



HAL
open science

Recognition of Human Activity Based on Probabilistic Finite-State Automata

Kevin Viard, Maria Pia Fanti, Gregory Faraut, Jean-Jacques Lesage

► **To cite this version:**

Kevin Viard, Maria Pia Fanti, Gregory Faraut, Jean-Jacques Lesage. Recognition of Human Activity Based on Probabilistic Finite-State Automata. 22nd IEEE International Conference on Emerging Technologies And Factory Automation, Sep 2017, Limassol, Cyprus. hal-01529721

HAL Id: hal-01529721

<https://hal.science/hal-01529721>

Submitted on 8 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Recognition of Human Activity Based on Probabilistic Finite-State Automata

K. Viard^{1,2}, MP. Fanti², G. Faraut¹, J-J. Lesage¹

¹ LURPA, ENS Cachan, Univ Paris Sud, Université Paris-Saclay
F-94235 Cachan, France

{kevin.viard; gregory.faraut; jean-jacques.lesage}@ens-paris-saclay.fr

²Department of Electrical and Information Engineering, Polytechnic of Bari,
Via Orabona 4, 70125 Bari, Italy
mariapia.fanti@poliba.it

Abstract—Smart home technologies are a promising way to improve health safety of frail people living alone at home. They allow for example on-line recognition of Activities of Daily Living (ADLs) performed by a person, in order to detect dangerous or unusual behaviour. Since human behaviour is not deterministic, probabilistic approaches are often used for ADL recognition, despite difficulties encountered in model building and probabilistic indicators computing. In this paper, it is proposed an approach, based on a Probabilistic Finite State Automata, to detect which activity is being performed. For that a new indicator, called the normalised likelihood, is proposed. The robustness of this indicator to the size of the observed behaviour as well as its computational complexity are also addressed. Finally, the quality of the obtained results are discussed on the basis of an experiment performed in a living lab.

Keywords— *Activity Recognition, Probabilistic Finite-State Automata, Normalised Likelihood, Activity of daily Living, Smart Home*

I. INTRODUCTION

Average age of population in industrialised countries has continuously increased over the last decades and the part of people aged 60 or over will reach 30% in many countries in 2050 [1]. This ageing of the population is both a crucial societal and an important economic problem that makes a priority for research in the field of ubiquitous environments, health technologies and their applications to health at home (HaH). In major studies [2, 3], the knowledge of activities of daily living (ADL) is needed for monitoring the well-being of people living at home. An ADL is generally defined as an activity steadily performed by a person in his/her dwelling (e.g.: preparing meals, doing housework, having a wash, having a nap) [4, 2]. An activity can generally be decomposed into a set of elementary actions; for example, each performance of the activity “preparing meal” consists in a specific combination of actions like taking food in the fridge or in the freezer, washing vegetables at sink, putting a plate in the oven or by the microwave.

Off-line identification, also called discovering, and on-line recognition of ADLs are two major research issues. Both have to be performed from observation of inhabitant behaviour into

a smart home, thanks to a network of sensors of different technologies (motion detectors, cameras and wearable sensors), more or less intrusive for private life.

The goal of Activity Discovery (AD), which is generally performed off-line, is to generate one or several models of activities from learning databases. Images processing technics offer powerful tools for learning ADL models by allowing detecting characteristic poses using cameras [4]. Nevertheless presence of cameras poses numerous problems of acceptance from monitored people. For that reason only non-intrusive sensors will be used in this work. On the other hand, numerous AD approaches are based on wearable sensors [3]. Problems of acceptance are also encountered for this kind of sensor. Furthermore, old people or people suffering for specific pathologies, like Alzheimer, often forget to wear these sensors. Wearable sensors will therefore neither be used in this work. For all these reasons the approaches using binary sensors only, a technology which is not only non-intrusive but also a low cost one, are of special interest. In [5, 6] data mining technics are used for detecting frequent patterns in large sets of observed binary events. The obtained patterns, semantically closed to actions, can be associated with activities, which have not to be known *a priori*. Nevertheless, the use of pattern mining technics makes these approaches suffer of a lack of robustness to the variations of the order of occurrence of events involved in an activity. The labelling allows limiting this drawback but is difficult to put in practice because it has to be done manually. In [2, 7], authors model ADLs by Hidden Markov Models (HMMs) using binary sensors signals. The presented methods require a perfect knowledge of activities actually performed during the learning phase, which is information difficult to obtain with the required completeness and accuracy [8]. In a previous work [9], we proposed an approach of discovering ADLs models by HMMs, but which not requires the knowledge of the activities actually performed during the learning period. In this paper, the same approach is used for modelling ADLs, but under the form of standard Probabilistic Finite-State Automata (PFA).

The goal of Activity Recognition (AR) is to detect on-line the activity which is actually being performed by a person in a

smart home using ADL models. In numerous works [4, 10] video analysis is used as well, but for the same reasons that for AD such approaches cannot be selected within the hypothesis of non-intrusive and non-wearable sensors. When AR is based on HMM models of activities, Viterbi algorithm is frequently used [11]. This is for example the case in [7], where one unique HMM is used for modelling all activities, what leads to a complexity of the model which strongly increases with the number of activities to observe and with the number of considered sensors. Therefore AR methods based on several activity models in which a sequence likelihood computation allows recognition of the performed activity are more efficient [12]. Nevertheless, in these approaches a strong hypothesis is that alphabets of all activity models are the same, what is not verified in practice. Indeed, ADLs are performed in different areas of a dwelling (e.g.: “preparing meals” is performed in the kitchen and “personal hygiene” in the bathroom), they are therefore defined onto different alphabets of sensor events.

In this paper, a new method to recognise activities is presented. Section II proposes a statement of the problem and recalls basis concepts and notations of probabilistic finite-state automata (PFA). Section III defines normalised likelihood allowing comparing sequences with or without shared alphabet of sensor events. A study case is presented in section IV in order to show the efficiency of the proposed method.

II. ACTIVITY RECOGNITION PROBLEM STATEMENT

A. Problem informal definition

Given a set of activities $\{\mathcal{A}_i\}$ and an observation of the behaviour of a person under the form of a sequence $w = e_j \dots e_k$ of observed binary events, the goal of activity recognition is to determine what activity is actually being performed by the person (Fig. 1). Furthermore, the complexity of AR algorithms has to be compatible with their use on-line. Considering the pace of life of a person, a computation time of some tens of seconds is acceptable.

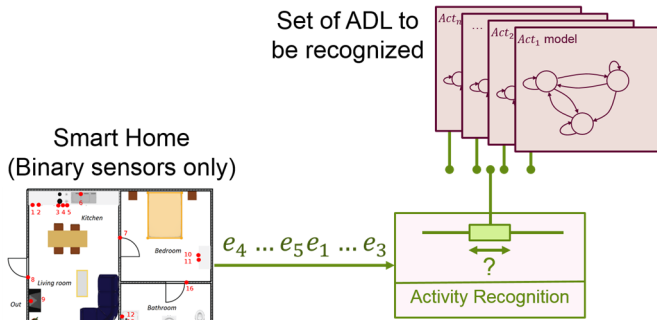


Fig. 1. Activity Recognition objective

In this work, it is assumed that the set of activities \mathcal{A}_i to recognise is fixed by a medical staff. At the end of the AD step, each activity is modelled by a probabilistic finite-state automaton obtained by using the discovery method described in [9].

B. Probabilistic Finite-State Automata : Basis and Notation

Definition 1 (from [13]): A PFA is a tuple

$\mathcal{A} = \langle Q_{\mathcal{A}}, \Sigma_{\mathcal{A}}, \delta_{\mathcal{A}}, I_{\mathcal{A}}, F_{\mathcal{A}}, P_{\mathcal{A}} \rangle$, where:

- $Q_{\mathcal{A}}$ is a finite set of states;
 - $\Sigma_{\mathcal{A}}$ is an alphabet of events;
 - $\delta_{\mathcal{A}} \subseteq Q_{\mathcal{A}} \times \Sigma \times Q_{\mathcal{A}}$ is a set of transitions;
 - $I_{\mathcal{A}} : Q_{\mathcal{A}} \rightarrow \mathbb{R}^+$ is a set of initial-state probabilities;
 - $P_{\mathcal{A}} : \delta_{\mathcal{A}} \rightarrow \mathbb{R}^+$ is a set of transition probabilities;
 - $F_{\mathcal{A}} : Q_{\mathcal{A}} \rightarrow \mathbb{R}^+$ is a set of final-state probabilities;
- $I_{\mathcal{A}}, P_{\mathcal{A}}$ and $F_{\mathcal{A}}$ are such that:

$$\sum_{q \in Q_{\mathcal{A}}} I_{\mathcal{A}}(q) = 1, \quad (1)$$

and:

$$\forall q \in Q_{\mathcal{A}}, F_{\mathcal{A}}(q) + \sum_{a \in \Sigma_{\mathcal{A}}, q' \in Q_{\mathcal{A}}} P_{\mathcal{A}}(q, a, q') = 1 \quad (2)$$

In what follows, \mathcal{A}_i will represent both the i -th activity and its associated PFA. When used in subscript, \mathcal{A}_i represents that the considered element is related to activity \mathcal{A}_i .

C. Problem Redefinition

A classical approach to recognise what activity is performed is to compute the probability for an observed sequence w to be played by each model of activity. The activity, which obtains the maximum probability can be considered as being the activity actually performed. In [14], different probabilistic indicators are presented for the computation of the distance between a set of sequences, called “sample” and denoted as Sp , and the probabilistic language generated by a PFA \mathcal{A}_i . The most used ones are the perplexity $PP(Sp|\mathcal{A}_i)$ and the entropy $\hat{\chi}(Sp, \mathcal{A}_i)$, defined as follows:

$$PP(Sp|\mathcal{A}_i) = \left[\prod_{w \in Sp} P_{\mathcal{A}_i}(w) \right]^{\frac{1}{|Sp|}}$$

$$\hat{\chi}(Sp, \mathcal{A}_i) = -\frac{1}{|Sp|} \sum_{w \in Sp} \log(P_{\mathcal{A}_i}(w))$$

In both cases, the computed value of probability strongly depends on several parameters:

- the observed sequences w with size ($|w|$) and alphabet (Σ_w),
- the activity \mathcal{A}_i with alphabet ($\Sigma_{\mathcal{A}_i}$).

Let us recall that in the general case for two activities \mathcal{A}_i and \mathcal{A}_j , $\Sigma_{\mathcal{A}_i} \neq \Sigma_{\mathcal{A}_j}$ and for an observed sequence w , $\Sigma_w \neq \Sigma_{\mathcal{A}_i}$.

In order to compute the probability for a sequence w to be reproduced by an activity automaton \mathcal{A}_i (often called likelihood- $P_{\mathcal{A}_i}(w)$) despite the differences between their alphabets, it is possible to compute this probability from the projection of w onto $\Sigma_{\mathcal{A}_i}$.

On the other hand, the more a sequence is long, the less high is its probability to be reproduced in an activity \mathcal{A}_i . The comparison between likelihoods associated with a given activity for different observed sequences of different lengths (or projected lengths) require therefore to define a normalised likelihood.

III. NORMALISED LIKELIHOOD

A. –Definition

The normalised likelihood is constructed so that the probability to generate w in \mathcal{A}_i is brought back to the maximum probability to generate a sequence of length $|w|$ in \mathcal{A}_i (3).

$$\begin{aligned} \|\text{Likelihood}(w|\mathcal{A}_i)\| &= \|P(w|\mathcal{A}_i)\| \\ &= \frac{P(w|\mathcal{A}_i)}{\max_{v \in \Sigma_{\mathcal{A}_i}^{|w|}} [P(v|\mathcal{A}_i)]} \end{aligned} \quad (3)$$

Thanks to this normalisation, the value of $\|\text{Likelihood}(w|\mathcal{A}_i)\|$ is between 0 and 1, whatever the length of the considered sequence w is.

Nonetheless the computation of $\max_{v \in \Sigma_{\mathcal{A}_i}^{|w|}} [P(v|\mathcal{A}_i)]$ requires the knowledge of the likelihood of all sequences of length $|w|$ ($1 \leq |w| \leq |w|_{max}$) that can be generated by the PFA. Even if this computation can be done off-line, it could be very long without optimisation. Therefore, a complexity analysis and reduction is presented in the next subsection..

B. Computational issue

The computational complexity of the normalised likelihood is:

$$C_{\|L\|} = \max(C_L, C_M) \quad (4)$$

where:

$$\begin{cases} C_L \text{ the complexity of the classical likelihood} \\ \text{computation} = O(|Q_{\mathcal{A}}|^2 \times |w|) \\ C_M \text{ the complexity of the maximum likelihood} \\ \text{computation} \end{cases}$$

A way to compute the maximum likelihood of a sequence of length $|w|$ in a PFA \mathcal{A} is to compute the likelihood of each possible sequence $v \in \Sigma_{\mathcal{A}}^{|w|}$ of length $|w|$ (for example thanks to the optimised forward algorithm [11]) and to take the maximum value. In this case, the complexity is:

$$C_{\|L\|} = C_M = O(|\Sigma_{\mathcal{A}}|^{|w|} \times |Q_{\mathcal{A}}|^2 \times |w|) \quad (5)$$

Fig. 2, orange curve, shows that in practice this computation time is too high for being acceptable for on-line recognition of ADLs, even by considering classical size sequences ($|w| \leq 10$).

However, the size of this problem can be reduced by performing the computation onto a reduced PFA in which, for each couple of states, only the transition with higher probability is kept. Property 1 expresses that, from one hand the maximum value of the Likelihood for a given length $|w|$ is the same when computed onto the original PFA or onto the reduced PFA (6). On the other hand, the algorithmic

complexity is time linear with respect to the length of w (7). The proof of this property cannot be presented in this paper due to the lack of space. Nevertheless, the idea of the proof is to establish that the right value of $\max_{v \in \Sigma_{\mathcal{A}_i}^{|w|}} [P(v|\mathcal{A}_i)]$ is obtained

by playing in the PFA the sequence (of length $|w|$) computed by concatenation of the most probable outgoing transitions of each state of the PFA.

Property 1:

Let $\mathcal{A} = \langle Q_{\mathcal{A}}, \Sigma_{\mathcal{A}}, \delta_{\mathcal{A}}, I_{\mathcal{A}}, F_{\mathcal{A}}, P_{\mathcal{A}} \rangle$, be a PFA, $\mathcal{A}^r = \langle Q_{\mathcal{A}^r}, \Sigma_{\mathcal{A}^r}, \delta_{\mathcal{A}^r}, I_{\mathcal{A}^r}, F_{\mathcal{A}^r}, P_{\mathcal{A}^r} \rangle$ its reduced PFA and $|w| \in \mathbb{N}^*$,

$$\max_{u \in \Sigma_{\mathcal{A}^r}^{|w|}} [P(u|\mathcal{A}^r)] = \max_{v \in \Sigma_{\mathcal{A}}^{|w|}} [P(v|\mathcal{A})] \quad (6)$$

$$\begin{aligned} C_{\|L\|} &\leq O(|Q_{\mathcal{A}}|^{|w|} \times |Q_{\mathcal{A}}|^2 \times |w|) \\ C_{\|L\|} &\leq O(|Q_{\mathcal{A}}|^{|w|+2} \times |w|) \end{aligned} \quad (7)$$

By comparing (5) and (7), it is evident that computing $\max_{v \in \Sigma_{\mathcal{A}_i}^{|w|}} [P(v|\mathcal{A}_i)]$ on \mathcal{A}_i^r allows in practice a great complexity reduction since in most cases $|Q_{\mathcal{A}}| \ll |\Sigma_{\mathcal{A}}|$. That is what is shown in Fig. 2. This time, this computation time can be considered as acceptable for on-line ADL recognition.

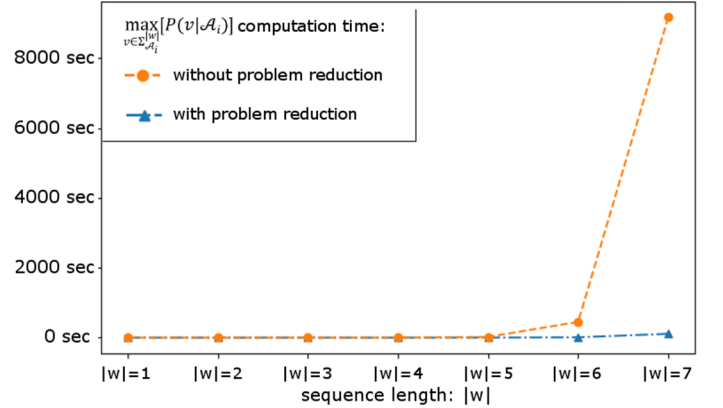


Fig. 2. Example of computation time of $\max_{v \in \Sigma_{\mathcal{A}_i}^{|w|}} [P(v|\mathcal{A}_i)]$ for $1 \leq |w| \leq 7$, for the automaton presented Fig. 6 using a standard laptop with a processor frequency of 2.50GHz. In orange: calculation time without model reduction, in blue: calculation time with reduction.

In order to illustrate the efficiency of the proposed method, next section presents an application performed in a living lab.

IV. CASE STUDY

A. Experimental Protocol

In this paper, for the sake of the brevity, a simple activity \mathcal{A}_{PHB} “preparing hot beverage (PHB)” is considered in the paper. This activity has been performed several times in a

smart flat and the corresponding observed sequences of events have been stored in a database. From this database, a model of this activity has been first built by applying the AD method presented in [9]. Afterwards, the performance of normalised likelihood for ADL on-line recognition has been studied.

Fig. 3 provides a global view of the smart flat equipped with different kinds of binary sensors. A vibration sensor linked to the water carafe is not visible in this picture.

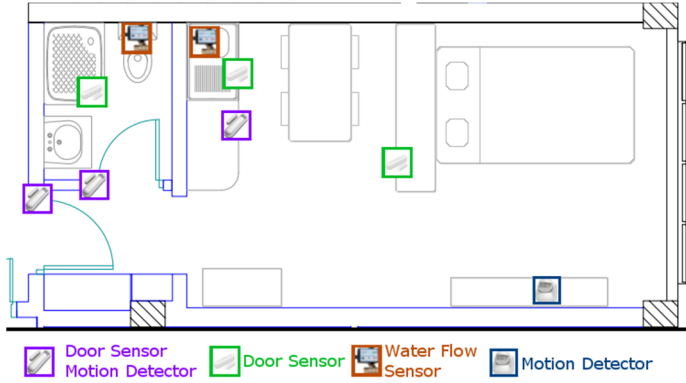


Fig. 3. The smart flat with part of its binary sensors (Sensors linked to elevated cupboards are not represented).

The activity \mathcal{A}_{PHB} is composed of two actions: “preparing coffee” and “preparing tea”. During the tests, these two actions are performed a lot of times and perturbations allowing to evaluate the robustness of the approach are introduced:

- noisy events (not included in $\Sigma_{\mathcal{A}_{PHB}}$) are added, for example by walking in the flat during the activity,
- action performances are interrupted,
- the order of execution of elementary movements composing actions are changed,
- cold water is added to cool down hot coffee.

In order to simulate a significant observation period, eleven instances of activity PHB (\mathcal{A}_{PHB}^j , $1 \leq j \leq 11$), separated by a random number of randomly chosen events not belonging to $\Sigma_{\mathcal{A}_{PHB}}$, are concatenated. The resulting sequence is composed of 545 events corresponding to approximately one hour of observation (Fig. 4).

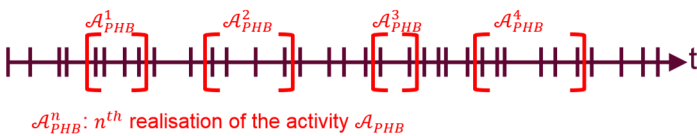


Fig. 4. Structure of the test sequence of observed events.

B. Activity Discovering

Before modelling \mathcal{A}_{PHB} , a decomposition of the activity into actions and an association to the events linked to these actions is needed. Let us recall that in this work it is assumed that activities to be recognised are chosen in order to allow monitoring specific pathologies; they are therefore given by the medical staff. The instrumentation of the smart dwelling is afterwards designed by an expert in order to allow observation of actions composing activities. For the instrumentation depicted in Fig. 3, the obtained decomposition in actions and the associated sensor events is given in Fig. 5. Some events, as those linked to the water flow, are shared between the two actions and can also be shared with actions from other activities.

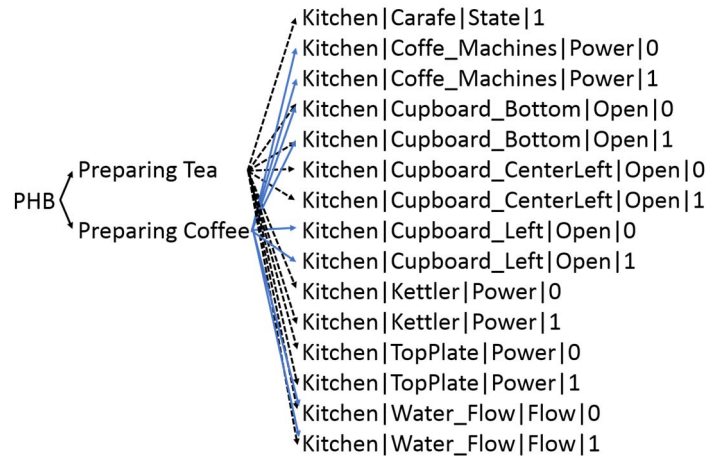


Fig. 5. Preparing Hot Beverage (PHB) activity decomposition

By applying the AD method presented in [9] to the learning database, the PFA presented Fig. 6 is automatically obtained. It is composed of three states: one for each action composing the activity plus one initial state whose outgoing transitions express the initial probability of emission of events associated with actions (destination states of the transitions).

C. Activity Recognition

The on-line observation of events generated by the monitored person constitutes a more or less long sequence in which events are more or less time-distant (Fig. 4). On-line AR requires therefore to observe this sequence through a sliding window of a certain number of events [2] (Fig. 7).

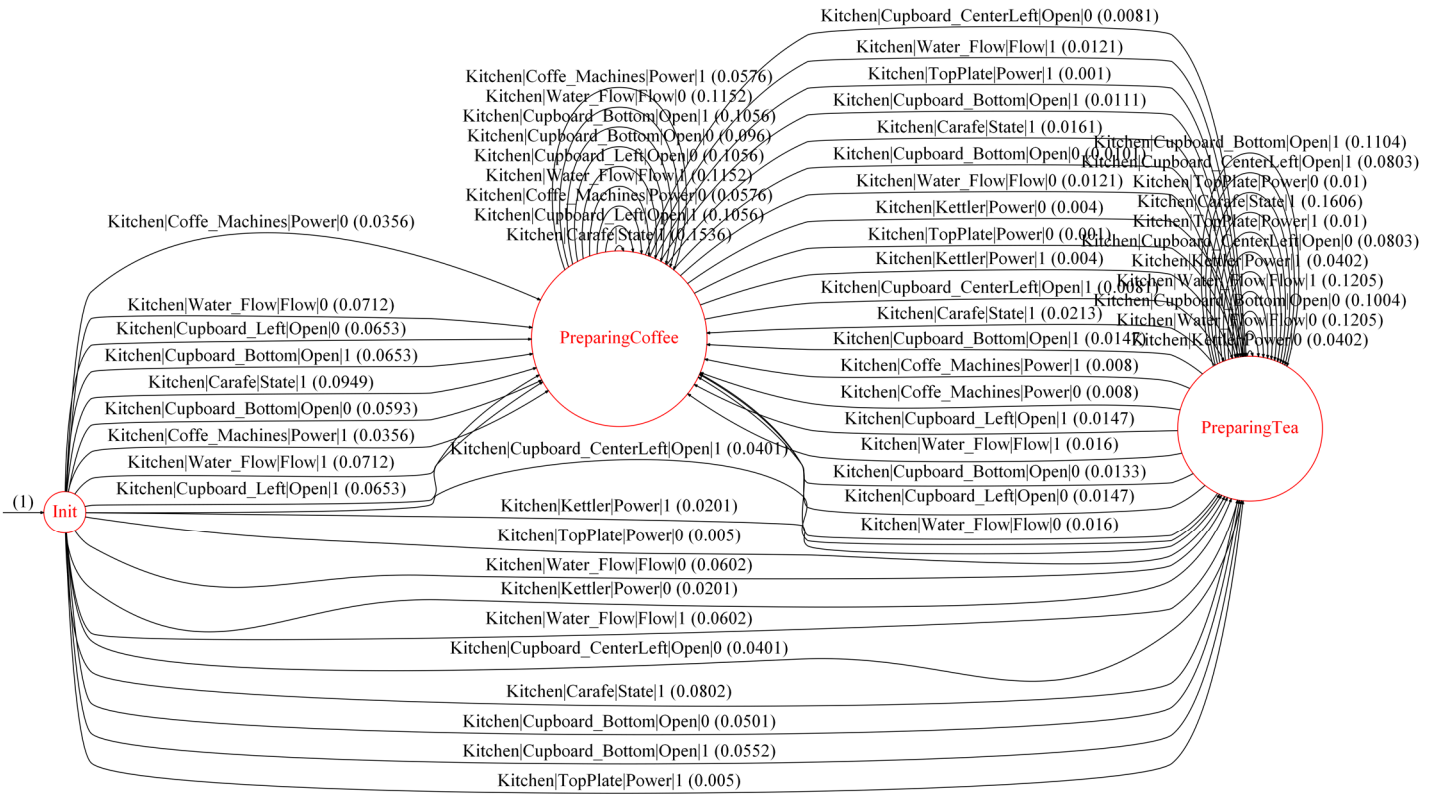


Fig. 6. Preparing Hot Beverage (PHB) activity model.

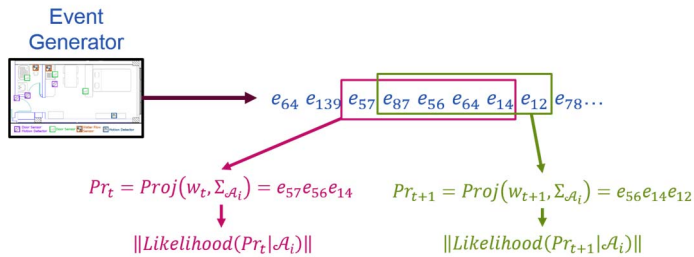


Fig. 7. Sequence projection normalised likelihood computation.

After the influence of the length of this window on the quality of AR is study, the number of events composing the sliding window is fixed to 5. Each time a new event occurs, the new sequence composing the sliding window is considered (Fig. 7). This sequence is first projected onto $\Sigma_{\mathcal{A}_{PHB}}$, in order to filter noisy events (events that not belong to $\Sigma_{\mathcal{A}_{PHB}}$). The computation of the probability of the projected sequence to be generated by the automaton \mathcal{A}_{PHB} is afterwards performed: using the normalised likelihood (Fig. 8) and using the classical likelihood (Fig. 9).

D. Discussion

The comparison of Fig. 8 and Fig. 9 allows underlining some advantages of the use of normalised likelihood:

- Probabilities computed thanks to classical likelihood are lower than 0.15; the range of

probabilities computed thanks to normalised likelihood is between 0.0 and 0.8. These values are more meaningful and easier to be interpreted by an expert. They allow for example the use of thresholds on the probability value for deciding when an activity is performed.

- The spectrum of PHB (Fig. 9) is composed of peaks (corresponding to the observation of the first event belonging to $\Sigma_{\mathcal{A}_{PHB}}$) followed by a quasi-null signal (because the probability, computed by the standard likelihood, of a sequence w to be generated by \mathcal{A}_{PHB} decreases very fast when $|w|$ increases). The spectrum of PHB (Fig. 8) reveals a more progressive evolution of PHB when computed thanks to normalised likelihood (see for example \mathcal{A}_{PHB}^2 or \mathcal{A}_{PHB}^4). This form of signal graphically expresses well the continuity of the performance of activity through the evolution of the probability value.

Furthermore, the spectrum of PHB when computed thanks to normalised likelihood will also allow comparing the probabilities of different ADL execution in further work in which the problem will be to distinguish which activity is actually being performed among a set of several activities sharing multiple events.

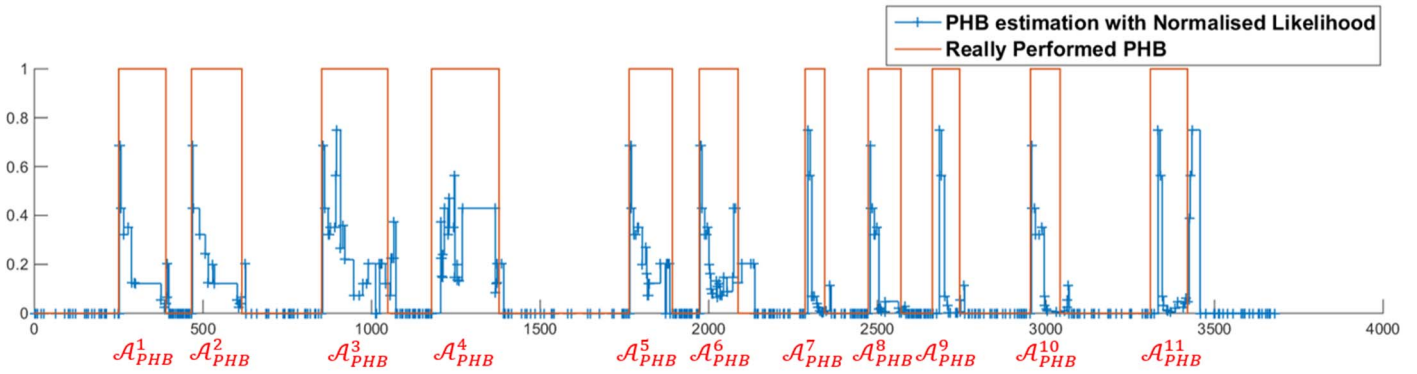


Fig. 8. In red : value = 1 when the activity \mathcal{A}_{PHB} is performed, in blue : *normalised likelihood* of the projected sequence at the instant t (in second).

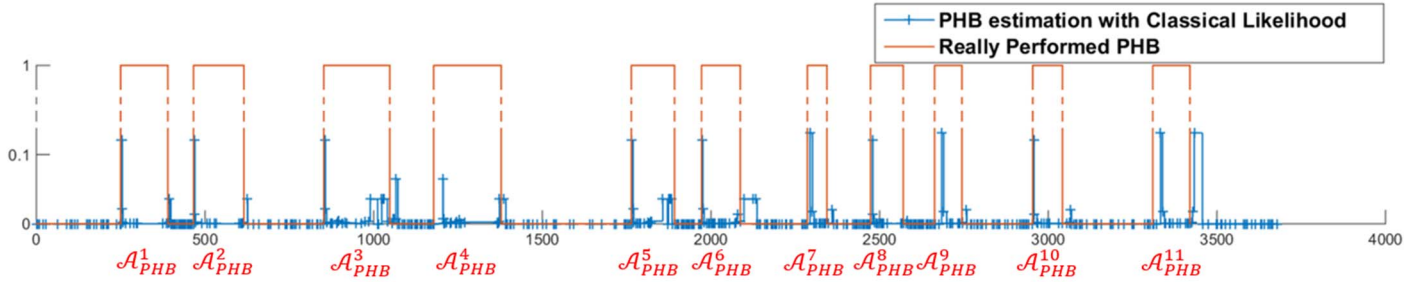


Fig. 9. In red: value = 1 when the activity \mathcal{A}_{PHB} is performed, in blue: *classical likelihood* of the projected sequence at the instant t (in second).

V. CONCLUSIONS

In this paper, an approach for Activity of Daily Living Recognition has been proposed. To do that, an estimator, called normalised likelihood, has been defined. It has also been shown that the normalised likelihood can be efficiently computed onto a reduced model of the activity without loss of accuracy. This makes efficient the on-line computation of the probability for an activity, modelled by a Probabilistic Finite-State Automaton, to be performed. Finally, the proposed approach has been applied to a test activity performed in a smart flat and the quality of the results obtained has been discussed.

Future works will be devoted to the improvement of the presented method and to the detection, if exist, of the drifts in the activities realisation.

ACKNOWLEDGMENT

The research of K. VIARD is supported by the French-Italian University, VINCI program 2016.

The authors would also like to thank ENS Paris-Saclay, which places the flat to their disposal.

REFERENCES

- [1] W. H. Organization *et al.*, “Good health adds life to years: Global brief for world health day 2012,” 2012.
- [2] D. J. Cook and N. C. Krishnan, *Activity Learning: Discovering, Recognizing, and Predicting Human Behavior from Sensor Data*. John Wiley & Sons, 2015.
- [3] O. D. Lara and M. A. Labrador, “A survey on human activity recognition using wearable sensors,” *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 3, pp. 1192–1209, 2013.
- [4] A. A. Chaaraoui, P. Climent-Pérez, and F. Flórez-Revuelta, “A review on vision techniques applied to human behaviour analysis for ambient-assisted living,” *Expert Systems with Applications*, vol. 39, no. 12, pp. 10873–10888, 2012.
- [5] J. Saives, C. Pianon, and G. Faraut, “Activity discovery and detection of behavioral deviations of an inhabitant from binary sensors,” *Automation Science and Engineering, IEEE Transactions on*, vol. 12, no. 4, pp. 1211–1224, 2015.
- [6] R. Agrawal and R. Srikant, “Mining sequential patterns,” in *Data Engineering, 1995. Proceedings of the Eleventh International Conference on*. IEEE, 1995, pp. 3–14.
- [7] T. Van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, “Accurate activity recognition in a home setting,” in *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 2008, pp. 1–9.
- [8] E. M. Tapia, S. S. Intille, and K. Larson, *Activity recognition in the home using simple and ubiquitous sensors*. Springer, 2004.
- [9] K. Viard, M. Fanti, G. Faraut, and J.-J. Lesage, “An Event-Based Approach for Discovering Activities of Daily Living by Hidden Markov Models,” in *15th IEEE Int. Conf. on Ubiquitous Computing and Communications, (IUCC 2016)*, Granada, Spain, Dec. 2016, pp. 85–92. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01378583>
- [10] P.-C. Chung and C.-D. Liu, “A daily behavior enabled hidden markov model for human behavior understanding,” *Pattern Recognition*, vol. 41, no. 5, pp. 1572–1580, 2008.

- [11] L. Rabiner, “A tutorial on hidden markov models and selected applications in speech recognition,” *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb 1989.
- [12] T. Duong, D. Phung, H. Bui, and S. Venkatesh, “Efficient duration and hierarchical modeling for human activity recognition,” *Artificial intelligence*, vol. 173, no. 7-8, pp. 830–856, 2009.
- [13] E. Vidal, F. Thollard, C. De La Higuera, F. Casacuberta, and R. C. Carrasco, “Probabilistic finite-state machines - Part I.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 7, pp. 1013–1025, May 2005, 12. [Online]. Available: <https://hal-ujm.archives-ouvertes.fr/ujm-00326243>
- [14] —, “Probabilistic finite-state machines - Part II.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 7, pp. 1026–1039, May 2005. [Online]. Available: <https://hal-ujm.archives-ouvertes.fr/ujm-00326250>