

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

ProCAVIAR: Hybrid Data-Driven and Probabilistic Knowledge-Based Activity Recognition

CLAUDIO BETTINI¹, (Senior Member, IEEE), GABRIELE CIVITARESE¹ (MEMBER, IEEE), DAVIDE GIANCANE², AND RICCARDO PRESOTTO.¹

¹University of Milan, Milan, Italy (e-mail: {claudio.bettini,gabriele.civitarese,riccardo.presotto}@unimi.it)

²University of Milan, Milan, Italy (e-mail: davide.giancane@studenti.unimi.it)

Corresponding author: Gabriele Civitarese (e-mail: gabriele.civitarese@unimi.it).

ABSTRACT

The recognition of physical activities using sensors on mobile devices has been mainly addressed with supervised and semi-supervised learning. The state-of-the-art methods are mainly based on the analysis of the user's movement patterns that emerge from inertial sensors data. While the literature on this topic is quite mature, existing approaches are still not adequate to discriminate activities characterized by similar physical movements. The context that surrounds the user (e.g., semantic location) could be used as additional information to significantly extend the set of recognizable activities. Since collecting a comprehensive training set with activities performed in every possible context condition is too costly, if possible at all, existing works proposed knowledge-based reasoning over ontological representation of context data to refine the predictions obtained from machine learning. A problem with this approach is the rigidity of the underlying logic formalism that cannot capture the intrinsic uncertainty of the relationships between activities and context. In this work, we propose a novel activity recognition method that combines semi-supervised learning and probabilistic ontological reasoning. We model the relationships between activities and context as a combination of soft and hard ontological axioms. For each activity, we use a probabilistic ontology to compute its compatibility with the current context conditions. The output of probabilistic semantic reasoning is combined with the output of a machine learning classifier based on inertial sensor data to obtain the most likely activity performed by the user. The evaluation of our system on a dataset with 13 types of activities performed by 26 subjects shows that our probabilistic framework outperforms both a pure machine learning approach and previous hybrid approaches based on classic ontological reasoning.

INDEX TERMS activity recognition, probabilistic reasoning, mobile computing

I. INTRODUCTION

Nowadays, mobile devices are increasingly capable of sensing and reasoning. This continuous evolution enables the development of intelligent context-aware applications also based on the recognition of human activities [1].

In mobile computing, activity recognition has been mainly tackled with supervised machine learning approaches on inertial sensors data [2] and more recently with semi-supervised learning [3]. While those data-driven approaches generally lead to high recognition rates considering few physical activities, their effectiveness on complex and context-dependent activities is still unclear. Moreover, discriminating

activities with similar movement patterns is still a challenge. For instance, activities like *walking* and *taking the stairs*, or *standing* and *standing on a bus* are easily confused between them by purely statistical methods based on inertial sensors.

The context that surrounds the user (e.g., semantic location, weather, traffic condition, speed, etc.) is valuable information to mitigate these issues [4], [5]. However, especially considering semi-supervised learning settings, it is not realistic to acquire a comprehensive labeled dataset that includes the large number of possible context conditions in which activities can be performed. Moreover, since the number of context variables may be high and dynamic, the resulting

machine learning model would be extremely complex.

The common-sense knowledge about the relationships between context and activities can be represented through formal models (e.g., ontologies) [6]. Hence, hybrid approaches have been proposed to refine and improve the recognition rate of purely data-driven solutions using semantic reasoning on context data. [7]. Following this promising research direction, we recently proposed CAVIAR [8], a hybrid semi-supervised and knowledge-based activity recognition framework. The main drawback of those hybrid approaches is that they are based on knowledge-based models that use rigid rules to model only the most common scenarios to exclude from the final prediction activities performed in unlikely (but still possible) context conditions. For instance, CAVIAR assumes that the user can not perform certain activities in certain semantic locations (e.g., running in an indoor environment). Such rigid rules inevitably lead to recognition errors when confronted with unusual but entirely possible real-life scenarios. Indeed, the domain of human activities is complex and the relationships between a type of activity and the context in which it can be performed are inherently uncertain [9].

The above-mentioned limitations of hybrid knowledge-based and data-driven methods highlight an open research problem: how to include uncertainty when modeling semantic relationships between activities and context? How to use this probabilistic knowledge to significantly improve the recognition rate obtained by state of the art machine learning methods?

In this paper, we tackle this research problem by proposing ProCAVIAR (Probabilistic Context-aware ActiVe and Incremental Activity Recognition), a novel activity recognition framework that combines semi-supervised learning and probabilistic context-aware reasoning. An incremental machine learning classifier is in charge of inferring from inertial sensor data a candidate probability distribution over the possible activities. A probabilistic knowledge-based reasoning engine is then used to refine the probability distribution considering context-data. The output is a context-refined probability distribution. ProCAVIAR also uses active learning to continuously improve the semi-supervised machine learning model.

Differently from CAVIAR [8], that uses rigid ontological reasoning for context-refinement, ProCAVIAR proposes an original application of probabilistic ontologies, obtaining a more realistic model of common knowledge by capturing the probabilistic relationships between activities and the context in which they are performed.

Our contributions are threefold:

- We propose a novel framework to model context data based on a probabilistic ontology.
- We propose a technique that combines machine learning and probabilistic reasoning to overcome known issues of activity recognition based on machine learning.
- We evaluate our system through extensive experiments, showing the potential benefits of ontological probabilis-

tic reasoning in hybrid systems for sensor-based human activity recognition.

II. RELATED WORK

Activity recognition using mobile devices has been mainly addressed with supervised machine learning methods [1], [2], [10]. The major drawback of those approaches is that they require the acquisition of a large amount of labeled data to initialize the recognition model.

For this reason, semi-supervised learning is emerging as a powerful tool to initialize the recognition model with few labeled data points and to continuously improve it over time [3], [11]–[13]. Among the many semi-supervised techniques (e.g., self-learning, co-training), active learning is one of the most effective [14]–[17]. Existing semi-supervised activity recognition methods in mobile computing mainly recognize a restricted number of physical activities using inertial sensors data. Moreover, discriminating activities with similar motion patterns is still challenging.

The information about the context that surrounds the user (e.g., semantic location, weather condition, time of the day, etc.) can be used to significantly expand the set of considered activities and to better discriminate activities with similar motion patterns that are generally executed in different context conditions (e.g., sitting and sitting on a bus) [8].

However, the acquisition of a comprehensive training set where activities are performed in all the possible context conditions is prohibitive. The abstraction ability of common-sense knowledge can be used to generate formal models representing the relationships between context and activities [5]. Several approaches have been proposed in the literature to formally represent context data [9]. Ontologies have been preferred over other formalisms mainly for their expressive power and automatic reasoning capabilities [6], [18], [19]. There are several well-known ontologies that propose a formalism for context and activities, like *SOUPA* [20], *MetaQ* [21], and the so called *foundational ontology* [22]. In our work we extend the *Activo* ontology proposed in [5].

The combination of ontological context reasoning tools and machine learning algorithms on inertial sensors data has been explored before. Banos et al. [23], proposed the integration of machine learning, used to derive low-level activities, with ontological reasoning, used to infer higher-level context based on the derived activities and other context sources (e.g., mood, semantic location). ProCAVIAR has a different goal since it is focused on improving the predictions, but we foresee an interesting application to that framework: our ontological probabilistic approach could replace the rigid rules of the standard ontology in [23], gaining the ability to deal with uncertainty, and hence making it more flexible.

Ontological reasoning has also been used to integrate context data derived from machine learning processes in complex industrial IoT scenarios [24]. This is a typical application of ontologies, particularly useful when data is gathered by different sources and organizations. However, this type of

combination of the two technologies differs from ours, since its focus is not on improving the activity recognition rate.

Closer to our approach are previous works on hybrid methods, in which the machine learning prediction is refined by knowledge-based reasoning over context data [5], [8]. Indeed, semantic reasoning can exclude from the prediction those activities that are highly unlikely according to the current context (e.g., if the user is at the park, it is highly unlikely that she is brushing teeth). The major issue of those hybrid solutions is the rigidity of the ontological formalism that is based on logical rules that despite offer the power of abstraction, struggle to capture the probabilistic nature of the relationship between context and activities.

There is a vast literature on logic formalisms that support some form of uncertainty reasoning. Some efforts also specifically for applications in the area of pervasive computing [25]. Considering description logics as the underlying logics of ontologies, an integration with fuzzy logic has been proposed to express confidence values for each axiom [26]. ProCAVIAR is not based on fuzzy reasoning but on probabilities, since it associates to each possible activity the probability of occurring in certain context conditions.

One of the well known formalisms that combine logic with probability theory is *Markov Logic Network* (MLN) [27]. MLN can model both hard and soft constraints using weights associated to each rule. Generally, the weights associated to soft constraints are learned from labelled data. MLNs have been proposed for smart-home activity recognition [28]. However, they are less suitable than ontologies to model the complex hierarchical relationships between context data and activities.

More recently, *probabilistic ontologies* have been proposed. Examples of such ontologies are *PR-OWL* [29], [30], *DISPONTE* [31], [32] and *Log-linear Description Logics* [33]. Such tools combine the expressiveness of description logics with probabilistic reasoning, and hence they are potentially suitable for ProCAVIAR. PR-OWL relies on OWL2 as a high-level overlay for multi-entity Bayesian networks. Its major drawback is that extending an existing ontology into a PR-OWL one requires significant engineering efforts, since it requires to map ontological concepts to probabilistic entities in the underlying Bayesian network. Since re-using and extending existing ontologies is desirable in our domain, PR-OWL is not suitable. DISPONTE extends description logics with probabilistic reasoning using Binary Decision Diagrams (BDDs) to compute the probability of a query based on the explanations generated by the reasoner. However, DISPONTE is based on the closed-world assumption and it considers axioms as independent and identically distributed random variables, which is unrealistic in our domain. Log-linear description logics integrate description logics with probabilistic log-linear models to compute a probability distribution over coherent and consistent worlds. ProCAVIAR relies on log-linear description logics, since it is the most suitable probabilistic modeling tool for our target domain.

To the best of our knowledge, the only application of prob-

abilistic ontologies to activity recognition has been to smart-home activities [34]. In this work, we consider different types of sensors and activities. Moreover, the way ProCAVIAR represents knowledge and exploits the reasoning capabilities of the ontology is also very different.

III. OVERALL ARCHITECTURE OF PROCARIAR

In this section, we describe the overall architecture of ProCAVIAR, as depicted in Figure 1.

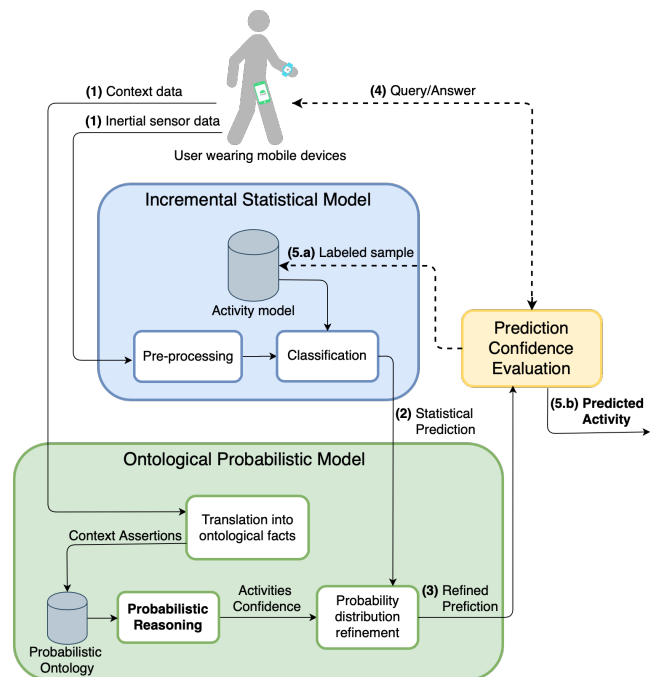


FIGURE 1: Overall architecture of our system

Inertial sensors data (e.g., accelerometer, gyroscope, and magnetometer) coming from multiple mobile devices are processed by the Incremental Statistical Model that relies on an incremental semi-supervised classifier to produce a candidate probability distribution over the possible activities. The mobile devices also dynamically acquire context data both exploiting built-in sensors (e.g., GPS, luminosity sensor) and by querying publicly available Web services (e.g., Google APIs to obtain the user's semantic location). Note that, in the literature, "*context*" is a very broad term used to define a situation at different levels of abstractions [35]. For the sake of this work, with context data we denote the information about the environment that surrounds the user (e.g., user's semantic location, temperature, the time of the day) while the user is interacting with the system. Context data are analyzed by the Ontological Probabilistic Model. This module generates confidence values on the possible activities given the current context. These confidence values are then used to refine the statistical prediction. Finally, the Prediction Confidence Evaluation module adopts an active learning strategy based on a dynamic threshold. When the confidence on the refined prediction is lower than this threshold, a query is triggered to

the user in order to obtain the ground truth. Upon receiving usable feedback, the module sends a new labeled sample to the incremental statistical model.

More details about the Incremental Statistical Model and the Prediction Confidence Evaluation module can be found in [8]. In the following, we focus on the Ontological Probabilistic Model.

IV. ONTOLOGICAL PROBABILISTIC MODEL

The core of ProCAVIAR is the Ontological Probabilistic Model, implemented through a *probabilistic ontology* which models relationships between activities and context data. Differently from a standard ontology, it takes into account the intrinsic uncertainty that characterizes these relationships. Context data obtained from the mobile devices are automatically translated into ontological facts, which are then added to the ontology as a description of the *current* context condition. Then, probabilistic reasoning is in charge of inferring, given the current context situation C , a confidence value $\text{conf}(C, A_i)$ for each activity $A_i \in \mathbf{A}$. Intuitively, $\text{conf}(C, A_i)$ estimates the “*semantic compatibility*” of A_i being performed by the user whose current context is C . Finally, these confidence values are used to refine the probability distribution $h(fv)$ derived from inertial sensors data.

The context-refined probability distribution is then forwarded to the Prediction Confidence Evaluation module.

1) Probabilistic ontology modeling

Our framework combines *soft* and *hard* constraints to model the relationships between activities and context. *Hard* constraints capture context conditions that should always be satisfied to consider a given activity as possible. For instance, *Walking* is an activity that requires the user to have a positive speed. On the other hand, *soft* constraints are useful to capture context conditions that are likely to occur when an activity is performed, but not necessarily they have to be verified; this can be captured by associating a certain degree of confidence to the axiom. Intuitively, the highest the confidence and the more value will have the presence of that context for the likelihood of the corresponding activity to occur. For instance, it is more likely that the activity *Running* is carried out in a sunny day rather than in a stormy day. Hence, the confidence value associated with the soft constraint “*running can be performed on a sunny day*” should be high, while the one associated with the soft constraint “*running can be performed on a stormy day*” should be lower.

In ProCAVIAR, we modified the publicly available OWL2 *ActivO* ontology [5] into a probabilistic ontology based on log-linear description logic [36]. A log-linear description logic is characterized by a *CBox* (i.e., Constraint Box) defined as $C = (C^D, C^U)$, where C^D is a set of *hard* axioms and $C^U = \{(c_1, w_{c_1}), (c_2, w_{c_2}), \dots, (c_n, w_{c_n})\}$ is a set of *soft* axioms. Each soft axiom c_i is associated with a real-valued weight w_{c_i} .

The inclusion of an axiom in C^D and C^U is mutually exclusive. C^D is also assumed to be coherent and consistent (i.e., it is not possible to derive inconsistencies). A log-linear description logic relies on a log-linear probability distribution over the *coherent* and *consistent* subsets of the CBox. Each subset of the CBox represents a world that, if coherent and consistent, is associated with a probability computed using the weights of its soft axioms. Incoherent and/or inconsistent subsets of the CBox are considered as impossible. More details about log-linear description logics can be found in [36].

In our probabilistic ontology, activities are explicitly grouped according to context conditions. Examples of these groups could be “activities that can be performed indoor” or “activities that can be performed at a positive speed”. We refer to these groups as *activity characterizations*. Clearly, an activity may belong to more than one characterization. Characterizations provide an abstraction layer that improves the ontology readability. Moreover, characterizations can be used to define mutually exclusive sets of activities. This approach also makes it possible to easily add new context conditions by creating a new characterization and binding it with the desired activities. Figure 2 and Figure 3 show how characterizations are represented in our ontology.

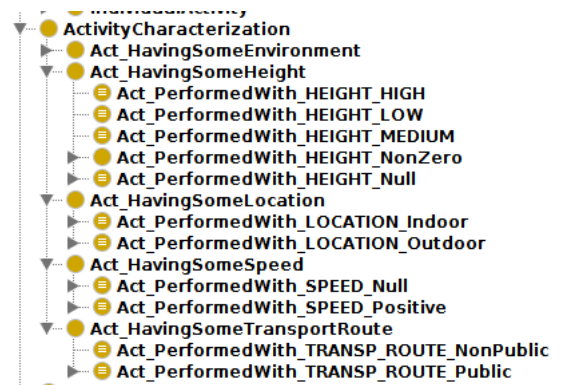


FIGURE 2: A subset of the characterizations in our ontology

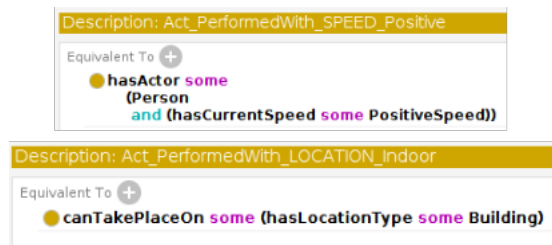


FIGURE 3: Examples of descriptions of characterizations in our ontology

Each characterization is modeled as an equivalence axiom which describes a specific context condition that an activity should satisfy. Therefore, each activity can be modeled in terms of *set membership* to specific context conditions by using subsumption or disjunction axioms with the character-

ization classes.

In order to understand how activities are modeled in our ontology, consider *Running*. It is clear that a person has to move with a positive speed in order to perform this activity. However, other context conditions related to *Running* should be modeled considering soft constraints:

- outdoor/indoor: even if it is more likely that a person is running outside, it is also possible to run inside a building;
- speed: a person may run with different speed rates and each rate has its own probability. Intuitively, a normal running speed rate is the most likely one for this activity, slow running (e.g., jogging) is slightly less probable, while running fast is the least likely one;
- height variation rate: a person may run on a flat or inclined road. Hence, users may run at varying height variations. The most likely scenario is probably running on flat roads.

Due to these considerations, a possible probabilistic modeling of *Running* is depicted in Figure 4, using subsumption and disjunction relationships with the corresponding characterizations. The hard rules are recognizable by the absence of the yellow OWLAnnotation marker, which is enabled on the soft rules. Indeed, the specific log-linear logic that we adopted in our system associates a weight to each soft axiom by using an OWL2 annotation called *confidence*.

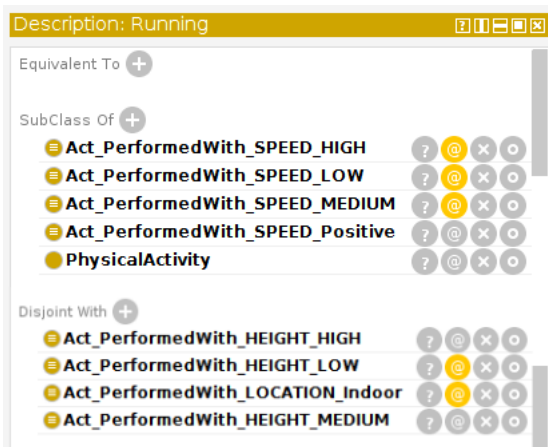


FIGURE 4: Description of *Running* using hard and soft axioms. The soft axioms are the ones associated with the yellow OWLAnnotation marker. Clicking on that marker it is possible to obtain the weight value.

Weighted subsumption axioms are used to describe uncertainty about the different values that a context condition can have. As we show in this example, our ontology includes a weighted subsumption for each possible speed rate related to *Running*.

Note that, for instance, the soft constraint of Figure 4 related to low speed can be formalized as follows:

$$Running \sqsubseteq Act_Performed_With_SPEED_LOW : w_1$$

where $w_1 \in \mathbb{R}$ is the weight associated with this axiom. Later in this section we will discuss how these weights are actually computed. The weight influences the veracity of other axioms related to the same context property (e.g. *Running* can be performed at medium/high speed rates). Therefore, when modelling weighted axioms, we need to pay attention to how a specific weighted axiom influences the others in the reasoning process.

In our model, weighted disjunctions are used to represent uncertainty about context conditions considered both in hard rules and soft rules. For instance, if we model using a weighted subsumption axiom that *Running* may be performed indoor, then *Running* and other outdoor activities would be associated with different output probability values given the same context conditions. This would happen because of the semantics of log-linear DL, which would take into account also the indoor subsumption axioms during the reasoning process of the current user's context, that may specify that the user is outdoor. On the other hand, weighted disjunctions express a degree of incompatibility between an activity and specific context information. In this case, the axiom would be taken into account during the reasoning process only if the current context contains that information. In this example, the weighted disjunction can be formalized as follows:

$$Running \sqcap Act_Performed_With_LOCATION_Indoor \sqsubseteq \perp : w_2$$

where $w_2 \in \mathbb{R}$ is the weight associated with the disjunction.

2) Axioms' Weights

In log-linear description logics, the weight associated with a soft axiom takes values in \mathbb{R} . In the literature, those weights are generally learned from labeled data. In our domain, the acquisition of a comprehensive annotated dataset that includes activities performed in a wide variety of context conditions is prohibitive. In this work, we associate with each axiom a probability value $p \in [0, 1]$ based on common-sense knowledge on context and activities. This knowledge should not necessarily come from the knowledge engineer and domain experts but it may be extracted semi-automatically in several ways, including:

- Proposing a survey to a large number of users;
- Scraping information about context and activities from the Web.

For example, suppose that, according to the common-knowledge, the activity *Running* is not very likely when performed in indoor environments. Hence, according to common-sense knowledge, our system associates the probability value 0.3 to the soft axiom "running can be performed indoor", while 0.7 to the soft axiom "running can be performed outdoor".

Note that directly using probability values as weights associated with soft axioms is not a good choice given the underlying log-linear probability distribution. Hence, as

proposed in other works [37], [38], we use the *logit* function to map each probability value p to a real number as follows:

$$\text{logit}(p) = \log(p) - \log(1 - p) = \log\left(\frac{p}{1 - p}\right)$$

The advantage of using *logit* is that it can approximate probability values for a log-linear model. Note that *logit* is not defined at 0 and at 1. When $p = 1$ or $p = 0$ we consider the axiom as a *hard* constraint. In the former case, it is a context condition that is always required for the corresponding activity; in the latter case, it is a context condition that should never occur.

3) Probabilistic reasoning

ProCAVIAR uses the previously described probabilistic ontology to compute, given the current context data, a confidence value for each activity. First, context data is translated into ontological facts: class instances and relationships that populate the assertional part of the ontology. Once the ontology has been extended with facts about the current context conditions, it is processed by the probabilistic reasoner ELOG [36]. ELOG is in charge of computing marginal inference to obtain, for each activity A_i , a confidence value. Each confidence value $\text{conf}(A_i, C)$ estimates the compatibility of A_i with the current context condition C .

The marginal inference algorithm implemented in ELOG, called *MisSampler (Minimal Inconsistent Subset Sampler)*, analyzes the entire ontology to generate a posterior probability value for each soft axiom according to the log-linear description logic semantics. In order to compute a posterior probability value for each activity possibly performed by the user, our ontology includes dedicated soft axioms. In particular, there is an additional soft axiom for each activity. Each one of these axioms is declared as subclass of the corresponding activity entity with 0 as default confidence value. According to the log-linear DLs semantics, the posterior probability of these axioms is 0.5 if they do not conflict with other axioms. Indeed, without conflicts, the posterior probability of an axiom c with weight w_c is defined by $\text{alogit}(w_c)$ where $\text{alogit}(\mathbb{R}) \rightarrow (0, 1)$ is the logit inverse function. Figures 5 and 6 show those additional soft axioms and their relationships with the rest of the ontology.

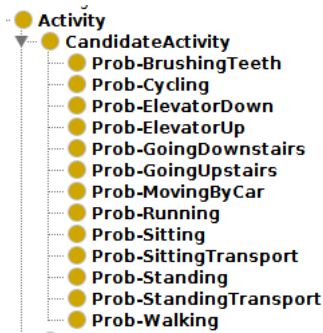


FIGURE 5: Probabilistic terminological overlay



FIGURE 6: CandidateActivity and Prob-Running classes

The output of the marginal inference algorithm is a vector of confidence values:

$$\text{confidences}(C) = \langle c_1, c_2, \dots, c_n \rangle$$

where C is the input context data and $c_i \in \mathbb{R}^+$ is the confidence value $\text{conf}(A_i, C)$ associated to the activity $A_i \in \mathbf{A}$. Note that $\text{conf}(C)$ is not a probability distribution over the activities. Each c_i is a posterior probability value computed by the underlying log-linear probability distribution over coherent and consistent ontologies. Hence, these values should be considered as confidence values associated with the activities given the current context condition. Since we use the default value 0 for the probabilistic axioms in the ontology, the value of each c_i is in the range $[0, 0.5]$ (due to *alogit* as we previously discussed).

Given a confidence value c_i :

- $c_i = 0.5$ reveals that the context satisfies the hard rules related to A_i without the involvement of probabilistic axioms;
- $0 < c_i < 0.5$ reveals that the context satisfies the hard rules of A_i but some soft axioms were used in the inference process, thus decreasing the output confidence. The more soft axioms are involved, the lower the output confidence;
- $c_i = 0$ reveals that the context does not satisfy at least one hard constraint for A_i . Therefore, according to our ontology, the activity is impossible in that specific scenario.

4) Refined activity prediction

The confidence values inferred by ELOG are used to refine the probability distribution obtained from the Incremental Statistical Model on inertial sensor data. In particular, given the probability distribution $\langle p_1, p_2, \dots, p_n \rangle$ and $\text{confidences}(C) = \langle c_1, c_2, \dots, c_n \rangle$ such that A_i is an activity label, p_i is the probability associated to A_i by the statistical model and c_i is the ontological confidence value of A_i given the current context condition, we compute the following vector:

$$v = \langle p_1 * c_1, p_2 * c_2, \dots, p_n * c_n \rangle$$

Hence, confidence values are used as weights associated to the probability values. Finally, the vector v is normalized in order to obtain a probability distribution over the possible activities:

$$\text{predictions} = \langle P_1, P_2, \dots, P_n \rangle$$

such that $\sum_{i=1}^n P_i = 1$ and $P_i \in [0, 1]$. This probability distribution is the output of the Ontological Probabilistic Model and is forwarded to the Prediction Confidence Evaluation module.

V. EXPERIMENTAL EVALUATION

A. DATASET

In order to evaluate the effectiveness of ProCAVIAR, we used the dataset collected in [8]. To the best of our knowledge, there is no other dataset with rich contextual information that can be exploited by ProCAVIAR. The considered activities are the following: *walking*, *running*, *standing still*, *lying*, *sitting still*, *stairs up*, *stairs down*, *elevator up*, *elevator down*, *cycling*, *moving by car*, *sitting on transport*, *standing on transport* and *brushing teeth*. Those activities were executed by 26 volunteers in different context conditions, including working at the office, going around in the city (Milan), driving, using public transportation, cycling, and staying at home. For each user, the dataset includes inertial sensor data and context data collected from a smartphone in the front pocket and a smartwatch on the wrist. Overall, the dataset includes 9 hours of recorded activity data (~ 350 activity instances). More details about this dataset can be found in [8].

B. SIMULATING UNUSUAL SCENARIOS

The dataset that we use to evaluate our system does not include activities executed in context conditions that are unlikely but not impossible in realistic scenarios. For instance, the *Running* activity was never executed in indoor environments and/or with lower speed rates (e.g., jogging). Another example is the *Stairs up* and *Stairs Down* activities, that were never executed outdoor despite it is entirely possible to find stairs outside. Consider a rigid knowledge-based approach to model context data, like the one proposed in [8], where most of the uncommon context conditions for an activity are considered as impossible. Using such a rigid approach in realistic scenarios would negatively impact on the recognition rate. Since we want to quantitatively show that our probabilistic reasoning framework overcomes these drawbacks, we slightly modified the dataset in order to incorporate unusual context scenarios. To the best of our knowledge, there is no public dataset where activities are performed in a wide variety of usual and unusual context conditions.

We implemented a probabilistic simulator for context data which is based on the considered dataset. Hence, we replaced the original context data with simulated context data. For each activity class in the dataset, our simulator considers:

- context information which characterizes the activity regardless of the scenario (i.e., context data needed to satisfy the hard constraints of the ontology);
- a probability distribution over the context data that may be relevant for estimating the probability for the activity (i.e., context data captured by soft constraints in the ontology).

Our simulator relies on a probabilistic representation of the common knowledge of the activity domain to generate possible scenarios for each activity class. For each activity instance in the dataset, the simulator generates, based on the label, a scenario that includes context data related to hard constraints and some of the context data related to soft constraints. The latter are sampled from a probability distribution.

For instance, consider the activity *Walking*. Based on common-sense, our simulator incorporates the following probabilistic knowledge:

- 1) it is very common that users walk slowly (80% of probability), while they sometimes walk faster (20% of probability);
- 2) in the majority of the cases, users walk on flat surfaces, hence with no height variation (70% of probability), while they can walk ascending/descending paths with a lower probability (30% of probability);
- 3) *Walking* can be performed indoor or outdoor with equal probability.

For each activity instance, our simulator generates context data by sampling from these probability distributions. Continuing the example of *Walking*, a wide variety of context scenarios can be generated, like the following ones:

- **Scenario A:** {low speed, no height variation, indoor location}
- **Scenario B:** {low speed, positive small height variation, outdoor location}
- **Scenario C:** {medium speed, no height variation, indoor location}
- **Scenario D:** {medium speed, negative small height variation, outdoor location}

Intuitively, scenario A is the most common one for *Walking* and it would be frequently generated by our simulator. The other examples of scenarios are least common, so they would be rarely generated by the simulator.

C. RESULTS

In the following, we present the results of ProCAVIAR evaluated on the dataset presented in Section V-A with context data simulated as described in Section V-B. The probability values associated to soft axioms in the ontology used for the experiments have been chosen as a result of an internal small survey. In order to evaluate the effectiveness of our technique with respect to alternative approaches, we considered two additional methods. The former does not rely on any kind of semantic refinement procedure and it is called *Data-driven approach*: it combines the incremental activity recognition module and the prediction confidence evaluation module without applying context-refinement. Note that *Data-driven approach* can be considered as a solid baseline, since it is the standard approach in the literature for activity recognition on mobile devices [1]. The latter is the CAVIAR method [8], that uses a state-of-the-art context-refinement procedure based on standard (deterministic) ontologies. This approach does not

take into account the intrinsic uncertainty and incompleteness of common knowledge, and it excludes from the prediction based on inertial sensors those activities that are very unlikely when considering the current context conditions. To the best of our knowledge, there are no other works in the literature that use context data to achieve the same goal. In order to guarantee a fair comparison, in our evaluation we consistently used the same hyper-parameters and classifier proposed in [8].

We performed leave-one-subject-out cross validation to evaluate the recognition rate of our system and the ones of the baselines. Table 1 shows the results in terms of overall F1 score).

Activity	Data-driven approach	CAVIAR [8]	Pro CAVIAR
Elevator up	0.0	0.95	0.95
Elevator down	0.27	0.94	0.94
Moving by car	0.78	0.81	0.81
Brushing teeth	0.77	0.77	0.82
Running	0.98	0.80	0.98
Sitting still	0.94	0.98	0.99
Going upstairs	0.50	0.63	0.86
Going downstairs	0.51	0.67	0.83
Cycling	0.95	0.95	0.97
Standing still	0.84	0.95	0.97
Walking	0.76	0.84	0.94
Sitting transport	0.31	0.86	0.90
Standing transport	0.48	0.94	0.97
Avg F1	0.63	0.86	0.92

TABLE 1: Recognition rate (F1 score) of ProCAVIAR compared with alternative approaches

Our results confirm that context data has a significant impact on the overall recognition rate. Most importantly, the probabilistic context refinement implemented in ProCAVIAR significantly outperforms the baseline methods reaching an overall F1 score of 0.92. Thanks to its probabilistic perspective, our method can recognize activities performed in unusual scenarios, considered as impossible by the deterministic reasoning approach. Looking closely at the results, some of the activities related to the highest improvements are *Going downstairs*, *Going upstairs*, *Walking*, *SittingTransport*, *StandingTransport* and *BrushingTeeth*. For these activities, our simulator generated a wide range of unusual scenarios, thanks to a higher number of combinations of context data with respect to other activities. Thus, the dataset contains more "unusual samples" for those activities respect to the others, which are characterized by a smaller range of possible scenarios.

Considering the solution based on deterministic reasoning, it is possible to observe a significant decrease in the recognition rate of *Running*. Indeed, while this activity can be reliably recognized only by analyzing inertial sensors data, the deterministic semantic refinement often considers it inconsistent considering unusual scenarios. For instance, the deterministic ontology in [8] considers as impossible the fact that *Running* can be carried out indoor, since it is unlikely.

Our method can overcome these problems thanks to soft axioms.

Besides the recognition rate, a crucial evaluation parameter is the number of questions triggered by the system, since it has a significant impact on usability. Figure 7 shows how both *ProCAVIAR* and *Deterministic reasoning* generates a significantly lower number of questions (respectively 6% and 8%) compared to *No context* (22%).

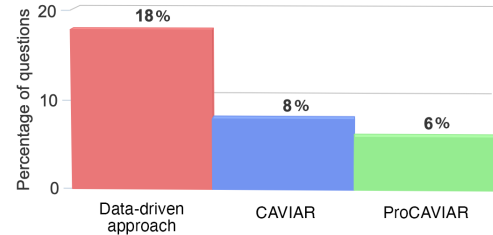
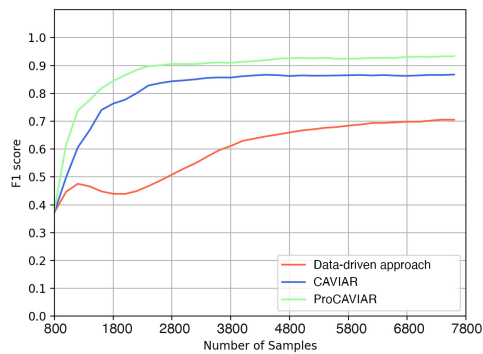


FIGURE 7: Percentage of triggered queries of ProCAVIAR compared with alternative approaches

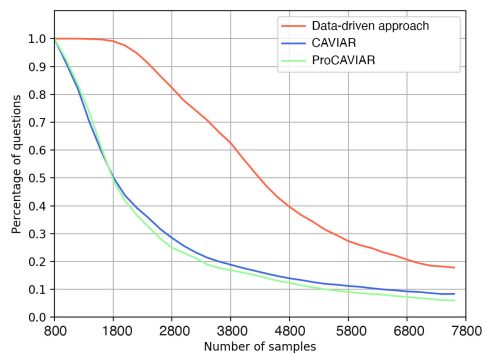
ProCAVIAR slightly decreases the number of user questions with respect to *Deterministic reasoning*. Hence, on average, our probabilistic context refinement method further reduces the uncertainty on the output probability distributions compared to the deterministic solution.

In order to evaluate how the recognition rate and the number of triggered questions evolve over time, we use the evaluation method proposed in [39]. We classify each data sample of the dataset (considering all 26 subjects) with the current model and, depending on the prediction's confidence, we update the recognition model. The classification's output (i.e., the most likely activity), and the corresponding ground truth are collected in sliding windows of 800 samples with an overlap of 75% to periodically compute the overall F-1 score and the percentage of triggered questions. Samples coming from different users are randomly interleaved. Figure 8 shows the evolution of the F-1 score and the number of questions of *ProCAVIAR* with respect to the baselines. Compared to *No Context*, both *Deterministic Reasoning* and *ProCAVIAR* quickly reach high recognition rates and a significantly lower number of questions. *ProCAVIAR* significantly outperforms *Deterministic Reasoning*, showing a faster learning curve. On the other hand, the number of questions generated by *ProCAVIAR* is only slightly lower than the ones generated by the deterministic approach, thus reflecting the results presented in Figure 7.

As we mentioned in Section V-A, to the best of our knowledge there are no other datasets with rich context data that can be used to directly evaluate our context-aware method. Nonetheless, we decided to perform additional experiments on a well known public benchmark for activity recognition: the PAMAP2 dataset [40]. This dataset only includes inertial sensors data gathered from multiple mobile devices. The set



(a) F1 score



(b) percentage of questions

FIGURE 8: Evolution of the recognition model over time. Considered activities: Running, Sitting, Cycling, Standing, Walking, Elevator up, Elevator down, Going Upstairs, Going Downstairs, Brushing Teeth, Moving by car, Sitting transport, Standing transport

of target activities is quite rich and includes the following: *lying still*, *sitting still*, *standing still*, *walking*, *running*, *cycling*, *nordic walking*, *going upstairs*, *going downstairs*, *vacuum cleaning*, *ironing*, *rope jumping*. From the experimental point of view, we compared the *Data-driven component* of ProCAVIAR with a state-of-the-art method based on deep learning [41]. Similarly to our previous experiments, we performed a leave-one-subject-out cross-validation. Table 2 shows that the machine learning module of ProCAVIAR outperforms the baseline confirming that we build our refinement on competitive machine learning techniques.

While we could not quantitatively evaluate the effectiveness of the refinement obtainable by the Ontological Probabilistic Model on the PAMAP2 dataset for the lack of context data, in the following we explain how additional context information could be used by ProCAVIAR to further improve the recognition rate. We first analyzed the confusion matrix of our *Data-driven approach* on the PAMAP2 dataset depicted in Figure 9.

A first observation concerns the activities *Going Upstairs* and *Going Downstairs* that have a lower recognition rate

than others, and are confused among themselves. Since these activities are also part of our dataset, the reader can see in Table 1 the improvement that could be achieved by probabilistic ontological reasoning. By looking at the confusion matrix, we also note that the activities *Sitting still* and *Lying still* are often confused by the classifier. Since *Lying still* is mostly performed in those indoor environments that have beds or sofas (e.g., home environment, hotel rooms, etc.), by considering this knowledge and the semantic location of the user, an ontology-based refinement would be able to improve the recognition rate by excluding that activity as incompatible when the semantic location is some outdoor space. However, even if much less likely, this activity can also be performed in other scenarios, for instance in a park or on the beach on a sunny day (i.e., a weather condition that favors leisure time outside). Hence, ProCAVIAR would be more flexible, defining probabilistic knowledge about semantic locations, time (e.g., it is more likely to lie during the night while sleeping) and weather conditions (e.g., it is less likely to be lying in the park during a rainy day). Similarly, this type of probabilistic reasoning can be applied to other activities that appear to be confused from the matrix, like *Running* and *Rope jumping*; *Ironing* and *Standing still*. There are clear limitations to this approach when activities can be performed in very similar context conditions, like it may be the case for *Standing still* and *Sitting still*.

Activity	Our data-driven approach	LSTM ensembles [41]
Lying still	0.89	0.97
Sitting still	0.83	0.92
Standing still	0.86	0.4
Walking	0.97	0.97
Running	0.95	0.97
Cycling	0.96	0.99
Nordic walking	0.98	0.97
Going upstairs	0.84	0.83
Going downstairs	0.86	0.81
Vacuum cleaning	0.92	0.76
Ironing	0.93	0.89
Rope jumping	0.87	0.76
Avg F1 Score	0.90	0.85

TABLE 2: The results of our *Data-driven approach* on the PAMAP2 dataset compared to a state-of-the-art approach based on deep learning

VI. CONCLUSION AND FUTURE WORK

In this paper we presented ProCAVIAR, a hybrid real-time activity recognition framework based on semi-supervised learning and probabilistic knowledge-based reasoning. Our method applies machine learning algorithms on inertial sensors data to obtain a candidate probability distribution over the activities possibly carried out by the users. Then, a probabilistic ontology that captures probabilistic relationships between context data and activities is in charge of refining the candidate prediction using available context data (e.g., the semantic location of the user, its speed, weather conditions, etc.). Thanks to active learning, ProCAVIAR can

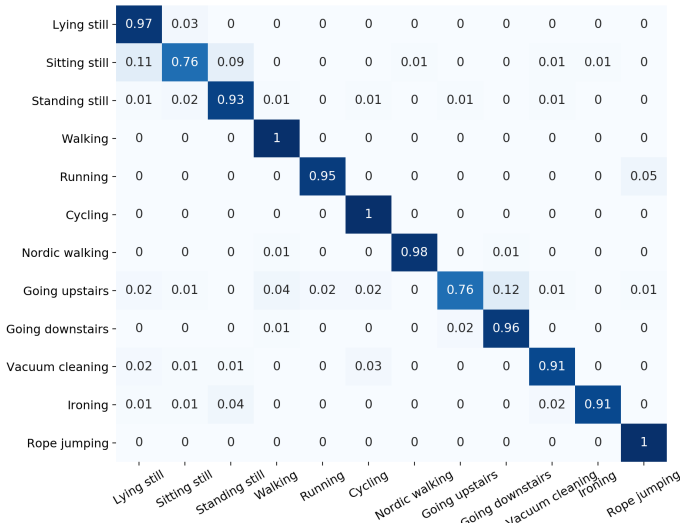


FIGURE 9: The confusion matrix obtained by our *Data-driven approach* on the PAMAP2 dataset.

continuously improve the semi-supervised classifier, which is initialized only with a limited set of labeled examples.

Differently from existing solutions based on a rigid formalism, ProCAVIAR takes advantage from probabilistic reasoning to capture the intrinsic uncertainty of context modeling. A preliminary evaluation shows that ProCAVIAR actually mitigates the problems of hybrid ontology based solutions while increasing the advantage of hybrid solutions over purely statistical approaches.

This work only represents a preliminary investigation of the effectiveness of using probabilistic logics in context-aware and hybrid activity recognition systems. We foresee several promising research directions.

First, a critical aspect is the setting of weights for the soft axioms determining the influence of context on activities. We plan to investigate how to populate the probabilistic ontology in a semi-automatic fashion, by extracting knowledge about context and activities from external sources. For instance, some works proposed to extract information from textual description [42] and images [43] of activities from the Web. Those works were mainly focused on building models for smart-home activity recognition.

Beside uncertainty on the association of context with activities, a probabilistic ontology may also capture the fact that context data may have an associated confidence value. Indeed, it is not always advisable to completely trust input context data (e.g., geographical positioning, as well as semantic place identification, can have different levels of approximation and reliability). Including uncertainty on input data has the potential of making our system more robust with respect to inaccurate information.

Finally, we aim to study personalization aspects. Each user may have her personal habits, and hence personal context situations. Incrementally adapting the probabilistic ontology to each user would allow our system to learn personalized

contexts and hence improving the accuracy and scalability of our approach.

REFERENCES

- [1] O. D. Lara, M. A. Labrador et al., “A survey on human activity recognition using wearable sensors.” *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.
- [2] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, “Activity recognition using cell phone accelerometers,” *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.
- [3] M. Stikic, K. Van Laerhoven, and B. Schiele, “Exploring semi-supervised and active learning for activity recognition,” in *2008 12th IEEE International Symposium on Wearable Computers*. IEEE, 2008, pp. 81–88.
- [4] L. Liao, D. Fox, and H. Kautz, “Location-based activity recognition,” in *Advances in Neural Information Processing Systems*, 2006, pp. 787–794.
- [5] D. Riboni and C. Bettini, “COSAR: Hybrid reasoning for context-aware activity recognition,” *Personal and Ubiquitous Computing*, vol. 15, no. 3, pp. 271–289, 2011.
- [6] —, “Owl 2 modeling and reasoning with complex human activities,” *Pervasive and Mobile Computing*, vol. 7, no. 3, pp. 379–395, 2011.
- [7] —, “Cosar: hybrid reasoning for context-aware activity recognition,” *Personal and Ubiquitous Computing*, vol. 15, no. 3, pp. 271–289, 2011.
- [8] C. Bettini, G. Civitarese, and R. Presotto, “Caviar: Context-driven active and incremental activity recognition,” *Knowledge-Based Systems*, p. 105816, 2020.
- [9] C. Bettini, O. Brdiczka, K. Henriksen, J. Indulska, D. Nicklas, A. Ranganathan, and D. Riboni, “A survey of context modelling and reasoning techniques,” *Pervasive and Mobile Computing*, vol. 6, no. 2, pp. 161–180, 2010.
- [10] A. Bulling, U. Blanke, and B. Schiele, “A tutorial on human activity recognition using body-worn inertial sensors,” *ACM Computing Surveys*, vol. 46, no. 3, pp. 33:1–33:33, 2014.
- [11] Z. S. Abdallah, M. M. Gaber, B. Srinivasan, and S. Krishnaswamy, “Activity recognition with evolving data streams: A review,” *ACM Computing Surveys (CSUR)*, vol. 51, no. 4, p. 71, 2018.
- [12] D. Guan, W. Yuan, Y.-K. Lee, A. Gavrilov, and S. Lee, “Activity recognition based on semi-supervised learning,” in *Embedded and Real-Time Computing Systems and Applications, 2007. RTCSA 2007. 13th IEEE International Conference on*. IEEE, 2007, pp. 469–475.
- [13] B. Longstaff, S. Reddy, and D. Estrin, “Improving activity classification for health applications on mobile devices using active and semi-supervised learning,” in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference on Pervasive Computing Technologies for Healthcare*. IEEE, 2010, pp. 1–7.
- [14] E. Hoque and J. Stankovic, “Aalo: Activity recognition in smart homes using active learning in the presence of overlapped activities,” in *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. IEEE, 2012, pp. 139–146.
- [15] T. Miu, P. Missier, and T. Plötz, “Bootstrapping personalised human activity recognition models using online active learning,” in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*. IEEE, 2015, pp. 1138–1147.
- [16] Z. S. Abdallah, M. M. Gaber, B. Srinivasan, and S. Krishnaswamy, “Adaptive mobile activity recognition system with evolving data streams,” *Neurocomputing*, vol. 150, pp. 304–317, 2015.
- [17] H. S. Hossain, M. A. A. H. Khan, and N. Roy, “Active learning enabled activity recognition,” *Pervasive and Mobile Computing*, vol. 38, pp. 312–330, 2017.
- [18] F. Baader, D. Calvanese, D. McGuinness, P. Patel-Schneider, D. Nardi et al., *The description logic handbook: Theory, implementation and applications*. Cambridge university press, 2003.
- [19] L. Chen and C. Nugent, “Ontology-based activity recognition in intelligent pervasive environments,” *International Journal of Web Information Systems*, 2009.
- [20] H. Chen, T. Finin, and A. Joshi, “The soupa ontology for pervasive computing,” in *Ontologies for agents: Theory and experiences*. Springer, 2005, pp. 233–258.
- [21] G. Meditskos, S. Dasiopoulou, and I. Kompatsiaris, “Metaq: A knowledge-driven framework for context-aware activity recognition combining sparql and owl 2 activity patterns,” *Pervasive and Mobile Computing*, vol. 25, pp. 104–124, 2016.

- [22] Q. Ni, I. Pau de la Cruz, and A. B. Garcia Hernando, "A foundational ontology-based model for human activity representation in smart homes," *Journal of Ambient Intelligence and Smart Environments*, vol. 8, no. 1, pp. 47–61, 2016.
- [23] O. Banos, C. Villalonga, J. Bang, T. Hur, D. Kang, S. Park, V. Le-Ba, M. B. Amin, M. A. Razzaq, W. A. Khan et al., "Human behavior analysis by means of multimodal context mining," *Sensors*, vol. 16, no. 8, p. 1264, 2016.
- [24] P. Patel, A. Gyrard, D. Thakker, A. P. Sheth, and M. Serrano, "Swotsuite: A toolkit for prototyping cross-domain semantic web of things applications." in *International Semantic Web Conference (Posters & Demos)*, 2016.
- [25] A. Ranganathan, J. Al-Muhtadi, and R. H. Campbell, "Reasoning about uncertain contexts in pervasive computing environments," *IEEE Pervasive computing*, vol. 3, no. 2, pp. 62–70, 2004.
- [26] G. Stoilos, G. Stamou, J. Z. Pan, V. Tzouvaras, and I. Horrocks, "Reasoning with very expressive fuzzy description logics," *Journal of Artificial Intelligence Research*, vol. 30, pp. 273–320, 2007.
- [27] M. Richardson and P. Domingos, "Markov logic networks," *Machine learning*, vol. 62, no. 1-2, pp. 107–136, 2006.
- [28] G. Civitarese, T. Sztyley, D. Riboni, C. Bettini, and H. Stuckenschmidt, "Polaris: Probabilistic and ontological activity recognition in smart-homes," *IEEE Transactions on Knowledge and Data Engineering*, 2019.
- [29] R. N. Carvalho, K. B. Laskey, and P. C. Costa, "Pr-owl—a language for defining probabilistic ontologies," *International Journal of Approximate Reasoning*, vol. 91, pp. 56–79, 2017.
- [30] L. L. dos Santos, R. N. Carvalho, M. Ladeira, L. Weigang, and G. L. Mendes, "Pr-owl 2 rl-a language for scalable uncertainty reasoning on the semantic web," in *11th International Workshop on Uncertainty Reasoning for the Semantic Web (URSW 2015)*, 2015, p. 14.
- [31] E. Bellodi, E. Lamma, F. Riguzzi, S. Albani et al., "A distribution semantics for probabilistic ontologies." *URSW*, vol. 778, pp. 75–86, 2011.
- [32] F. Riguzzi, E. Bellodi, E. Lamma, and R. Zese, "Probabilistic description logics under the distribution semantics," *Semantic Web*, vol. 6, no. 5, pp. 477–501, 2015.
- [33] M. Niepert, J. Noessner, and H. Stuckenschmidt, "Log-linear description logics," in *Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.
- [34] R. Helaoui, D. Riboni, and H. Stuckenschmidt, "A probabilistic ontological framework for the recognition of multilevel human activities," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 2013, pp. 345–354.
- [35] C. Bettini, O. Brdiczka, K. Henriksen, J. Indulska, D. Nicklas, A. Ranganathan, and D. Riboni, "A survey of context modelling and reasoning techniques," *Pervasive and Mobile Computing*, vol. 6, no. 2, pp. 161–180, 2010.
- [36] J. Noessner and M. Niepert, "Elog: a probabilistic reasoner for owl el," in *Proceedings of the 5th international conference on Web reasoning and rule systems*. Springer, 2011, pp. 281–286.
- [37] R. Helaoui, D. Riboni, and H. Stuckenschmidt, "A probabilistic ontological framework for the recognition of multilevel human activities," in *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. New York, NY, USA: ACM, 2013, pp. 345–354.
- [38] G. Civitarese, C. Bettini, T. Sztyley, D. Riboni, and H. Stuckenschmidt, "newnectar: Collaborative active learning for knowledge-based probabilistic activity recognition," *Pervasive and Mobile Computing*, vol. 56, pp. 88–105, 2019.
- [39] J. Gama, R. Sebastião, and P. P. Rodrigues, "On evaluating stream learning algorithms," *Machine learning*, vol. 90, no. 3, pp. 317–346, 2013.
- [40] A. Reiss and D. Stricker, "Introducing a new benchmarked dataset for activity monitoring," in *2012 16th International Symposium on Wearable Computers*. IEEE, 2012, pp. 108–109.
- [41] Y. Guan and T. Plötz, "Ensembles of deep lstm learners for activity recognition using wearables," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 2, pp. 1–28, 2017.
- [42] P. Palmes, H. K. Pung, T. Gu, W. Xue, and S. Chen, "Object relevance weight pattern mining for activity recognition and segmentation," *Pervasive and Mobile Computing*, vol. 6, no. 1, pp. 43–57, 2010.
- [43] D. Riboni and M. Murtas, "Sensor-based activity recognition: One picture is worth a thousand words," *Future Generation Computer Systems*, vol. 101, pp. 709–722, 2019.



CLAUDIO BETTINI is full professor of Computer Science at the University of Milan, where he leads the EveryWare laboratory. His research interests cover the areas of mobile and pervasive computing, data privacy, spatio-temporal data management, and knowledge reasoning. He acted as associate editor for IEEE TKDE, the VLDB Journal, and Pervasive and Mobile Computing journal. He is a IEEE senior member.



GABRIELE CIVITARESE is a researcher at the Computer Science department of the University of Milan. He is a member of the EveryWare laboratory since 2014. His research interests are mainly focused in the area of context-awareness and activity recognition. He published his results at major international conferences and journals.



DAVIDE GIANCANE is a post-graduated student in Computer Science at the University of Milan. He is member of EveryWare laboratory. His master thesis was focused on context modeling with probabilistic ontologies for activity recognition.



RICCARDO PRESOTTO is a Ph.D. student at the Department of Computer Science at the University of Milan, Italy. He is a member of the EveryWare laboratory. His research interests are focused on hybrid activity recognition methods which combine context-aware and semi-supervised learning.

...