, \cdot\rangle$
$\Sigma_1^{2\mathrm{f}} = \langle a_2, \cdot, \cdot\rangle$
$\Sigma_1^{2\mathrm{g}} = \langle a_*, \cdot, \cdot\rangle$
$\Sigma_1^{2\mathrm{h}} = \langle a_*, \cdot, \cdot\rangle$

Fig. 4. Transient tree of the model of Fig. 2 enumerated from class $\Sigma_1^2 = \langle a_2, \cdot, \cdot\rangle$ of Fig. 3 up to the next observable events (only observable transitions are shown, using the same graphical notation of Fig. 2). Class $\Sigma_1^{2\mathrm{d}}$ reached through the observation of $\omega_2 = \langle e_4, 400\,\mathrm{s}\rangle$ is filled with light gray.

fact that Σ_n^m was entered at time t_n. Then, the likelihood of activity a is derived from $p(a, t, \Sigma_n^m)$ through Theorem 4.

Theorem 4. *$\forall\, n > 0$ and $\forall\, t \in (t_n, t_{n+1})$, it holds that:*

$$L^{\text{on}}(a, t, \Sigma_0, \overrightarrow{\omega}_{1,n}) = \sum_{\Sigma_n^m \in \Gamma_n} p(a, t, \Sigma_n^m)\, l^{\text{on}}(\Sigma_n^m). \tag{6}$$

C. Offline classification

Given a sequence of observations $\overrightarrow{\omega}_{1,N} = \langle \omega_1, \ldots, \omega_N \rangle$, with $\omega_i = \langle e_i, t_i \rangle\ \forall\, i \in [1, N]$ and $t_1 \le t_2 \le \ldots \le t_N$, at any time $t \in [0, t_N]$ each activity $a \in A_*$ is associated with a *measure of likelihood* $L^{\text{off}}(a, t, \Sigma_0, \overrightarrow{\omega}_{1,N})$ defined as the probability that a is performed at time t conditioned on the overall observation sequence $\overrightarrow{\omega}_{1,N}$, considering not only the observations taken within the interval $[0, t]$ (as in the online classification) but also those taken within the interval $(t, t_N]$:

$$L^{\text{off}}(t, a, \Sigma_0, \overrightarrow{\omega}_{1,N}) := P\{\alpha(t) = a \,|\, \omega_1, \ldots, \omega_N\} \tag{7}$$

As for the online classification, at any time t, the predicted class $\rho^{\text{off}}(t)$ is the activity with the largest measure of likelihood, i.e., $\rho^{\text{off}}(t) := \arg\max_{a \in A_*} L^{\text{off}}(a, t, \Sigma_0, \overrightarrow{\omega}_{1,N})$, derived in a different manner depending on whether t is the time instant of an observation (Section IV-C1) or a time instant between two consecutive observations (Section IV-C2).

1) Classification right after an observation: The likelihood of $a \in A_*$ right after the n-th observation $\omega_n = \langle e_n, t_n \rangle$ is evaluated as the sum of the likelihoods of the classes with logical location a that can be reached from the initial class Σ_0 conditioned on the entire observation sequence $\overrightarrow{\omega}_{1,N}$:

$$L^{\text{off}}(t_n, a, \Sigma_0, \overrightarrow{\omega}_{1,N}) = \sum_{\Sigma \in \Gamma_n \,|\, \Sigma = \langle a, \cdot, \cdot \rangle} l^{\text{off}}(\Sigma_n^m) \tag{8}$$

where $l^{\text{off}}(\Sigma_n^m) := P\{\Sigma_n^m \,|\, \Sigma_0, \overrightarrow{\omega}_{1,N}\}$ is defined as the probability that Σ_n^m is entered at time t_n conditioned on Σ_0 and $\overrightarrow{\omega}_{1,N}$, and can be evaluated through a kind of Forward-Backward procedure [4] according to Theorem 5.

Theorem 5. *$\forall\, n > 0$ and $\forall\, \Sigma_n^m \in \Gamma_n$, it holds that:*

$$l^{\text{off}}(\Sigma_n^m) = \frac{\chi(\Sigma_0, \overrightarrow{\omega}_{1,n}, \Sigma_n^m)\beta(\Sigma_n^m, \overrightarrow{\omega}_{n+1,N})}{\displaystyle\sum_{\Sigma_n^i \in \Gamma_n} \chi(\Sigma_0, \overrightarrow{\omega}_{1,n}, \Sigma_n^i)\beta(\Sigma_n^i, \overrightarrow{\omega}_{n+1,N})}, \tag{9}$$

where $\beta(\Sigma_n^m, \overrightarrow{\omega}_{n+1,N}) := P\{\overrightarrow{\omega}_{n+1,N} \,\,|\Sigma_n^m\}$ is a backward variable defined as the probability of the observation sequence $\overrightarrow{\omega}_{n+1,N}$ conditioned on the fact that Σ_n^m is entered at time t_n.

In turn, $\beta(\Sigma_n^m, \overrightarrow{\omega}_{n+1,N})$ is derived iteratively by Theorem 6.

Theorem 6. *$\forall\, 0 < n < N$ and $\forall\, \Sigma_n^m \in \Gamma_n > 0$, it holds that:*

$$\begin{aligned} &\beta(\Sigma_n^m, \overrightarrow{\omega}_{n+1,N}) \\ &= \sum_{\Sigma_{n+1}^i \in \Gamma_{n,m}^{\text{child}}} \beta(\Sigma_{n+1}^i, \overrightarrow{\omega}_{n+2,N}) \cdot P\{\omega_{n+1}, \Sigma_{n+1}^i \,|\, \Sigma_n^m\} \end{aligned} \tag{10}$$

where: $\Gamma_{n,m}^{\text{child}} \subseteq \Gamma_{n+1}$ is the set of classes yielded by Σ_n^m conditioned on observation ω_{n+1}; $P\{\omega_{n+1}, \Sigma_{n+1}^i \,|\, \Sigma_n^m\}$ is derived through Eq. (5); and, $\beta(\Sigma_N^i, \overrightarrow{\omega}_{N+1,N}) = 1\ \forall\, \Sigma_N^i \in \Gamma_N$.

2) Classification between consecutive observations: It is performed similarly to online classification using Theorem 4.

Theorem 7. *$\forall\, n > 0$ and $\forall\, t \in (t_n, t_{n+1})$, it holds that:*

$$L^{\text{off}}(a, t, \Sigma_0, \overrightarrow{\omega}_{1,N}) = \sum_{\Sigma_n^m \in \Gamma_n} p(a, t, \Sigma_n^m)\, l^{\text{off}}(\Sigma_n^m). \tag{11}$$

D. Complexity

The complexity of CT-HSMMs linearly depends on the number L_{ct} of observed events rather than on the number L of time ticks at which classification is emitted, as in HSMMs. Note that L may be huge to attain sufficient granularity when temporal parameters of the model have different time scale, and typically it is significantly larger than L_{ct}, e.g., in Section V, we consider time ticks of 60 s as in [36] to compare with HSMMs, yielding $L = 40,320$, against $L_{\text{ct}} = 1319$.

After each event, the set Γ' of plausible classes is derived from the set Γ of classes that were plausible after the previous event. For $U \le |A|$ classes in Γ representing the start of an annotated activity, the successor classes in Γ' and the corresponding transition probabilities can be pre-computed. For the remaining $V \ge |\Gamma| - |A|$ classes, the logical location of each successor class in Γ' and the PDF of the remaining sojourn time θ in successor classes representing the start of an annotated activity can be derived a-priori, while all transition probabilities and the PDF of θ in successor classes representing the continuation of an activity must be computed based on the observed time. Given that each class in Γ yields at most $|A|+2$ classes in Γ', the complexity is proportional to $(V - U)(|A| + 2) \approx Q_{\max}|A|$, where $Q_{\max}$ is the maximum number of plausible classes.

Moreover, after each event, transient analysis is performed from each plausible class until the next event. Probabilities are computed in a constant number of time points, yielding linear complexity in $Q_{\max}$. Note that the number of time points could be significantly reduced, e.g., not performing the analysis if the most likely activity after the next event remains the same.

By Theorem 2, $Q_{\max} = (|A| + 1)2^C$, where C is the maximum number of consecutive events being the start or continuation of an activity, yielding overall complexity $O(|A|^2 2^C L_{\text{ct}})$. As typical in continuous-time solution techniques, worst case analysis of complexity provides overly conservative results: Theorem 2 guarantees termination, but in principle indicates an exponential complexity. In practice, the number of concurrent plausible classes turns out to be much lower than $Q_{\max}$, e.g., in the experiments of Section V, the number of plausible classes per day is at most 225 and 69 on average. Note that, to improve scalability, the actual cost could be largely reduced by pruning hypotheses with low likelihood, e.g. discarding classes with likelihood at least one order of magnitude lower than that of other classes with the same logical location.

V. Computational Experience

A. Description of the data set under consideration

The approach is experimented on a publicly available data set [6] containing: 1319 raw events collected by 14 state-change sensors deployed at different locations in a 3-room apartment (e.g., doors and household appliances); 245 activities of 7 distinct types (i.e., *Preparing breakfast*, *Preparing a beverage*, *Preparing dinner*, *Taking shower*, *Toileting*, *Sleeping*, and *Leaving house*) plus *Idling*, performed and manually annotated over a period of 28 days. Activities are annotated in a mutually exclusive way (i.e., one activity at a time), with the only exception of a few instances of *Toileting* nested in *Sleeping* (21 instances) or in *Preparing dinner* (1 instance).

B. Experimental setup

The data set is split into training and test sets according to a *leave one day out* approach, using sensor data of each day for testing and those of the remaining days for training (as discussed in Section II, events natively encoded by sensors using the raw representation are converted into a dual change-point representation). Events observed during the periods of absence of the subject (i.e., *Leaving house*) are removed to avoid inconsistencies (except for *Opening/Closing front door*).

A prototype implementation based on the API of the ORIS tool [37], [38] was used to perform experiments for both the proposed online and offline classification techniques, considering both non-Markovian durations described in Section III-B, and exponentially distributed durations that fit only the expected value. In the latter case, the model has an underlying stochastic process in the class of Continuous Time Markov Chains (CTMC) and becomes conceptually close to CT-HMMs of [14], with a notable difference: using CT-HMMs, each state would be made of an activity and an observed event. This increases the expressivity, permitting the next transition and the rate of the (exponential) sojourn time be dependent also on the last observed event. However, this also increases the number of states, with impact on computational complexity and, more importantly, on the volume of data needed to avoid overfitting. Despite this difference, for notational convenience, in the rest of this Section, CT-HSMM with exponential durations and inter-event times will be referred to as CT-HMM.

Results achieved by CT-HSMMs and CT-HMMs are compared with those attained by two discrete-time generative models, namely offline HMMs and offline HSMMs, running the implementation used in [36] on the considered data set. To permit a fair comparison, the predicted activity $\rho(t)$ is compared with the ground truth activity $\xi(t)$ using time slices of duration of 60 s as in [36]. For each activity $a \in A_*$, we obtain the number of time slices representing *True Positives* (TP_a), *False Positives* (FP_a), and *False Negatives* (FN_a), permitting evaluation of *precision* $\mathrm{Pr}_a = \mathrm{TP}_a/(\mathrm{TP}_a + \mathrm{FP}_a)$ (i.e., number of correctly predicted time slices for activity a out of the total number of time slices predicted for a), and *recall* $\mathrm{Re}_a = \mathrm{TP}_a/(\mathrm{TP}_a + \mathrm{FN}_a)$ (i.e., number of correctly predicted time slices for a out of the total number of time slices of a).

Two aggregated performance measures are considered, i.e., the *accuracy* $\mathrm{Acc} = \sum_{a\in A_*} \mathrm{TP}_a / \sum_{a\in A_*}(\mathrm{TP}_a + \mathrm{FN}_a)$ (i.e., the number of correctly predicted time slices out of the total number of time slices) and the *average recall* $\mathrm{Re} = \sum_{a\in A_*} \mathrm{Re}_a/|A_*|$ (i.e., the average of all recall values).

Moreover, the following classification errors are considered for each activity: *i*) *never recognized* (NR), when all the time slices of the activity are misclassified; *ii*) *lately started* (LS), when the initial time slices of the activity are misclassified but all the subsequent ones are correctly classified; *iii*) *early ended* (EE), when the last time slices of the activity are misclassified but all the previous ones are correctly classified; and, *iv*) *interrupted* (IN), when misclassified and correctly classified time slices are interleaved. In the limited number of cases where an occurrence of *Toileting* is nested in an occurrence of *Sleeping* or *Preparing dinner*, the nested activity is removed from the ground truth, and thus a time-slice is correctly classified if the predicted activity type is equal to *Sleeping* or *Preparing dinner*, respectively.

C. Experimental results

1) Performance measures: Table I compares accuracy and average recall achieved by the proposed techniques and by offline HMMs and offline HSMMs. In particular, the reference method proposed in this paper is offline CT-HSMMs, a *continuous-time* approach which is based on a model with *non-Markovian* durations and implements *offline* classification: it achieves an accuracy of 0.956 and an average recall of 0.685.

Continuous-time vs discrete-time. On the one hand, offline CT-HSMMs can be compared with the corresponding discrete-time approach of offline HSMMs, which achieve an accuracy of 0.761 and an average recall of 0.691. According to this, offline CT-HSMMs attain slightly lower average recall than offline HSMMs, while significantly gaining in accuracy.

Non-Markovian vs Markovian. On the other hand, offline CT-HSMMs can be compared with offline CT-HMMs, which achieve an accuracy of 0.956 and an average recall of 0.648, showing that the use of non-Markovian durations slightly improves average recall with respect to offline CT-HSMMs.

In the discrete-time domain, the comparison between the non-Markovian approach of offline HSMMs and the Markovian approach of offline HMMs yields a counterintuitive outcome: offline HMMs achieve an accuracy of 0.895 and an average recall of 0.739, which definitely outperform the 0.761 and the 0.691 attained by offline HSMMs, respectively. This result may be due to the limited volume of data available for training, which may cause overfitting, and to implementation choices requested by the discrete-time approach, e.g., in [37], time is discretized into ticks of duration equal to 60 s and distributions are represented as histograms with 5 bins.

It is remarkable that performance measures attained by offline HMMs can be compared also with those of offline CT-HMMs, showing that, as in the comparison of offline HSMMs with offlline CT-HSMMs, the continuous-time approach improves accuracy while showing lower average recall.

Offline vs online. Online CT-HSMMs achieve an accuracy of 0.953 and an average recall of 0.668, which compare with 0.956 and 0.685 attained by offline CT-HSMMs, respectively, while online CT-HMMs achieve an accuracy of 0.957 and an average recall of 0.626, which compare with 0.956 and 0.648 attained by offline CT-HMMs, respectively. Intuitively, exploiting only the forward variable reduces the available information and produces a (minor) reduction in performance measures, which however affects average recall but not accuracy.

Remark. Overall, the proposed continuous-time techniques improve accuracy of the considered discrete-time approaches, while showing lower average recall. Notably, the proposed methods provide consistent results: non-Markovian techniques improve on Markovian techniques (i.e., CT-HSMMs improve on CT-HMMs, both in the online and in the offline version), and offline techniques improve on online techniques (i.e., offline CT-HSMMs and offline CT-HMMs improve on online CT-HSMMs and online CT-HMMs, respectively). This result is not achieved by the discrete-time approaches, with offline HMMs definitely outperforming offline HSMMs.

2) Insight and discussion: Table II shows performance measures per class, highlighting particularly low precision values for *Preparing a beverage* and *Preparing dinner*. Confusion matrices, also reported in Table II, show that these wrong predictions mainly depend on low recall values of *Idling*, due to a non negligible number of time slices misclassified as *Preparing a beverage* and *Preparing dinner* (e.g., 10 and 440 for offline CT-HSMMs, respectively). These misclassifications have no significant impact on the recall of *Idling*, which is the third most frequent class (i.e. 4974 time slices in the ground truth), but reduce the precision of *Preparing a beverage* and *Preparing dinner* due to their lower volume (i.e. only 17 and 342 time slices in the ground truth, respectively).

Table II shows also particularly low recall values for *Preparing a beverage* and *Preparing dinner* due to misclassifications as *Idling* (e.g., 4 and 128 time slices of *Preparing a beverage* and *Preparing dinner* are misclassified as *Idling* by offline CT-HSMMs, respectively), and also due to a significant number of time slices of *Preparing a beverage* misclassified as *Preparing dinner* (e.g., 6 for offline CT-HSMMs). Overall, misclassifications of *Preparing a beverage*, *Preparing dinner*, and *Idling* may be due to the ambiguity associated with events that may start or continue multiple activities, e.g., *Opening fridge* may be the initial event of both *Preparing a beverage* and *Preparing dinner*, while *Starting dishwasher* may be a continuing event of both the activities. Moreover, these classification errors may be also due to the fact that the proposed model of CT-HMMs and CT-HSMMs distinguishes the starting event of an activity but not the concluding event.

Table III provides further insight showing the number of classification errors per activity. In more than half of the cases, misclassifications of *Preparing a beverage*, *Preparing dinner*, and *Idling* by online CT-HSMMs are due to NR or IN errors. Though, as expected, the number of IN errors is significantly reduced using offline CT-HSMMs, NR and IN errors comprise a structural limitation of the proposed approach.

Conversely, LS and EE errors inherently depend on the fact that the duration of an activity is evaluated as the time elapsed from its first event to its last event, as illustrated in Section II-B. In principle, these errors could be reduced by characterizing the random variables representing the time elapsed from the start of an activity to its first event and from the last event of an activity to its completion, respectively. Though this model variant is expected to reduce the number of time slices misclassified as *Idling*, it would make the computed statistics dependent on the ground truth, which would be used not only in a qualitative perspective to identify the performed activity, but also in a quantitative perspective to determine its duration. Given that ground truth annotations are structurally imprecise, and a completely supervised data set may not be always available, this solution would weaken the approach.

TABLE I
ACCURACY AND AVERAGE RECALL.

	Accuracy	Average recall
Online CT-HMMs	0.957	0.626
Online CT-HSMMs	0.953	0.668
Offline CT-HMMs	0.956	0.648
Offline CT-HSMMs	0.956	0.685
Offline HMMs	0.895	0.739
Offline HSMMs	0.761	0.691

TABLE II
CONFUSION MATRICES USING CT-HSMMS: (a_i,a_j) IS THE NUMBER OF TIME SLICES OF ACTIVITY a_i PREDICTED AS ACTIVITY a_j.

Online CT-HSMMs									
	Breakfast	Beverage	Dinner	Shower	Toileting	Sleeping	Leaving	Idling	Recall
Breakfast	34	2	1	0	0	7	0	22	0.515
Beverage	0	8	3	0	0	0	0	6	0.471
Dinner	24	9	137	0	4	0	1	167	0.401
Shower	0	0	0	152	5	0	0	63	0.691
Toileting	0	0	1	6	89	10	0	53	0.560
Sleeping	0	0	0	10	23	11549	0	76	0.991
Leaving	0	0	0	3	0	1	22557	6	0.999
Idling	27	8	331	183	60	425	357	3583	0.720
Precision	0.400	0.296	0.290	0.429	0.492	0.963	0.984	0.901	

Offline CT-HSMMs									
	Breakfast	Beverage	Dinner	Shower	Toileting	Sleeping	Leaving	Idling	Recall
Breakfast	35	1	9	0	0	7	0	14	0.530
Beverage	0	7	6	0	0	0	0	4	0.412
Dinner	23	8	179	0	4	0	0	128	0.523
Shower	0	0	0	152	5	0	0	63	0.691
Toileting	0	0	1	1	94	3	0	60	0.591
Sleeping	0	0	0	7	22	11564	0	65	0.992
Leaving	0	0	0	2	0	0	22543	22	0.999
Idling	27	10	440	135	61	302	326	3673	0.738
Precision	0.412	0.269	0.282	0.512	0.505	0.974	0.986	0.912	

TABLE III
NUMBER OF CLASSIFICATION ERRORS USING CT-HSMMS.

	Online CT-HSMMs					Offline CT-HSMMs				
	NR	IN	LS	EE	Total	NR	IN	LS	EE	Total
Breakfast	13	5	5	9	32	15	3	2	11	31
Beverage	0	4	1	4	9	5	2	2	1	10
Dinner	153	12	33	7	205	152	1	5	5	163
Shower	5	11	51	1	68	6	11	48	3	68
Toileting	2	55	13	0	70	3	55	7	0	65
Sleeping	36	1	60	12	109	57	0	24	13	94
Leaving	2	0	8	0	10	1	1	10	12	24
Idling	199	675	119	398	1391	213	396	96	596	1301
Total	410	763	290	431		452	469	194	641	

VI. CONCLUSIONS

We presented a continuous-time approach for online and offline AR in AAL, based on a model where durations are either Markovian (CT-HMMs) or non-Markovian (CT-HSMMs). Experiments performed on a real data set [6] prove feasibility of the approach, showing that offline CT-HMMs achieve better accuracy but lower average recall than offline HMMs, while offline CT-HSMMs attain largely better accuracy and comparable average recall with respect to offline HSMMs.

Achieving comparable performance with respect to state-of-the-art discrete-time machine learning techniques is valuable, since the proposed continuous-time approach is more general, being equivalently applicable when the observed phenomena are asynchronous or have different timescales. In fact, the approach avoids time discretization, which would require the achievement of a trade-off between granularity and complexity, and may become critical when temporal parameters have different orders of magnitude. Moreover, while parameter optimization in machine learning typically requires a large amount of data, model enhancement in the proposed continuous-time setting is less exposed to overfitting, since only few parameters have to be fit (e.g., the first two moments). Indeed also the proposed approach would benefit of the availability of more data, which would enable the derivation of finer approximants.

Insight gained on results opens the way to many extensions. Specifically, misclassifications due to activities characterized by the same events could be reduced extending the model with the notion of *time of day* and using observed time-stamps to derive the current time slot. The model could be extended further with possible *sequences* of initial/concluding events of activities, rather than considering individual initial events only. Classification could be also improved exploiting the set of most likely activities natively provided by the approach, which could be useful if multiple outcomes were plausible. Benefits in activity disambiguation could be also obtained by exploiting the computed statistics to plan new sensor acquisitions.

The approach is open to extension to a fully data-driven model elicitation, which would permit application to cases where domain knowledge is not available or requires a long and expensive acquisition process. Conversely, extension to unsupervised data sets and concurrent activities would require major advancements both in the modeling and in the solution.

REFERENCES

[1] M. R. Alam, M. B. I. Reaz, and M. A. M. Ali, “A review of smart homes - past, present, and future,” *IEEE Trans. Sys., Man, Cyber., Part C (Appl. and Rev.)*, vol. 42, no. 6, pp. 1190–1203, Nov 2012.

[2] J. Ye, S. Dobson, and S. McKeever, “Situation identification techniques in pervasive computing: A review,” *Pervasive and mobile computing*, vol. 8, no. 1, pp. 36–66, 2012.

[3] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, “Sensor-based activity recognition,” *IEEE Trans. on Sys., Man, and Cybernetics, Part C: Appl. and Reviews*, vol. 42, no. 6, pp. 790–808, 2012.

[4] L. Rabiner, “A tutorial on hidden markov models and selected applications in speech recognition,” *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[5] J. Lafferty, A. McCallum, and F. Pereira, “Conditional random fields: Probabilistic models for segmenting and labeling sequence data,” in *Int. Conf. Machine Learn. (ICML)*, vol. 1, 2001, pp. 282–289.

[6] T. van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, “Accurate activity recognition in a home setting,” in *Proc. Int. Conf. on Ubiquitous Computing*, ser. UbiComp ’08. ACM, 2008, pp. 1–9.

[7] P. Rashidi, D. J. Cook, L. B. Holder, and M. Schmitter-Edgecombe, “Discovering activities to recognize and track in a smart environment,” *IEEE Tran Know. Data Eng.*, vol. 23, no. 4, pp. 527–539, 2011.

[8] S. Zhang, S. McClean, B. Scotney, P. Chaurasia, and C. Nugent, “Using duration to learn activities of daily living in a smart home environment,” in *Int. Conf. Perv. Comp. Tech. for Healthcare*. IEEE, 2010, pp. 1–8.

[9] T. Van Kasteren, G. Englebienne, and B. J. Kröse, “Activity recognition using semi-markov models on real world smart home datasets,” *J. Amb. Intell. Smart Env.*, vol. 2, no. 3, pp. 311–325, 2010.

[10] K. P. Murphy, “Dynamic bayesian networks: representation, inference and learning,” Ph.D. dissertation, Univ. of California, Berkeley, 2002.

[11] S. Sarawagi and W. W. Cohen, “Semi-markov conditional random fields for information extraction,” in *Int. Conf. on Neural Information Processing Systems*, ser. NIPS’04. MIT Press, 2004, pp. 1185–1192.

[12] H. Narimatsu and H. Kasai, “State duration and interval modeling in hidden semi-markov model for sequential data analysis,” *arXiv preprint arXiv:1608.06954*, 2016.

[13] U. Avci and A. Passerini, “Improving Activity Recognition by Segmental Pattern Mining,” *IEEE Tran Know. Data Eng.*, vol. 26, no. 4, pp. 889–902, 2014.

[14] W. Wei, B. Wang, and D. Towsley, “Continuous-time hidden markov models for network performance evaluation,” *Performance Evaluation*, vol. 49, no. 1, pp. 129–146, 2002.

[15] C. Krull and G. Horton, “The effect of rare events on the evaluation and decoding of hidden non-markovian models,” in *Proc. Int. Workshop on Rare Event Simulation (RESIM 2008)*, 01 2008.

[16] R. Buchholz, C. Krull, T. Strigl, and G. Horton, “Using Hidden non-Markovian Models to Reconstruct System Behavior in Partially-Observable Systems,” in *Int. Conf. Sim. Tools and Tech.*, 2010, p. 86.

[17] F. Salfner and M. Malek, “Using hidden semi-markov models for effective online failure prediction,” in *Int. Symposium on Reliable Distributed Systems(SRDS)*. IEEE, 2007, pp. 161–174.

[18] L. Chen, C. D. Nugent, and H. Wang, “A knowledge-driven approach to activity recognition in smart homes,” *IEEE Tran Know. Data Eng.*, vol. 24, no. 6, pp. 961–974, June 2012.

[19] L. Chen and C. Nugent, “Ontology-based activity recognition in intelligent pervasive environments,” *International Journal of Web Information Systems*, vol. 5, no. 4, pp. 410–430, 2009.

[20] L. Chen, C. Nugent, and G. Okeyo, “An Ontology-Based Hybrid Approach to Activity Modeling for Smart Homes,” *IEEE Trans. on Human-Machine Systems*, vol. 44, no. 1, pp. 92–105, Feb 2014.

[21] J. Ye, G. Stevenson, and S. Dobson, “A top-level ontology for smart environments,” *Perv. Mob. Comp.*, vol. 7, no. 3, pp. 359 – 378, 2011.

[22] W. Van Der Aalst, A. Adriansyah, A. K. A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. van den Brand, R. Brandtjen, J. Buijs *et al.*, “Process mining manifesto,” in *Business process management workshops*. Springer, 2012, pp. 169–194.

[23] B. F. van Dongen, A. K. A. de Medeiros, H. M. W. Verbeek, A. J. M. M. Weijters, and W. M. P. Van Der Aalst, “The ProM framework: A new era in process mining tool support,” in *Applications and Theory of Petri Nets 2005*. Springer, 2005, pp. 444–454.

[24] W. M. P. van der Aalst, H. A. Reijers, A. J. M. M. Weijters, B. F. van Dongen, A. K. A. De Medeiros, M. Song, and H. M. W. Verbeek, “Business process mining: An industrial application,” *Information Systems*, vol. 32, no. 5, pp. 713–732, 2007.

[25] R. S. Mans, M. H. Schonenberg, M. Song, W. M. P. van der Aalst, and P. J. M. Bakker, “Application of Process Mining in Healthcare - A Case Study in a Dutch Hospital,” in *Biomedical Eng. Sys. and Technologies*. Springer, 2009, pp. 425–438.

[26] A. Rogge-Solti and M. Weske, “Prediction of remaining service execution time using stochastic petri nets with arbitrary firing delays,” in *Service-Oriented Comp.* Springer, 2013, pp. 389–403.

[27] L. Carnevali, C. Nugent, F. Patara, and E. Vicario, “A Continuous-Time Model-Based Approach to Activity Recognition for Ambient Assisted Living,” in *Quant. Eval. Sys.* Springer, 2015, pp. 38–53.

[28] G. Blair, N. Bencomo, and R. B. France, “Models@run.time,” *Computer*, vol. 42, no. 10, pp. 22–27, 2009.

[29] A. Horváth, M. Paolieri, L. Ridi, and E. Vicario, “Transient analysis of non-Markovian models using stochastic state classes,” *Perform. Eval.*, vol. 69, no. 7-8, pp. 315–335, Jul. 2012.

[30] E. Vicario, L. Sassoli, and L. Carnevali, “Using stochastic state classes in quantitative evaluation of dense-time reactive systems,” *IEEE Trans. Softw. Eng.*, vol. 35, no. 5, pp. 703–719, 2009.

[31] W. Whitt, “Approximating a point process by a renewal process, i: Two basic methods,” *Op. Res.*, vol. 30, no. 1, pp. 125–147, 1982.

[32] S. Katz, T. D. Downs, H. R. Cash, and R. C. Grotz, “Progress in development of the index of ADL,” *The gerontologist*, vol. 10, no. 1 Part 1, pp. 20–30, 1970.

[33] H. Choi, V. G. Kulkarni, and K. Trivedi, “Markov regenerative stochastic petri nets,” *Perf. Eval.*, vol. 20, pp. 337–357, 1994.

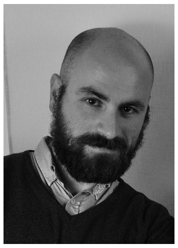
[34] G. Ciardo, R. German, and C. Lindemann, “A characterization of the stochastic process underlying a stochastic Petri net,” *IEEE Trans. Softw. Eng.*, vol. 20, no. 7, pp. 506–515, 1994.

[35] A. Bobbio, A. Horváth, and M. Telek, “Matching three moments with minimal acyclic phase type distributions,” *Stochastic models*, vol. 21, no. 2-3, pp. 303–326, 2005.

[36] T. L. van Kasteren, G. Englebienne, and B. J. Kröse, “Human activity recognition from wireless sensor network data: Benchmark and software,” in *Activity rec. in per. int. env.* Springer, 2011, pp. 165–186.

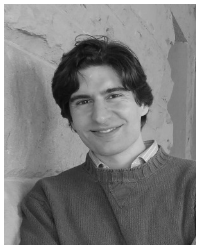
[37] M. Biagi, L. Carnevali, M. Paolieri, and E. Vicario, “An introduction to the ORIS tool,” in *VALUETOOLS*. ACM, 2017, pp. 9–11.

[38] http://www.oris-tool.org.



Marco Biagi received the M.S. in Computer Engineering (2011) and the B.S. in Computer Engineering (2008) from the University of Florence, Italy. He has been a Ph.D. student in the Smart Computing program at University of Florence, Italy. His research activity mainly focused on quantitative approaches for diagnosis, prediction and action scheduling in partially observable systems. He is currently a senior software engineer in a video advertising company.

Laura Carnevali is Assistant Professor of Computer Science (with tenure track) at the School of Engineering, University of Florence, Italy. She received the Ph.D. in Informatics, Multimedia, and Telecommunications Engineering (2010) and the M.S. in Computer Engineering (2006) from the University of Florence, Italy. Her research is focused on correctness verification and quantitative evaluation of real-time and concurrent systems.



Marco Paolieri is a senior research associate at the University of Southern California, Los Angeles. He received his Ph.D. in Computer Science, Systems and Telecommunications (2015) and his M.S. in Computer Engineering (2011) from the University of Florence, Italy. His research interests focus on stochastic modeling and quantitative evaluation of performance and reliability in concurrent and distributed systems.

Fulvio Patara is a postdoctoral research fellow at the University of Florence, Italy, where he received his Ph.D. in Computer Science, Systems and Telecommunications (2016) and his M.S. in Computer Engineering (2011). Since 2018 he is Adjunct Professor of Software Engineering at the School of Engineering, University of Florence, Italy. His research is presently focused on stochastic modeling for Activity Recognition in Ambient Assisted Living and performance evaluation of multi-level meta-modeling architectures applied to e-Health..

Enrico Vicario (M ’95) received the Ph.D. in Informatics and Telecommunications Engineering (1994) and the M.S. in Electronics Engineering (1990) from the University of Florence, Italy. He is Full Professor of Computer Science and Head of the Information Engineering Department at the School of Engineering, University of Florence, Italy. His research is presently focused on model based approaches for development, verification, and quantitative evaluation of concurrent systems with non-deterministic and stochastic temporal parameters.