Addressing multi-users open challenge in habit mining for a process mining-based approach

Daniele Sora
DIAG "A.Ruberti"
Sapienza Università di Roma

Francesco Leotta
DIAG "A.Ruberti"
Sapienza Università di Roma

Massimo Mecella
DIAG "A.Ruberti"
Sapienza Università di Roma

Abstract

Models of human habits in smart spaces can be expressed by using a multitude of formalisms, whose readability influences the possibility of being validated by human experts. Given the growing availability of low-cost sensing devices promoted by the emerging Internet-of-Things, the analysis of huge amount of data produced by these systems will assume an utmost importance in the near future. But most of them are designed for single user cases. Moving forward in their development, often they hardly fit a realistic environment with many users. In this paper, we first review the most relevant approaches in the area during the last decade, and then we present an analysis pipeline that allows, starting from the sensor log of a smart space, to model human habits in a multi-user environment. The approach is based on exploit BLE beacons to discriminate the different users, then applying techniques borrowed from the area of business process automation and mining on a version of the sensor log preprocessed in order to translate raw sensor measurements into human actions. The paper also presents some hints of how the proposed method can be employed to automatically extract models to be reused for ambient intelligence in a multi-users environment.

1 Introduction

The aim of a smart space is providing people with automatic or semi-automatic services realizing the concept of ambient intelligence (AmI). The input for these intelligent services is represented by a sensor log, which is a sequence of measurement values acquired from sensors deployed across the monitored space. According to the number and type of installed sensors, and on the number of users acting in the environment, the amount of data produced may vary grandly in terms of size and rate. Additionally, the increasing availability of low-cost sensing technologies, even on wearable devices and smartphones, makes it likely to imagine a near future where a space (e.g., an house) and its inhabitants produce huge volumes and rates of data, according to the vision of the Internet-of-Things (IoT). Many approaches have been proposed in the literature to automatically analyze sensor logs at runtime to understand the current context and to make decisions on the basis of user

In: Marco Schaerf, Massimo Mecella, Drozdova Viktoria Igorevna, Kalmykov Igor Anatolievich (eds.): Proceedings of REMS 2018 – Russian Federation & Europe Multidisciplinary Symposium on Computer Science and ICT, Stavropol – Dombay, Russia, 15–20 October 2018, published at http://ceur-ws.org

preferences and habits. All of these solutions are based on models that relate the output of the sensors during a (potentially very short) temporal window, to a specific contextual information that can be then employed to act or reason on the state of the environment. Models can be either manually defined (specification-based methods) or obtained through machine learning techniques (learning- based methods). In the first case, models are usually based on logic formalisms, relatively easy to read and validate (once the formalism is known to the reader), but their creation requires an heavy cost in terms of expert time. In the latter case, the model is automatically learned from a training set (whose labeling cost may vary according to the proposed solution), but employed formalism are often less immediate to understand[1] . Another difference between the two approaches is that whereas specification-based methods use human actions as main modeling elements, learning-based ones directly refer to sensor measurements, thus loosing the focus on human actions and making even more difficult to visually inspect and validate produced models. On the other hand, taking as input raw sensor measurements usually makes learning-based methods easier to apply in a practical context; whereas, in the vast majority of cases, specification-based methods do not face the problem of translating sensor measurements into actions. As argued in [LMM15], applying methods originally taken from the area of *(business) process mining* [vdA16] to human habits may represent a compromise between specification-based and learning-based methods, provided that the gap between raw sensor measurements and human actions can be filled in by performing a log preprocessing step. Such a log preprocessing step may consist of simple inferences on data or complex machine learning algorithms.

Moreover other challenges are related to many users management. The most of the algorithms and approaches in the state of the art are devised for single-user scenario. In other words they make the hypothesis that the environment is populated only by an user at the time. But this is a very strong condition, commonly not true. Here we propose a first approach, based on BLE beacons, to organize the multi-users sensor log to be used with the state of the art activities also devised for single user conditions.

2 Related Works

In this paper, we aim at proposing an approach that merges human habits mining and multi-user managing methodology. We present in this section some relevant research papers, that adopted different approaches for the modeling technique and in the application context. In defining habits mining approaches there are two phases to approach that are the modeling phase and the runtime phase. The modeling phase is in charge of creating the models of human habits and environmental dynamics, whereas the runtime phase covers the aspect related to how these models are employed at runtime to recognize the context and to act onto the environment. These two phases can even overlap, in the cases where the system is able to refine models at runtime (in collaboration or not with the user) or even to completely create them from scratch at runtime. Models in the literature can be roughly divided into *specification-based* and *learning-based*. Specification-based models are usually more *human-readable* (even though a basic experience with formal logic languages is required), but creating them is very expensive in terms of human resources. Most learning-based models are instead represented using mathematical and statistical formalisms (e.g., HMMs), which make them difficult to be revised by experts and understood by final users. Learning-based techniques can be divided into *supervised* and *unsupervised* techniques. The former expect the input to be previously labeled according to the required output function, hence they require a big effort for organizing input data in terms of training examples, even though active learning can be employed to ease this task. Unsupervised techniques (or weakly supervised ones, i.e., those where only a part of the dataset is labeled) can be used to face this challenge but a limited number of works is available in the literature. Unsupervised techniques for AmI knowledge modeling can be useful since knowledge should not be considered as a static resource; instead it should be updated at runtime without a direct intervention of the users [RC09a], hence updating techniques should rely on labeling of sensor data as little as possible.

The learning based techniques rely on some well known algorithms. Widely used are approaches like Bayesian classification techniques, that are based on the well known Bayes theorem or $P(H|X) = \frac{P(X|H)P(H)}{P(X)}$, where H denotes the hypothesis (e.g., a certain activity is happening) and X represents the set of evidences (i.e., the current value of context objects). As calculating $P(X|H)$ can be very expensive, different assumptions can be made to simplify the computation. For example, *naïve Bayes* (NB) is a simple classification model, which supposes

[1] Notably, the lack of understandability of machine learning solutions is increasingly becoming a concern in both the AI and CS scientific communities [ACM17,RDT15], and has been recently taken up by DARPA, through the DARPA-BAA-16-53 “Explainable Artificial Intelligence (XAI)” program, cf. `https://www.darpa.mil/program/explainable-artificial-intelligence`. For example, the ACM Statement on Algorithmic Transparency and Accountability [ACM17] says: “There is also growing evidence that some algorithms and analytics can be opaque, making it impossible to determine when their outputs may be biased or erroneous”.

the n single evidences composing X independent (that the occurrence of one does not affect the probability of the others) given the situational hypothesis; this assumption can be formalized as $P(X|H) = \prod_{k=1}^{n} P(x_k|H)$. The inference process within the naïve Bayes assumption chooses the situation with the maximum a posteriori (MAP) probability. Hidden Markov Models (HMMs), that represents one of the most widely adopted formalism to model the transitions between different states of the environment or humans [Coo12, RC09b, SCSE10]. Here hidden states represent situations and/or activities to be recognized, whereas observable states represent sensor measurements. HMMs are a statistical model where a system being modeled is assumed to be a Markov chain, which is a sequence of events. A HMM is composed of a finite set of hidden states and observations that are generated from states. Some other *learning-based* techniques generally assume that all observations are independent, which could possibly miss long-term trends and complex relationships. Some of the most relevant algorithms in this field are used in Ambient Intelligence field. Follows an analysis of some examples. *Conditional Random Fields* - CRFs, for instance, eliminate the independence assumptions by modeling the conditional probability of a particular sequence of hypothesis given a sequence of observations. Modeling the conditional probability of the label sequence rather than the joint probability of both the labels and observations, as done by HMMs, allows CRFs to incorporate complex features of the observation sequence X without violating the independence assumptions of the model. The graphical model representations of a HMM (a directed graph) and a CRF (an undirected graph) makes this difference explicit. For example in [VKNEK08] a comparison between HMM and CRF is shown, where CRF outperforms HMM in terms of timeslice accuracy, while HMM outperforms CRF in terms of class accuracy.

Other statistical tools often employed is represented by Markov Chains (MCs), which are based on the assumption that the probability of an event is only conditional to the previous event, or Support Vector Machines (SVMs), that allow to classify both linear and non-linear data. SVMs are good at handling large feature spaces since they employ overfitting protection, which does not necessarily depend on the number of features. Binary Classifiers are built to distinguish activities. Artificial Neural Networks (ANNs) are a sub-symbolic technique, originally inspired by biological neuron networks. They can automatically learn complex mappings and extract a non-linear combination of features. Some other techniques stem from data mining methods for *market basket analysis* (e.g., the Apriori algorithm [AS+94]), which apply a windowing mechanism in order to transform the event/sensor log into what is called a *database of transactions*.

Initial approaches to the development of context-aware systems able to recognize situations were based on *predicate logic*. Loke [Lok04] introduced a PROLOG extension called LogicCAP; here the "in-situation" operator captures a common form of reasoning in context-aware applications, which is to ask if an entity E is in a given situation S. *Ontologies* (denoted as ONTO) represent the last evolution of logic-based approaches and have increasingly gained attention as a generic, formal and explicit way to "capture and specify the domain knowledge with its intrinsic semantics through consensual terminology and formal axioms and constraints" [YCDN07]. They provide a formal way to represent sensor data, context, and situations into well-structured terminologies, which make them understandable, shareable, and reusable by both humans and machines. A considerable amount of knowledge engineering effort is expected in constructing the knowledge base, while the inference is well supported by mature algorithms and rule engines. Example of using ontologies in identifying situations is given by [RB09] (later evolved in [HRS13, RSCS16]). Instead of using ontologies to infer activities, they use ontologies to validate the result inferred from statistical techniques. The way an AmI system makes decisions on the actions can be compared to decision making in AI agents. As an example, *reflex agents with state*, as introduced in [RN95], take as input the current state of the world and a set of Condition-Action rules to choose the action to be performed. Similarly, Augusto [AN04] introduces the concept of Active DataBase (ADB) composed by Event-Condition-Action (ECA) rules. An ECA rule basically has the form "ON *event* IF *condition* THEN *action*", where conditions can take into account time.

First attempts to apply techniques taken from the business process management - BPM [DLRM+13] area were the employment of workflow specifications to anticipate user actions. A workflow is composed by a set of tasks related by qualitative and/or quantitative time relationships. Authors in [GGH06] present a survey of techniques for *temporal calculus* (i.e., Allen's Temporal Logic and Point Algebra) and *spatial calculus* aiming at decision making. The SPUBS system [AIB+09, AIB+10] automatically retrieve these workflows from sensor data.The Table 1 recaps details about some techniques modeling approaches: in particular about their type (Specification, Supervised/Unsupervised learning) and multiple users support.

It appears clear how only few works support multiple users with ad hoc systems, so multi-users management is still an open challenging problem.

Table 1: Model construction (Y - Yes, N - No, S - Single, M - Multiple)

Technique	**Modeling Type**	**Single/Multiple users**	**Additional labeling**
AUG-ECA [ALM+08]	Specification	S	N
CHEN-ONT [CNW12]	Specification	S	N
RIB-PROB [HRS13, RSCS16]	Specification	S	N
NUG-EVFUS [HNM+09]	Specification	S	N
CASAS-HMM [SCSE10]	Supervised	S	N
CASAS-SVM [KC14]	Supervised	S	N
CASAS-HMMNBCRF [Coo12]	Supervised	S	N
WANG-EP [GWW+11]	Supervised	S	N
KROS-CRF [VKNEK08]	Supervised	S	N
REIG-SITUATION [BCR09]	Supervised	M	N
YANG-NN [YWC08]	Supervised	S	N
LES-PHI [LCB06]	Supervised	S	N
BUE-WISPS [BPPW09]	Supervised	S	N
FLEURY-MCSVM [FVN10]	Supervised	M	Y
CASAS-DISCOREC [CKR13, RCHSE11, RC13]	Weakly Sup.	S	N
STIK-MISVM [SLES11]	Weakly Sup.	M	Y
AUG-APUBS [AAB+12]	Unsupervised	S	N
CASAS-HAM [RC09b]	Unsupervised	S	N
WANG-HIER [WGTL12]	Unsupervised	S	N
PALMES-OBJREL [PPG+10]	Unsupervised	S	N

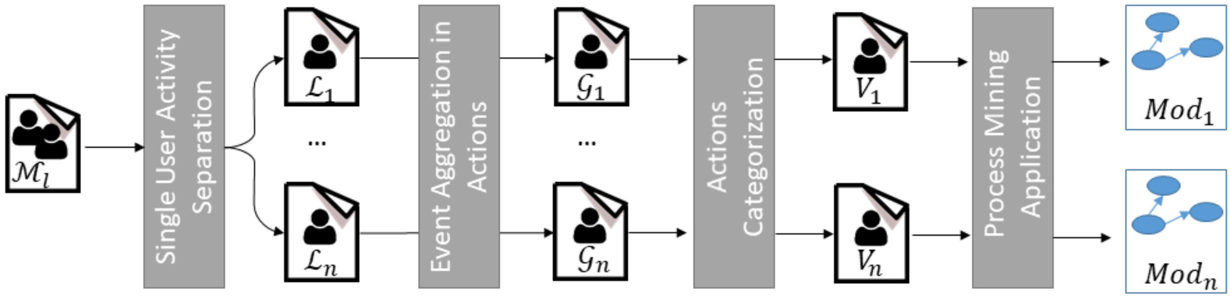

Figure 1: System overview of the proposed application of Process Mining techniques on multi-users smart space log

3 System Description

As introduced in the previous sections, this paper presents an approach that allows a multiuser management adaptation for already existing systems, originally designed on single user hypothesis. An extended description of the approach can be read in [SA18].

Figure 1 shows an high level description of the pipeline proposed. In the first phase a log separation is performed. In other words from a sensor log created by the activity of many users into the space, the system automatically extracts single user sub-logs. The context which we are going to consider is a classical generic smart home: each room of the home is enriched with a set of heterogeneous sensors(i.e. motion Passive Infra Red based, doors contact, energy monitoring and so on). This sensors set can be formalized as: $S = \{S_1 \dots S_m\}$ with $|S| = m$ total number of installed sensors. Each sensor is associated to a state. We indicate the state of a sensor $S_i, 1 \leq i \leq m$ with the notation $S_i[t]$ where t is a temporal indication. Moreover the set of the possible states is $S_i[t] \in \{0, 1\}$. This means that the state activation is considered as a boolean value (not activated/activated).

3.1 Users separation: from multiple users log to single user traces

The users, acting inside the environment, cause sensors activations records. They are collected by the smart home system, paired with activation time stamp ts and stored into a list, $A = [\langle ts_w, A_w \rangle], w \in \mathbb{N}$. This is the smart space described above. When assigned to a user H_i, they are going to compose data traces $T_{H_i} = \{A_1^{H_i}, A_2^{H_i}, \dots, A_{L_{T_{H_i}}}^{H_i}\}$ with $|T_{H_i}| = L_{T_{H_i}} \in \mathbb{N}, 1 \leq i \leq P$. Notice that $L_{T_{H_i}}$ is different for each user, since it depends on the user's activities. Hence, since multi-users assumptions, more than one trace will be produced. The set of data traces will be $T = \{T_{H_1}, T_{H_2}, \dots, T_{H_P}\}$ with $|T| = P, P \in \mathbb{N}$. This last formula implies $|T| = |H|$, so, we are imposing a trace for each user. This is not automatically obtained just recording activation data, but it is the final goal of this work. By now the dataset produced by this kind of environment would be a confused interleaving of different $T_{H_i}, 1 \leq i \leq P$, without the possibility of reconstructing the different traces.

To perform the log separation we use a special sensor kind, the Bluetooth-Low-Energy sensor. The beacon set is defined B as $B = \{B_1, \dots B_n\}$ with B finite set and $|B| = n$ number of beacons in the system. When a user H_i enters into a proximity beacon domain B_{id}, the system creates an association between H_i and B_{id}. An association is expressed as a coupled value $\langle ts, B_{id} \rangle$ where ts the time stamp indicating when the association happens and B_{id} is the beacon involved. For each user, the multi user management system generates an association list $\Lambda^{H_i} = [\langle ts, B_{id} \rangle, \dots], 1 \leq id \leq n$ and ts is a progressive number. The set of all the association lists is $\Lambda = \Lambda^{H_i}$ *for each* $i : 1 \leq i \leq P$.

Figure 2 shows how the beacons subsystem works. For hypothesis, each user has a registered smartphone that brings always with him. Moving near the beacon, the smartphone registers the change of the nearest beacon encountered. In this way the subsystem has a complete beacon log. The system has a Specification Based definition of the sensors related to that beacon. So combining the beacon log with the multiple users data log, we can separate the portion of the log related to every user. However there is the possibility that two different users get the registration to the same beacon: in this case the log records are duplicated and inserted into both traces of both the users. The generated traces are not completely clean, but they are good enough to apply the techniques.

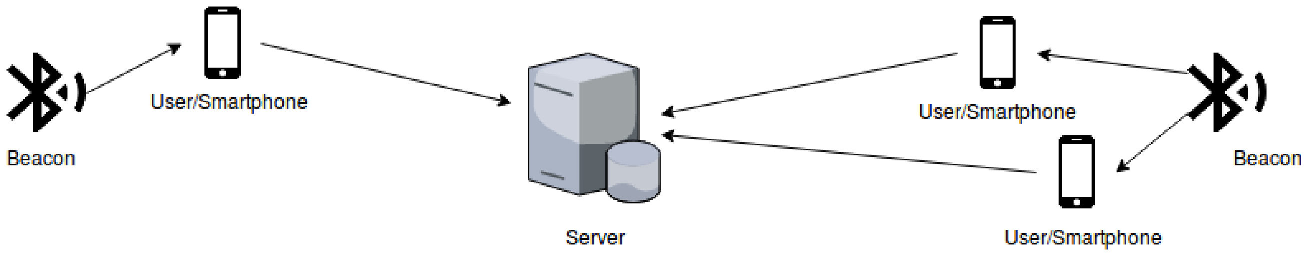

Figure 2: Beacons Subsystem Architecture

4 Conclusions and future works

In this paper we focus the attention on the well known problem of multi-users in a smart space system management. We provide a feasible solution to this problem: the main idea is to be able to distinguish between different users, for building the data traces of the sensor activations correlated to their activities. Naturally, the optimal result would be to obtain a log related to a given user containing only the correct measurements. The approach exploits simple technology like BLE beacons. Moreover we showed the generality of this approach involving it in a development pipeline that relies on Process Mining techniques (Fuzzy Mining) to model the Smart Space inhabitant habits. Effectiveness and precision depend on several factors, that can have an impact more or less important. For instance technology factors related to the proximity BLE. This kind of technology is rapidly and continuously evolving. There exists some protocols and formats, for instance Eddystone that can provide more data and metadata useful for improving precision. Moreover more complex computation on BLE signals can help in improving the system precision. In this first approach the nearest beacon information is mined just looking at the maximum RSSI value given by all the reached beacons. So each beacon is analyzed independently from the others. Other more complex algorithms that involves triangulation techniques between many beacons can bring more precision in localizing the users and, consequently, in separating the user traces.

Models of human habits have an utility that goes beyond the mere analysis. A way to recognize their occurrences at runtime must be devised. At this point the models can be used either for conformance checking or to make the smart space work as a proactive agent. In the second case, controllable services provided by the smart house (e.g., automatic shutters) should be part of the model also in a real multi-user environment.

References

[AAB+12] Asier Aztiria, Juan Carlos Augusto, Rosa Basagoiti, Alberto Izaguirre, and Diane J Cook. Discovering frequent user–environment interactions in intelligent environments. *Personal and Ubiquitous Computing*, 16(1):91–103, 2012.

[ACM17] ACM U.S. Public Policy Council and ACM Europe Policy Committee. Statement on algorithmic transparency and accountability. Online at: https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf, 2017.

[AIB+09] Asier Aztiria, Alberto Izaguirre, Rosa Basagoiti, Juan Carlos Augusto, and Diane J Cook. Discovering frequent sets of actions in intelligent environments. In *Intelligent Environments*, pages 153–160. Citeseer, 2009.

[AIB+10] Asier Aztiria, Alberto Izaguirre, Rosa Basagoiti, Juan Carlos Augusto, and Diane Cook. Automatic modeling of frequent user behaviours in intelligent environments. In *Intelligent Environments (IE), 2010 Sixth International Conference on*, pages 7–12. IEEE, 2010.

[ALM+08] Juan Carlos Augusto, Jun Liu, Paul McCullagh, Hui Wang, and Jian-Bo Yang. Management of uncertainty and spatio-temporal aspects for monitoring and diagnosis in a smart home. *International Journal of Computational Intelligence Systems*, 1(4):361–378, 2008.

[AN04] Juan Carlos Augusto and Chris D Nugent. The use of temporal reasoning and management of complex events in smart homes. In *ECAI*, volume 16, page 778. Citeseer, 2004.

[AS+94] Rakesh Agrawal, Ramakrishnan Srikant, et al. Fast algorithms for mining association rules. In *Proc. 20th Int. Conf. Very Large Data Bases, VLDB*, volume 1215, pages 487–499, 1994.

[BCR09] Oliver Brdiczka, James L Crowley, and Patrick Reignier. Learning situation models in a smart home. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1):56–63, 2009.

[BPPW09] Michael Buettner, Richa Prasad, Matthai Philipose, and David Wetherall. Recognizing daily activities with rfid-based sensors. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 51–60. ACM, 2009.

[CKR13] Diane J Cook, Narayanan C Krishnan, and Parisa Rashidi. Activity discovery and activity recognition: A new partnership. *IEEE transactions on cybernetics*, 43(3):820–828, 2013.

[CNW12] Liming Chen, Chris D Nugent, and Hui Wang. A knowledge-driven approach to activity recognition in smart homes. *IEEE Transactions on Knowledge and Data Engineering*, 24(6):961–974, 2012.

[Coo12] Diane J Cook. Learning setting-generalized activity models for smart spaces. *IEEE intelligent systems*, 27(1):32–38, 2012.

[DLRM+13] Marlon Dumas, Marcello La Rosa, Jan Mendling, Hajo A Reijers, et al. *Fundamentals of business process management*, volume 1. Springer, 2013.

[FVN10] Anthony Fleury, Michel Vacher, and Norbert 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*, 14(2):274–283, 2010.

[GGH06] Bjrn Gottfried, Hans Guesgen, and Sebastian Hbner. Spatiotemporal reasoning for smart homes. In Juan Augusto and Chris Nugent, editors, *Designing Smart Homes*, volume 4008 of *Lecture Notes in Computer Science*, pages 16–34. Springer Berlin / Heidelberg, 2006.

[GWW+11] Tao Gu, Liang Wang, Zhanqing Wu, Xianping Tao, and Jian Lu. A pattern mining approach to sensor-based human activity recognition. *IEEE Transactions on Knowledge and Data Engineering*, 23(9):1359–1372, 2011.

[HNM+09] Xin Hong, Chris Nugent, Maurice Mulvenna, Sally McClean, Bryan Scotney, and Steven Devlin. Evidential fusion of sensor data for activity recognition in smart homes. *Pervasive and Mobile Computing*, 5(3):236–252, 2009.

[HRS13] Rim Helaoui, Daniele Riboni, and Heiner 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*, pages 345–354. ACM, 2013.

[KC14] Narayanan C Krishnan and Diane J Cook. Activity recognition on streaming sensor data. *Pervasive and mobile computing*, 10:138–154, 2014.

[LCB06] Jonathan Lester, Tanzeem Choudhury, and Gaetano Borriello. A practical approach to recognizing physical activities. In *International Conference on Pervasive Computing*, pages 1–16. Springer, 2006.

[LMM15] Francesco Leotta, Massimo Mecella, and Jan Mendling. Applying process mining to smart spaces: Perspectives and research challenges. In *Advanced Information Systems Eng.ing Workshops*, pages 298–304. Springer, 2015.

[Lok04] Seng W Loke. Logic programming for context-aware pervasive computing: Language support, characterizing situations, and integration with the web. In *Proceedings of the 2004 IEEE/WIC/ACM International Conference on Web Intelligence*, pages 44–50, 2004.

[PPG+10] Paulito Palmes, Hung Keng Pung, Tao Gu, Wenwei Xue, and Shaxun Chen. Object relevance weight pattern mining for activity recognition and segmentation. *Pervasive and Mobile Computing*, 6(1):43–57, 2010.

[RB09] D. Riboni and C. Bettini. Context-aware activity recognition through a combination of ontological and statistical reasoning. *Ubiquitous Intelligence and Computing*, pages 39–53, 2009.

[RC09a] P. Rashidi and D.J. Cook. Keeping the resident in the loop: Adapting the smart home to the user. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 39(5):949–959, Sept 2009.

[RC09b] Parisa Rashidi and Diane J Cook. Keeping the resident in the loop: Adapting the smart home to the user. *IEEE Transactions on systems, man, and cybernetics-part A: systems and humans*, 39(5):949–959, 2009.

[RC13] Parisa Rashidi and Diane J Cook. Com: A method for mining and monitoring human activity patterns in home-based health monitoring systems. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 4(4):64, 2013.

[RCHSE11] Parisa Rashidi, Diane J Cook, Lawrence B Holder, and Maureen Schmitter-Edgecombe. Discovering activities to recognize and track in a smart environment. *Knowledge and Data Engineering, IEEE Transactions on*, 23(4):527–539, 2011.

[RDT15] Stuart J. Russell, Daniel Dewey, and Max Tegmark. Research priorities for robust and beneficial artificial intelligence. *AI Magazine*, 36(4), 2015.

[RN95] Stuart J Russell and Peter Norvig. Artificial intelligence: A modern approach, 1995.

[RSCS16] Daniele Riboni, Timo Sztyler, Gabriele Civitarese, and Heiner Stuckenschmidt. Unsupervised recognition of interleaved activities of daily living through ontological and probabilistic reasoning. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1–12. ACM, 2016.

[SA18] Daniele Sora and Juan Carlos Augusto. Managing multi-user smart environments through ble based system. *IWSSS2018: International Workshop on Smart Sensing System*, 2018.

[SCSE10] Geetika Singla, Diane J Cook, and Maureen Schmitter-Edgecombe. Recognizing independent and joint activities among multiple residents in smart environments. *Journal of ambient intelligence and humanized computing*, 1(1):57–63, 2010.

[SLES11] Maja Stikic, Diane Larlus, Sandra Ebert, and Bernt Schiele. Weakly supervised recognition of daily life activities with wearable sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2521–2537, 2011.

[vdA16] Wil MP van der Aalst. *Process mining data science in action*. Springer, 2016.

[VKNEK08] Tim Van Kasteren, Athanasios Noulas, Gwenn Englebienne, and Ben Kröse. Accurate activity recognition in a home setting. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 1–9. ACM, 2008.

[WGTL12] Liang Wang, Tao Gu, Xianping Tao, and Jian Lu. A hierarchical approach to real-time activity recognition in body sensor networks. *Pervasive and Mobile Computing*, 8(1):115–130, 2012.

[YCDN07] J. Ye, L. Coyle, S. Dobson, and P. Nixon. Ontology-based models in pervasive computing systems. *The Knowledge Engineering Review*, 22(4):315–347, 2007.

[YWC08] Jhun-Ying Yang, Jeen-Shing Wang, and Yen-Ping Chen. Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers. *Pattern recognition letters*, 29(16):2213–2220, 2008.