

# Context-Aware Data Association for Multi-Inhabitant Sensor-Based Activity Recognition

Luca Arrotta  
*University of Milan*  
 luca.arrotta@studenti.unimi.it

Claudio Bettini, Gabriele Civitarese, Riccardo Presotto  
*University of Milan*  
 {claudio.bettini, gabriele.civitarese, riccardo.presotto}@unimi.it

**Abstract**—Recognizing the activities of daily living (ADLs) in multi-inhabitant settings is a challenging task. One of the major challenges is the so-called *data association* problem: how to assign to each user the environmental sensor events that he/she actually triggered? In this paper, we tackle this problem with a context-aware approach. Each user in the home wears a smartwatch, which is used to gather several high-level context information, like the location in the home (thanks to a micro-localization infrastructure) and the posture (e.g., sitting or standing). Context data is used to associate sensor events to the users which more likely triggered them. We show the impact of context reasoning in our framework on a dataset where up to 4 subjects perform ADLs at the same time (collaboratively or individually). We also report our experience and the lessons learned in deploying a running prototype of our method.

**Index Terms**—multi-inhabitant, activity recognition, smart-home

## I. INTRODUCTION

The sensor-based recognition of Activities of Daily Living (ADL) in smart-home environments has been extensively studied in the last decades. Indeed, continuously monitoring the behavior of fragile subjects (e.g., elderly) in their homes enables several health-care applications, like the detection of early symptoms of cognitive decline [1].

Existing works mainly focus on single inhabitant scenarios [2], [3]. Some research efforts have also considered the more challenging multi-inhabitant scenario, characterized by different subjects concurrently performing activities in the same environment (e.g., Alice watches TV while Bob is cooking) [4], [5]. This problem has been often formulated in terms of recognizing activities executed in parallel [6], but in most cases the task has not considered identifying the specific user that is performing each activity.

Indeed, the major challenge in sensor-based multi-inhabitant activity recognition is how to map each environmental sensor event (e.g., the fridge opening as revealed by a magnetic sensor, or the TV turned on as revealed by a smart plug) to the subject which triggered the event. In the literature, this problem is called *data association*, and solving it is a crucial step to identify which subject is performing which activity [7]. Previous work has partially addressed this problem by associating behavior patterns to individuals [8] or by using identifying devices (e.g., RFID readers, apps on smartphones) carried by the inhabitants and tags or beacons on the home objects [9]. Other works proposed the use of probabilistic graphical models (e.g., Hidden Markov Model) [4], [7], [10],

[11]. In [12] these models are used to exploit the correlation between data from the user’s smartphone and motion sensors. The evaluation of these approaches was limited to two subjects and to a very restricted set of activities.

In this paper, we propose a novel multi-inhabitant ADL recognition framework which exploits context-aware reasoning to tackle the data association problem. As sensing devices, we consider a smartwatch at the wrist of each subject and environmental sensors in the smart home infrastructure. From the smartwatch we derive the identity of the user, inertial sensors readings, the user’s location in the home (using a positioning infrastructure), and the user’s posture (e.g., if the user is sitting or standing). While constantly wearing devices or wristbands may be considered unrealistic, smartwatches nowadays are becoming common and they represent a non-intrusive technology that can be continuously worn in home environments. From the environmental sensors we detect open/close events of doors and drawers, turn on/off events of electrical devices, and more. The combination of environmental and mobile devices sensors has the potential of significantly improving activity recognition [13]–[15]. Indeed, by exploiting common knowledge we extract context from these two different sources of data and we use it to address the data association problem. For instance, based on Alice’s presence in a given area, we associate her identity only with the events fired by sensors physically close to her. The actual activity recognition module will analyze a separate stream of inertial and ambient sensors for each subject according to the data association. One of the advantages of our approach is its modularity: since activity recognition is decoupled from data association, it is possible to seamlessly integrate any single-inhabitant activity classifier to recognize the activities performed by each inhabitant separately. The contributions of this paper are threefold:

- A multi-inhabitant activity recognition framework which uses context-aware reasoning to tackle the data association problem
- Preliminary results about the impact of context-aware data association on a multi-inhabitant dataset, where up to 4 subjects at the same time perform both independent and collaborative activities
- Our experience and the lessons learned on implementing and deploying a running prototype of our system

## II. THE PROBLEM

The final goal of the activity recognition system (named just *system* in the following) is to periodically predict, for each user, the activity that she has been performing. We also would like the system to identify situations in which activities are jointly performed by multiple users. Intuitively, a group of users jointly performing an activity should include users that are in the same place, and that, according to the system prediction are performing the same activity<sup>1</sup>.

Formally, let  $\mathbf{U} = \{u_1, u_2, \dots, u_n\}$  be the set of users (the inhabitants of the smart-home) and  $\mathbf{A} = \{A_1, A_2, \dots, A_k\}$  the set of considered ADLs. Given an instant  $t$ , the system should return a set of tuples  $PA_t = \{\{(u_r, \dots, u_s), A_i\} | u_r, \dots, u_s \text{ are all the users predicted to jointly perform activity } A_i \text{ at time } t\}$ . Since we assume that each user is performing a single activity at a time, the same user cannot appear in more than one tuple. Moreover, each user that is present in a monitored room should appear in a tuple.

In order to solve this problem, the system continuously records a stream of time-stamped events coming from inertial and environmental sensors. Given an instant  $t$  and a user  $u$ , the system needs to solve a *data association problem* by computing a personal stream  $s(u)^t$  of sensor events associated with user  $u$  and collected in a time window  $[t, t - k]$  where  $k$  is the window size parameter. For example, suppose that Alice turns on the cooker at time  $t'$ . The corresponding sensor event (and its timestamp) generated by the plug sensor connected to the cooker and recorded by our system should be associated with Alice and hence considered part of  $s(\text{Alice})^t$  when  $0 \leq t - t' \leq k$ .

The data association problem is straightforward for events coming from inertial sensors on personal devices, but challenging for environmental sensors.

## III. OUR FRAMEWORK

In this section, we describe our multi-inhabitant activity recognition framework, whose architecture is depicted in Figure 1. Raw data from smartwatches and environmental sensors are pre-processed by the CONTEXT AGGREGATION module to obtain high-level context data. The resulting context, as well as raw sensor data, is then used by the CONTEXT-AWARE DATA ASSOCIATION module to generate, for each user, a personal stream of sensor data. For each stream, SENSOR-BASED ACTIVITY RECOGNITION infers for each user the activity that she is performing. Finally, high-level context is used to refine the statistical prediction by the PREDICTION REFINEMENT module. In the following, we describe in detail each component.

### A. Context aggregation

The CONTEXT AGGREGATION module receives as input raw data from multiple sensing devices and it derives high-

<sup>1</sup>Note that here we make the assumption that users that are performing the same activity in the same semantic location at the same time are actually jointly performing the activity. This is indeed the case in our considered setting.

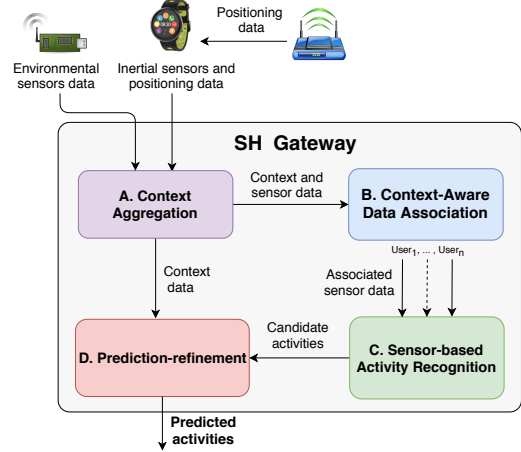


Fig. 1. Overall architecture of our framework

level context information, which is used to enable both data association and the refinement of statistical predictions. Given a time instant  $t$ , the system infers for each user  $u$  her personalized context  $C(u)^t = (l(u)^t, p(u)^t)$ , where  $l(u)^t$  is the location of  $u$  in the home at time  $t$  and  $p(u)^t$  is the posture of  $u$  at time  $t$ . For instance, if Bob is sitting in the kitchen at time  $t$  then  $C(\text{Bob})^t = (\text{kitchen}, \text{sitting})$ . Moreover, the system also infers the home context  $C_H^t$ , which keeps track of the status and the position of each sensor in the home. For instance, consider a home  $H$  instrumented with two plug sensors: one to detect the usage of the electrical stove in the kitchen and one to detect the usage of the television in the living room. Suppose that at time  $t$  Bob is watching TV and that no one is using the electrical stove. In this case  $C_H^t = \{(\text{stove}, \text{kitchen}, \text{OFF}), (\text{television}, \text{living\_room}, \text{ON})\}$ . In the following, we describe in more detail how this module infers these types of context data.

1) *User's position:* For each user  $u$ , the CONTEXT AGGREGATION module aims to derive her semantic position  $l(u)^t$  inside the home (e.g., kitchen, living room). Note that the organization of the home in semantic positions has to be performed in an offline phase, and its granularity depends on the accuracy of the underlying micro-localization system. A coarse granularity may consider room-level semantic positions (e.g., living room, kitchen, dining room), while a fine-grained granularity may map specific regions of each room into semantic positions (e.g., cooking area, dining table and sink area). Micro-localization can be implemented with a wide range of technologies (e.g., BLE Beacons, WiFi access points, passive localization). In our experimental setup, our positioning infrastructure is composed of BLE beacons and WiFi access points, whose signal strengths are collected by each user's smartwatch while we employ a Random Forest classifier to map the stream of positioning data into semantic positions at the room-level granularity.

2) *User's posture:* The posture  $p(u)^t$  (e.g., standing, sitting, lying) of a user  $u$  is computed by applying machine learning

methods to inertial sensors (i.e., accelerometer, gyroscope and magnetometer) data gathered by the user’s smartwatch. In our experimental setup, we use a fully connected deep neural network to distinguish between the *sitting* and *standing* postures, considering state-of-the-art time-domain and frequency-domain features from temporal windows of sensor data [16].

3) *Sensor status and position*: As previously mentioned, we consider as high-level context data of the home environment  $C_H^t$  the semantic position of each environmental sensor and its readings. The semantic position of each sensor is considered as prior knowledge defined during the deployment phase in the smart-home. Clearly, the granularity of sensors’ semantic positions should be the same as the one considered for the users. On the other hand, the status of environmental sensor events is dynamic. Each event is automatically mapped to a corresponding high-level concept. For example, when the magnetic sensor installed on the kitchen’s fridge door fires, it generates the high-level concept (*fridge\_door, kitchen, OPEN*), which means that the fridge door in the kitchen has been opened. This high-level concept will be part of the current home context  $C_H^t$ . In our experimental setup, we consider several types of environmental sensing devices like magnetic sensors, smart plugs, and pressure mats.

### B. Context-Aware Data Association

Given the high-level context information obtained from the CONTEXT AGGREGATION module and the raw sensor data gathered from inertial and environmental sensors, the CONTEXT-AWARE DATA ASSOCIATION module assigns to each user  $u$  a personal sensor data stream  $s(u)^t$ . A stream  $s(u)^t$  consists of: a) inertial sensor readings gathered from the personal device of  $u$  and b) environmental sensor events triggered by  $u$ . Note that an environmental sensor event ( $e, st, t$ ) (e.g. (*fridge\_door, OPEN, 12:32*)) can not directly identify the user who triggered it. This module approximates a stream  $s(u)^t$  by including all environmental events that are *consistent* with  $C(u)^t$  and  $C_H^t$ . For example, suppose that Alice and Bob are both in the living room, and a pressure mat on a chair generates an event at time  $t$ , thus indicating that someone is sitting. Suppose that Alice is standing, while Bob is the one which actually was sitting on that chair. Thanks to our posture detection module, our system infers that the pressure mat event is consistent with Bob’s context, while it is not-consistent with Alice’s context (i.e., a user can not activate a pressure mat if she is *standing*). Hence, in this case the pressure mat event will be included in  $s(Bob)^t$  and not in  $s(Alice)^t$ .

The output of CONTEXT-AWARE DATA ASSOCIATION is hence a personal stream  $s(u)^t \forall u \in \mathbf{U}$ . The solution is approximate since there may not be sufficient information to associate an event to a single user and in this case the event will be associated with the stream of each candidate user.

The notion of *consistency* used above is inherently related to the semantics of the context and of the action revealed by the event. In our experimental setup, the CONTEXT-AWARE DATA ASSOCIATION module is implemented with ontological

reasoning. We built an OWL2 ontology that defines the relationships between environmental sensor events and contexts. For instance, in the terminological part of our ontology there is an axiom that imposes that a user can trigger a sensor event only if she is in the same *semantic position* where the sensor is located. Other axioms combine *user’s posture* and *sensor status and position* to better associate environmental sensor events when more than one user is in the same semantic position at the same time. Referring to the example above, an axiom states that a pressure mat activation can be associated only to those users which recently switched to the *sitting* posture. Similarly, other rules state that the *sitting* posture is not compatible with sensor events that can be triggered only while *standing* (e.g., turning on the stove).

Whenever a sensor event ( $e, st, t'$ ) is triggered, our system computes its context-consistency for each user  $u$ . In particular, we add facts to the assertional part of the ontology to describe the sensor event and the context  $C(u)^{t'}$  and  $C_H$ . Then, by using the automatic consistency check of the resulting ontology, the system decides whether ( $e, st, t'$ ) should be included in  $s(u)^t$  (with  $t'$  in the time window defined by  $t$ ).

### C. Sensor-based Activity Recognition

The SENSOR-BASED ACTIVITY RECOGNITION module is in charge of inferring the activity performed by each user in the home, by analyzing each personal stream  $s(u)^t$  received from CONTEXT-AWARE DATA ASSOCIATION. The advantage of this approach is its modularity, since it is possible to use any single-inhabitant classifier  $h$  on each  $s(u)^t$  to periodically infer a probability distribution over the possible activities performed by the user  $u$  at time  $t$ :

$$h(s(u)^t) = \langle p_{A_1}, p_{A_2}, \dots, p_{A_n} \rangle$$

where  $p_{A_i}$  is the probability  $P(A_i | s(u)^t)$  that the user  $u$  at time  $t$  is performing activity  $A_i \in \mathbf{A}$ , based on the sensor data stream  $s(u)^t$ . We also have  $p_{A_i} \in [0, 1] \forall i$  and  $\sum_{i=1}^n p_{A_i} = 1$ .

In our experimental setup, we use a state-of-the-art deep-learning neural network that combines convolutional and recurrent layers [17]. By using this technique, features are automatically learned by the classifier from raw data. Environmental sensor events and inertial sensor data are provided as two separate channels to the network. On each channel, the network first applies 4 *convolutional* layers interleaved with *batch normalization* and *max pooling* layers. Then, we apply on each channel a *recurrent layer* (i.e., LSTM). In the following step we apply to each channel 5 *fully connected* layers. Finally, the two channels are concatenated into a single one and we apply 3 *fully connected* layers interleaved with a *dropout* layer. The probability distribution over the possible activities  $h(s(u)^t)$  is finally obtained by using a *softmax* layer.

### D. Prediction refinement

Finally, we refine each activity prediction  $h(s(u)^t)$  obtained from SENSOR-BASED ACTIVITY RECOGNITION using the high-level context  $C(u)^t$  and  $C_H^t$  computed by the CONTEXT AGGREGATION module. Indeed, activity recognition methods

are not completely accurate and sometimes they can derive a wrong activity that is not compatible with the user’s context. For instance, considering activity recognition methods based on machine learning, the training set used to create the model may not be representative and may not generalize on unseen activity patterns. To mitigate these issues, we adopt an approach similar to the one proposed in [18]. In particular, the PREDICTION-REFINEMENT module applies knowledge-based reasoning on context data to exclude from the probability distribution resulting from activity recognition those activities which are not context-consistent.

In our experimental setup this mechanism is based on the same ontology used by the CONTEXT-AWARE DATA ASSOCIATION module. Indeed, this ontology also contains axioms about the relationships between context data and activities. Our system evaluates whether an activity  $A$  is context-consistent by adding to the assertional part of the ontology the current context  $C(u)^t$  and  $C_H^t$  and the fact that  $u$  is currently performing activity  $A$ . The non context-consistent activities are removed from the probability distribution  $h(s(u)^t)$ , thus generating a refined probability distribution  $h'(s(u)^t)$  over the possible *context-consistent* activities.

For example, suppose that the system inferred that Bob is *watching television* with 60% of probability, *eating* with the 30% of probability and *working at pc* with the remaining 10%. According to our ontology, the *watching television* activity can be carried out only when: a) the user is sitting (*user posture*), b) the television is in the same semantic position of the user (*user and sensor position*), and c) the television is turned on (*sensor status*). Suppose that Bob is actually sitting at the dining table while eating and the television in the living room is turned on. Hence, *watching television* is not context-consistent for Bob considering how this activity is described in our knowledge-base. The resulting refined probability distribution of Bob in this case is 75% *eating* and 25% *working at pc*.

Finally, the output of the system is a set  $PA_t$  of tuples  $\langle (u_r, \dots, u_s), A \rangle$  such that for each user  $u$  in  $\{u_r, \dots, u_s\}$  the most likely activity according to  $h'(s(u)^t)$  is  $A$  and  $l(u_i)^t = l(u_j)^t \forall u_i, u_j \in \{u_r, \dots, u_s\}$ .

#### IV. EVALUATION

##### A. Dataset

To the best of our knowledge, there is no public multi-inhabitant annotated dataset suitable to experimentally evaluate our method. Hence, we collected a dataset in our research lab, simulating a smart-home environment. The lab area (about 80 square meters) was divided into 6 semantic areas, each representing a different room (hall, kitchen, dining room, closet room, living room and office). Different environmental sensors were installed to monitor the interaction of the users with their surrounding environment: 14 BLE beacons, 5 magnetic sensors, 9 pressure mats and 2 smart-plugs. We involved 12 different subjects in our acquisition, each one wearing a smartwatch on the dominant arm when performing activities. We simulated both single- and multi-inhabitants scenarios,

recording activities performed at the same time by one, two or four subjects. Before the acquisition, we informed the subjects about the sequence of activities they had to perform, but in order to increase the dataset variability, they were free to execute each activity in their own way. Overall, we acquired 12 scenarios in a single-inhabitant setting, 10 scenarios with 2 subjects, and 10 with 4 subjects involved. Table I shows, for each activity, the number of recording minutes and the average length expressed in seconds related to the data acquisition process.

TABLE I  
STATISTICS ON LABELED ACTIVITIES

	recording minutes	average length (s)
ANSWERING_PHONE	70.8	83.3
CLEARING_TABLE	44.3	61.8
COOKING	81.8	101.5
EATING	136.1	29.2
GETTING_IN	17.9	15.6
GETTING_OUT	15.0	18.7
MAKING_PHONE_CALL	63.9	68.4
PREPARING_COLD_MEAL	54.0	70.5
SETTING_UP_TABLE	52.5	58.4
TAKING_MEDICINES	36.1	31.0
USING_PC	102.2	107.6
WASHING_DISHES	54.3	66.5
WATCHING_TV	274.9	134.1

##### B. Results

We evaluated our method on the dataset described above, adopting a leave-one-subject-out cross-validation technique. Figure 2 shows the resulting confusion matrix. From the results

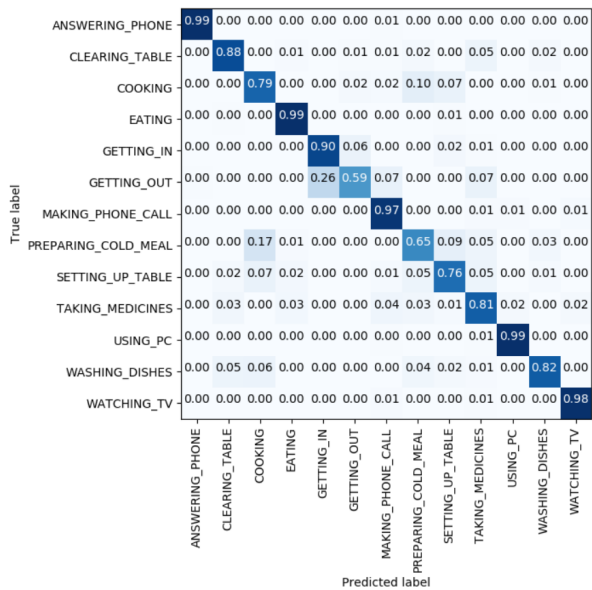


Fig. 2. Confusion matrix concerning the dataset activities

it emerges that some activities (e.g., Answering call, eating, watching TV) are recognized with a recall around 99%. This is due to the fact that, in our setup, these activities can be performed in unique semantic areas, triggering sensors that are

specific for that activity. For example, *watching television* is the only activity that can be performed in the *living room* and it involves the smart-plug sensor connected to the television. On the other hand, the accuracy of other activities decreases when the available context data can not provide additional information to perform data association (e.g. cooking, preparing a cold meal, setting up table). For instance, consider the scenario where Alice is preparing a salad and Bob is cooking pasta in the kitchen. Since both users are *standing* in the same semantic position, it is not possible to assign sensor events (e.g., the usage of the cooker) at a fine granularity. Hence, our system associates both to Alice and Bob the sensor events triggered in the kitchen, inferring their activities mainly relying on inertial sensor data. For the same reason, the prediction-refinement process does not have an impact in these situations.

Figure 3 shows the impact of our context-aware data association technique. Note that, to better highlight how data asso-

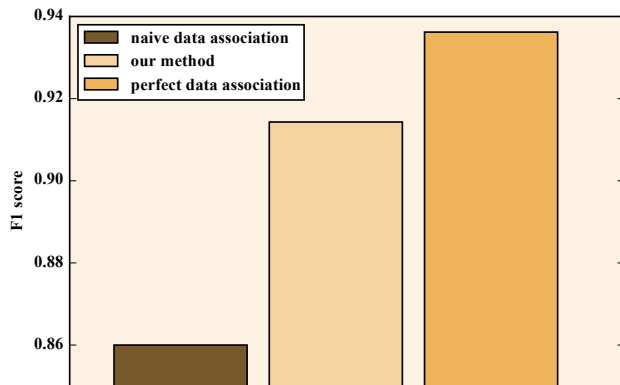


Fig. 3. Impact of data association.

ciation impacts activity recognition, those results are obtained without applying prediction refinement. We compared our method with a hypothetical perfect data association method where each sensor event is correctly assigned to the correct user that triggered it. Our solution is only 2% behind perfect data association in terms of F1 score averaged over all the activities. We also show that our method significantly outperforms a *naive data association* strategy, where each sensor event is simply assigned to all the users currently present in the home without considering the context.

The impact of our context-based prediction-refinement method is shown in Figure 4 where the increment in the recognition rate can be clearly observed in terms of the average F1 score.

## V. OUR EXPERIENCE ON THE DEPLOYMENT OF A PROTOTYPE

We implemented a running prototype of our system to detect multi-inhabitant activities in real-time. As expected, the implementation of this prototype turned out to be more challenging than simply applying our method to a dataset, since it required a significant amount of engineering (mainly

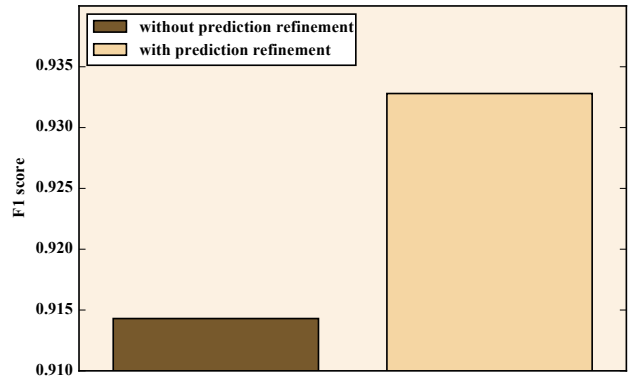


Fig. 4. Impact of prediction refinement.

due to temporal synchronization problems). Our prototype takes advantage of an activity recognition model built offline from the dataset described in Section IV-A. We deployed the prototype in a slightly different environmental setup with respect to the one of the dataset to evaluate the system’s robustness. The major difference regards the environment and the position of sensors: while the dataset was acquired in 6 semantic locations in the same large space, the prototype was deployed in 3 different rooms corresponding to 3 semantic locations: office, kitchen and living room equipped with the same type of sensors.

In order to quantitatively evaluate the recognition rate of our prototype, we asked some volunteers to perform scripted scenarios where up to 2 subjects at the same time performed both collaborative and individual activities. We only considered users that were not involved in the acquisition of the training set. The scripted scenarios involved a subset of the dataset activities: *cooking*, *preparing a cold meal*, *eating*, *setting up table*, *clearing table*, *washing dishes* and *watching TV*. Overall, the volunteers performed 45 activity instances. In order to compute the recognition rate of our prototype, we manually annotated those activities in real-time while the system automatically stored its predictions. Our prototype reached an overall F1 score of 88%. The drop in recognition rate with respect to the results on the dataset as presented in Section IV-B is due to the fact that the recorded predictions include periods of time in which the subjects transition from an activity to another. In these cases, the system exhibits a high uncertainty, since the dataset we used to build the recognition model does not capture transitions. Hence, we also evaluated our prototype considering “*stable*” predictions where the probability of the most likely activity is higher than 90%, since these are the ones we expect to be used in an application context. The average F1 score in this case was  $\sim 93\%$ , which is close to the one obtained using the dataset. We are currently working to introduce in our recognition model a ‘*transition*’ activity to better capture the uncertainty status between two subsequent activities.

While evaluating our prototype, we also observed the ca-

pability of our context-aware approach to overcome very common mistakes of machine learning only solutions. For instance, our dataset does not contain situations where posture detection has a significant impact on the overall recognition rate. However, during the evaluation of the prototype, the volunteers interleaved activities in a different way with respect to the dataset. Considering these new scenarios, we observed that the posture positively affects data association and the overall activity recognition rate. Despite this fact might seem a data collection design problem, this situation is very common in supervised machine learning, since it is always possible that the system is deployed in scenarios not considered in the training set.

## VI. CONCLUSION AND FUTURE WORK

In this paper, we presented a multi-inhabitant activity recognition framework which takes advantage of a novel data association mechanism based on context-aware reasoning. The preliminary results on a real rich dataset and on the deployment of a real-time prototype show the benefits of coupling context reasoning with machine learning for multi-inhabitant activity recognition.

Despite the clear advantage in exploiting context reasoning, the problem of building a reliable machine learning model from a sufficiently rich training set remains. As future work we plan to extend our framework with semi-supervised learning, so that the recognition model can be initialized with a limited training set avoiding the high cost of collecting annotated data. Moreover, this approach will continuously update the model with a combination of self-learning and active learning.

## ACKNOWLEDGMENT

This work has been partially funded by TIM S.p.A., Services Innovation Department, Joint Open Lab Milano. The authors would like to thank the former students Pierpaolo Albrici, Chiara Franceschetti and Daniele Guarnieri for their excellent work in software implementation.

## REFERENCES

- [1] G. Civitarese and C. Bettini, "Monitoring objects manipulations to detect abnormal behaviors," in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, 2017, pp. 388–393.
- [2] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, "Sensor-based activity recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 790–808, 2012.
- [3] 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.
- [4] G. Singla, D. J. Cook, and M. Schmitter-Edgecombe, "Recognizing independent and joint activities among multiple residents in smart environments," *Journal of ambient intelligence and humanized computing*, vol. 1, no. 1, pp. 57–63, 2010.
- [5] D. H. Wilson and C. Atkeson, "Simultaneous tracking and activity recognition (star) using many anonymous, binary sensors," in *International Conference on Pervasive Computing*. Springer, 2005, pp. 62–79.
- [6] T. Gu, Z. Wu, X. Tao, H. K. Pung, and J. Lu, "epSICAR: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition," in *Proceedings of the Seventh Annual IEEE International Conference on Pervasive Computing and Communications (PerCom)*. Washington, D.C.: IEEE Computer Society, 2009, pp. 1–9.
- [7] A. Benmansour, A. Bouchachia, and M. Feham, "Multioccupant activity recognition in pervasive smart home environments," *ACM Computing Surveys (CSUR)*, vol. 48, no. 3, p. 34, 2016.
- [8] A. S. Crandall and D. Cook, "Attributing events to individuals in multi-inhabitant environments," 2008.
- [9] L. Wang, T. Gu, X. Tao, H. Chen, and J. Lu, "Recognizing multi-user activities using wearable sensors in a smart home," *Pervasive and Mobile Computing*, vol. 7, no. 3, pp. 287–298, 2011.
- [10] A. Benmansour, A. Bouchachia, and M. Feham, "Modeling interaction in multi-resident activities," *Neurocomputing*, vol. 230, pp. 133–142, 2017.
- [11] H. Alemdar and C. Ersoy, "Multi-resident activity tracking and recognition in smart environments," *Journal of Ambient Intelligence and Humanized Computing*, vol. 8, no. 4, pp. 513–529, 2017.
- [12] N. Roy, A. Misra, and D. Cook, "Ambient and smartphone sensor assisted adl recognition in multi-inhabitant smart environments," *Journal of ambient intelligence and humanized computing*, vol. 7, no. 1, pp. 1–19, 2016.
- [13] A. Fleury, M. Vacher, and N. Noury, "Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results," *IEEE transactions on information technology in biomedicine*, vol. 14, no. 2, pp. 274–283, 2009.
- [14] L. Atallah, B. Lo, R. Ali, R. King, and G.-Z. Yang, "Real-time activity classification using ambient and wearable sensors," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 6, pp. 1031–1039, 2009.
- [15] H. F. Nweke, Y. W. Teh, G. Mujtaba, and M. A. Al-Garadi, "Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions," *Information Fusion*, vol. 46, pp. 147–170, 2019.
- [16] 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.
- [17] F. Ordóñez and D. Roggen, "Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition," *Sensors*, vol. 16, no. 1, p. 115, 2016.
- [18] D. Riboni and C. Bettini, "Context-aware activity recognition through a combination of ontological and statistical reasoning," in *Proceedings of the 6th International Conference on Ubiquitous Intelligence and Computing (UIC)*, ser. Lecture Notes in Computer Science, vol. 5585. Springer, 2009, pp. 39–53.