Department of
Information Engineering
and Computer Science DISI

UNIVERSITY
OF TRENTO - Italy

DISI - Via Sommarive 14 - 38123 Povo - Trento (Italy)
http://www.disi.unitn.it

Personal Context Recognition via Reliable Human-Machine Collaboration

Fausto Giunchiglia, Mattia Zeni, Enrico Bignotti

March 2018

Technical Report # DISI-18-016

Proceedings in: Pervasive Computing and Communications (PerCom), 2018 IEEE International Conference on IEEE.

Personal Context Recognition via Reliable Human-Machine Collaboration

Fausto Giunchiglia
DISI - University of Trento
Via Sommarive, 9
I-38123 Povo (TN)
Email: fausto.giunchiglia@unitn.it

Mattia Zeni
DISI - University of Trento
Via Sommarive, 9
I-38123 Povo (TN)
Email: mattia.zeni.1@unitn.it

Enrico Bignotti
DISI - University of Trento
Via Sommarive, 9
I-38123 Povo (TN)
Email: enrico.bignotti@unitn.it

***Abstract*—An effective context recognition system cannot rely only on sensor data but requires the user to collaborate with the system in providing her own knowledge. In approaches such as participatory sensing, which leverages on users to annotate and collect their own data, user-generated data is usually assumed to be accurate; however, in real life situations, this is not the case. Research in social sciences and psychology shows that humans are unreliable due to several behavioral biases when describing their everyday life. In this paper, we propose to parametrize two biases, i.e., cognitive bias and carelessness, in order to identify and evaluate their impact on the users' reliability when recognizing users' context. The parameters are part of an architecture for context modelling and recognition from previous work, which combines sensors and users as a source of information. We evaluate our approach on a dataset of location points from the SmartUnitn One experiment.**

***Index Terms*—context recognition; behavioral analysis; social sensing; smartphone; context-aware computing**

I. Introduction

Each individual interprets her surroundings differently because of her habits, routines, and intelligence; this represents her personal context. However, the user's notion and representation of her own context is radically different from what machines understand. In fact, machines interpret the world via sensors but lack the knowledge about it. For instance, while any location can be reduced to (a set of) coordinates for a machine, for humans locations can be differentiated via categories such as function, e.g., home vs. work. Thus, effective context recognition requires machines to collaborate with users to learn their habits and knowledge of the world.

Imparting human knowledge that machines have to learn for context recognition is an integral part of supervised machine learning [21]. In supervised approaches, humans are employed as experts that interact either in an offline fashion by annotating sensor data [22], or directly online by interactively asking the user to reduce the amount of labeling effort, see active learning [18], [13], [14]. Humans as annotators are extremely important also in real life scenarios such as participatory sensing, where mobile users actively participate in collecting data [17]. However, using experts for labeling another person's data is unfeasible, due to the workload and issues in aligning to users' experience [3]; thus, the annotation task is up to users themselves.

Being their own annotators, users should be motivated to provide the correct information to understand their own data. Unfortunately, this is not always the case, since humans are not a reliable source of information [20]. This issue is well known in social sciences and psychology, because of response biases in answering self-reports [23]. A major issue in sociology is that the evaluation of the impact of these biases on users' answers is still not well understood [5].

We propose an architecture, based on the work from [8], [24], where users annotate their own sensor data (in real time) on their smartphone, with the final goal of recognizing the context based only on the reliable annotations. To ensure the reliability of annotations, we account for two specific behavioral biases, i.e., cognitive bias and carelessness, by parametrizing them. These parameters are used to evaluate the impact of biases when recognizing users' context in real life scenarios where they are not assumed to be reliable. To the best of our knowledge, we are the first to explicitly consider and address user reliability in context recognition in the area of pervasive computing.

We evaluate our approach on a dataset of location points from the SmartUnitn One experiment [11], which aims at understanding how students' behavior affects their academic performance. Results show that we can detect and quantify biases in students annotations. We also compare the algorithm based on students' annotations with an unsupervised algorithm, proving that our solution performs better.

The paper is structured as follows. Section II illustrates how we enable reliable human-machine collaboration for context recognition, while we provide an evaluation in Section III based on the SmartUnitn One. Section IV describes works related to ours and Section V concludes our paper.

II. Addressing User Reliability

In our approach, we model the user context based on its definition from [7], i.e., "a theory of the world which encodes an individual's subjective perspective about it", as shown in Section II-A. Based on this model, we adapt our context model via sociological approaches to elicit users' input, illustrated in Section II-B. Then, we enable the collaboration between humans and machines to recognize context by asking for users feedback on her surroundings while also collecting real-life data on their smartphones, shown in Section II-C. To ensure their reliability, we account for biases in answering by quantifying them as parameters of users' responding behavior, detailed in Section II-D.

A. *Personal Context Modelling*

For humans, context is the mechanism to make sense of our environment to decide what is relevant to our current state of affairs [2]. For instance, a student focus will be on specific elements of her environment, e.g., a lesson taking place, which means that others may not be considered relevant, e.g., the teacher gesturing.

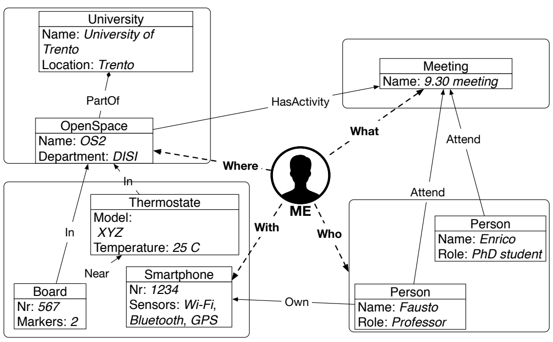


Fig. 1: The four dimensions of context, centered on the subject.

Figure 1 shows an example of context as a knowledge graph. Each node represents an entity with its respective attributes and their attribute values. For instance, attributes of Enrico in Figure 1 are "Class", "Name", and "Role", and their corresponding values are "Person", "Enrico", and "PhD student", respectively. Edges represent relations between entities, e.g., "Enrico" has one relation "Attend" for "Meeting".

We model context as a tuple:

$$Cxt =< me, WA, WE, WO, WI > \tag{1}$$

where:

- **me** is the person on which the context is centered, e.g., a student;
- WA is the Temporal component, i.e., the dimension that answers the question "**W**h**A**t are you doing?". It covers all the relevant activities for a person in the current context, e.g., attending a lesson;
- WE is the Spatial component, i.e., the dimension that answers the question "**W**h**E**re are you?". It covers all the relevant locations for a person in the current context, e.g., a classroom;
- WO is the Social component, i.e., the dimension that answers the question "**W**h**O** are you with?". It covers all the relevant people for a person in the current context, e.g., the teacher and classmates;
- WI is the Object component, i.e., the dimension that answers the question "**W**hat are you w**I**th?". It covers all the relevant objects for a person in the current context, e.g., his or her own smartphone.

Each dimension is modelled as an ontology based on the general ontology in [9] unifying human perception and knowledge representation.

B. *Involving Humans*

Ontologies can be seen as a hierarchy of labels to annotate users' life; to do so, they must be made deployable on smartphones and usable by people. Thus, we adapt them as time diaries, which are widely used in sociology to analyze human behavior and consist of logs where respondents report activities performed, locations visited and people encountered during their day [12].

This presentation of ontological information as time diaries consists in two main steps: *i)* the appropriate time use classification standard must be agreed upon, and *ii)* the context dimensions to be covered must be chosen. Then, the resulting ontologies must be adapted, in collaboration with sociologists, based on the research needs to become the coded entries of smartphone-based time diaries.

The time diary used in this work was presented in [8], so we will skip detailing the building process. To summarize, it relied on the America Time Use Survey (ATUS) [19] to obtain an ontology consisting in over 80 candidate labels for three dimensions, each being a question to be asked: activities, locations, and people. Then, sociologists halved the number of labels to 43.

TABLE I: The time diary used in [8].

What are you doing?	Where are you?	Who is with you?
Lesson	Class	Alone
Study	Study Hall	Classmate(s)
Eating	Library	Friend(s)
Personal Care	Other University place	Roomate(s)
En route (*)	Canteen	Partner(s)
Social life	Bar/ Pub/etc	Colleague(s)
Social media & internet	Relative(s)	Other
Cultural Activity	Home	
Sport	Other Home	**(*) How are you travelling?**
Shopping	Workplace	By Foot
Hobbies	Outdoors	By Bus
Other Free Time	Gym	By Train
Work	Shop	By Car
Housework	Other Place	By Bike
Volunteering		Other
Other		

In the resulting time diary, shown in Table I, each dimension is mirrored by a list of possible closed answers, where each question refers to its corresponding context dimension. Since WI, i.e., objects, were not deemed relevant, there are three questions: *i)* "What are you doing?", which accounts for activities (WA), *ii)* "Where are you?", which accounts for locations (WE), and *iii)* "Who is with you?", which accounts for social relations (WO). The asterisk links the question "How are you traveling?" and the "En route" activity. When a user selects it, instead of the "Where are you?" options, a list of means of transportation is provided.

C. *Human-Machine Collaboration for Context Recognition*

In addition to human input, we also need to involve machines to collaborate as a source of information and combine the two inputs to recognize the context. To do so, we rely on

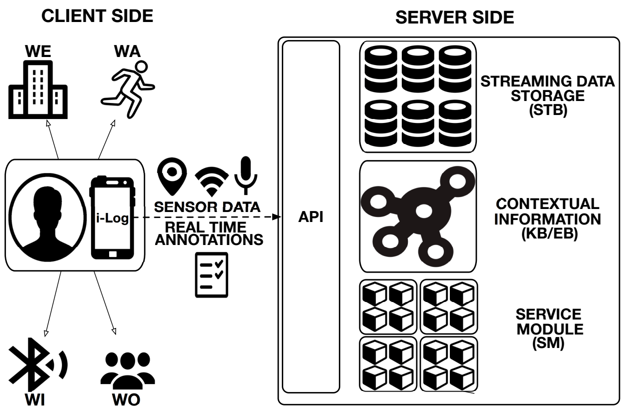


Fig. 2: System architecture showing client and server.

the i-Log infrastructure [24], which is composed of different elements seamlessly interacting with each other to move from sensor data to the level of the context model. i-Log is built according to a client-server architecture. infrastructure as shown in Figure 2.

The interaction with the user is guaranteed via a mobile application with two functionalities. With respect to sensor data input, i-Log can collect up to 30 different sensors, with the final number depending on the actual sensors availability on the user's smartphone. As for human input, i-Log can administer the time diary from Section II-B to users at configurable time intervals, thus adjusting the context information granularity. i-Log has been designed *i)* to be modular and adapt to each smartphone model, especially in terms of sensing strategies for both smartphones and their internal sensors (which can greatly vary among different models), *ii)* to consume as little battery as possible, by devising sensor-dedicated energy consumption strategies and delegating all computation server-side, and *iii)* to ensure users' privacy from data collection to its analysis.

The backend stores the data, and performs the mapping between sensor data and the context model. It has 4 components: 1) an *API* system enabling client and server communication, 2) the *Sensor data storage system (STB)*, which stores the streams of data coming from the sensors in a NOSQL database, to ensure scalability, 3) The *entity storage (EB/KB),* which stores the contextual information as entities, their attributes and their corresponding values, and 4) the *Service Module (SM)*, which analyzes the streams of sensors stored in the STB and updates the context model.

D. Parametrizing Biases

Since our approach relies on humans as a source of information, we must address their reliability. We base our solution on the sociological research showing how different response and behavioral biases cause a lack of congruence between subjects' answers and their true value [20].

Among the many biases, e.g., conditioning, memory bias, and unwillingness to report [23], we focus on two specific types of behavioral biases. The first one is **cognitive bias**, i.e., the inadequate recall of respondents when reporting their time use. The main cause is that the time diary answers are often given retrospectively, causing memory bias [5], or because respondents oversimplify their experience [20]. The second one is **carelessness**, i.e., the set of behaviors leading to hurriedness when answering, which may be caused by, e.g., boredom or annoyance [23]. The use of time diaries administered via smartphone in combination with the collection and exploitation of sensor data allows us to parametrize them. The two behavioral parameters are defined as:

1) $\mathbf{\Delta_{QA}}$, formalizing the cognitive bias, and defined as the time interval (in minutes) from when the question is presented to respondents to when the answer is given. The lower the value, the better in terms of reliability.
2) $\mathbf{\Delta_{A(X,Y)}}$ formalizing carelessness, and defined as the time interval (in seconds) elapsed from when the user starts answering one question of the time diary X and answers another question Y, where $X >= 1$, $Y <= Z$, and Z is the total number of questions and $X < Y$. The higher the value, the better in terms of reliability.

III. Evaluation

We test our approach and the effectiveness of our parameters in the SmartUnitn One experiment, which belongs to the family of projects called ƧMAЯTRAMS[1], detailed in Section III-A. Section III-B shows the strategy for *quantifying the biases* that depend on the user sample, and Section III-C demonstrates how to exploit the human-machine collaboration to *detect the context*. Section III-D explains how the results can be improved by addressing the biases.

Due to lack of space, for our evaluation we focused our attention only on one of the elements of the context, i.e., the spatial one (WE) and in particular on the students' homes location.

A. The SmartUnitn One Experiment

SmartUnitn One is an experiment that aims at understanding the relationship between students' behavior and their academic performance. It has been carried on in November/December 2016 on 72 students enrolled at the University of Trento in the Academic year 2015-2016.

SmartUnitn One lasted two weeks: during the first one, students were asked to answer a time diary on their smartphone about their time use, while the application was collecting sensor data in the background. During the second week, they were only required to have the application running for collecting sensor data. Students were incentivized with a fixed money compensation plus three final prizes assigned to random users that actively participated. We collected a total of user 110 Gb of labeled sensor data from the students for the whole duration of the experiment. Our architecture guarantees privacy is guaranteed through data anonymization in all the steps, from data acquisition to storage and processing. Moreover, before starting the data collection, we obtained the approval from the ethical committee of the University of Trento.

[1]http://trams.disi.unitn.it

B. Quantifying Biases

We adopted the following strategy to quantify the behavioral parameters in the SmartUnitn One use case:

(i) Δ_{QA} has a lower bound of 0 and an upper bound of $150min$ since the questions were available to the user for 2.5 hours. After this time, the question was discarded and considered empty. $\Delta_{A(X,Y)}$ has $X = 1$ and $Y = 3$ since the questions were composed by only 3 sub-questions as in [8]. The lower bound is 0 while the upper bound is left unconstrained.

(ii) Both values are calculated using the *mean* among *all the answers* (not only the ones to the "Home") given by all the students and not of the single user because while designing the experiment, but the attention was also focused on students as a population.

(iii) We used the values calculated with the strategy above as a *threshold*. Answers with values of Δ_{QA} above the threshold and with values of $\Delta_{A(X,Y)}$ below the threshold were considered unreliable.

Concerning Δ_{QA}, Table II presents the distribution of the values across all the 17207 answers generated by the students during the seven days of the experiment. The mean value was identified as 30.4 minutes (Std 37.5). The general trend is that many answers were given for a low value of Δ_{QA} and decreased upon nearing to the maximum limit of 150 minutes to answer (use case defined limit). If we consider the average value of 30.4 minutes obtained from our data, the coverage reaches a higher value of 66.4%.

TABLE II: Distribution of the Δ_{QA} parameter at different time slots.

	10min	**30min**	**60min**	**90min**	**120min**	**150min**
Δ_{QA}	41.5%	24.9%	15.9%	8.4%	5.4%	3.8%

Concerning $\Delta_{A(1,3)}$, we calculated an average value of 8.8 seconds on the 17207 answers (Std 33.2); Table III shows how the values are distributed. We defined six time intervals to represent all the samples,; moreover, at this stage, we were able to set the upper bound as 60 seconds since only 0.04% of the answers were given with a duration above this threshold.

TABLE III: Distribution of the $\Delta_{A(1,3)}$ parameter at different time slots.

	4sec	**8sec**	**12sec**	**20sec**	**40sec**	**60sec**
$\Delta_{A(1,3)}$	36.96%	40.04%	11.14%	7.05%	4.01%	0.76%

C. Detecting Students' Homes

Students were instructed to consider "Home" as their main residential locations. They were supposed to do so regardless of the fact that they were either residents in Trento, commuters or that they usually went back to their hometown during the weekends; this last case would, in fact, imply an additional residence. Nonetheless, if they were staying in other private building, e.g., houses belonging to members of their social circles, then they were supposed to mark them as "Other Private Home". These considerations define a limit for the student's homes to 1 for those students that were resident in Trento and 2 for the commuters.

From a sensor point of view, each answer was associated with a unique location point collected by the smartphone when the question was asked, either from the GPS sensor or calculated through the network Wi-Fi connection. By clustering together the points belonging to all the "Home" answers for each student, we identified a certain number of proxy locations. The clustering process accounts for the GPS inaccuracies. Ideally, the resulting clusters should identify one single location for all the "Home" answers (or two different locations for commuters) if the student replied reliably.

We chose DBSCAN [4] as a clustering algorithm, which is based on the spatial density of the points to cluster. Among its parameters, ε is the maximum distance between two points so that they are considered as being in the same neighborhood. To characterize this value, we averaged the accuracy of all the points collected by the location sensor at the time of the "Home" answer input. In fact, the location collected with i-Log is composed by a tuple of information: $< latitude, longitude, altitude, accuracy, provider >$. The *accuracy* variable contains information about the accuracy of the current measurement, where the lower, the better, represented in meters from the point centered at $< latitude, longitude, altitude >$. It usually varies from few meters when using the GPS up to tens of meters when using the Wi-Fi network provider, or even up to hundreds of meters when using the cellular network provider. In this experiment, the average value for the accuracy was 108.27 meters (Std 314.68 meters).

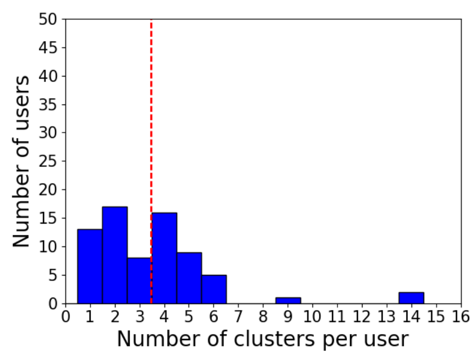
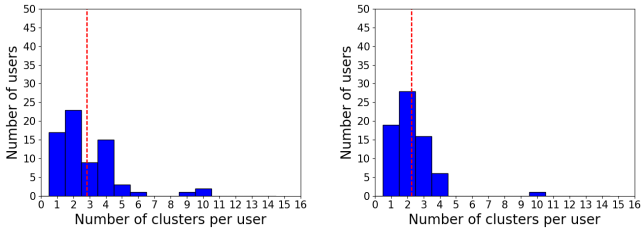


Fig. 3: Distribution of the number of clusters for "Home".

Figure 3 shows the distribution of the number of clusters for the "Home" location across all students, without considering biases. While two homes is a reasonable and expected upper bound, the results show that the majority of students' clusters are included between those of size 1 and 6, with two peaks for 2 and 4 clusters. Moreover, a few students claimed that their home is in 10 or even 14 different places. On average, we identified 3.47 (Std 2.47) homes per student. This result shows that leveraging on the user to annotate his or her data may generate inconsistent answers.

D. The Impact of Biases

To evaluate how the biases affect the user annotations, the system exploits the behavioral parameters to filter out the labels that: *i)* have a higher value than the threshold of Δ_{QA}, and *ii)* a lower value than the threshold of $\Delta_{A(1,3)}$.



(a) Δ_{QA} lower than 30min. The red dashed vertical line is the mean value of 2.81 clusters

(b) $\Delta_{A(1,3)}$ greater than 8.8sec. The red dashed vertical line is the mean value of 2.38 clusters

Fig. 4: Number of clusters distribution for "Home" with:

Figure 4a shows the cluster distribution considering only the location points of the answers with Δ_{QA} lower than the average of 30.4 minutes. In this case, the results improve in accordance to our expectations, with a mean value of 2.81 (Std 1.92) with a total of 4117 points considered belonging to 71 students (63.6% of all points). In this result, not all the students are present since some of them never answered "Home" with such constraints on Δ_{QA}.

In the case of $\Delta_{A(1,3)}$, the threshold value was calculated as 8.8. Figure 4b, shows the cluster distribution for the location points of the answers with $\Delta_{A(1,3)}$ higher than 8.8 sec. Also, in this case, the number of clusters moves towards the expected result and shows a correlation between label reliability and questions completion time. The average number of clusters per student is 2.38 (Std 1.47) calculated out of 1579 points, which represent all the students.

These results show that even the simple strategy we adopted in this experiment produces a significant improvement in the context recognition results when using the behavioral parameters to detect and account for the user unreliability.

The previous results are obtained via what could be considered a standard supervised approach. We present, as an additional result, the comparison between our results and those generated by an entirely unsupervised algorithm that does not leverage on the user helping in the recognition. Based on [16], the best window of time to detect a people in their own house is between 07:00-07:30 AM. We then run the same DBSCAN algorithm used for the other analysis on all the location points of every user in that timeslot over all the two weeks of the experiment regardless of the labels. In this case, the total points clustered are 11784, which belong to 63 students. Apparently, 9 of them did not have any location point in that timeslot. The distribution of the resulting clusters is shown in Figure 5. The results, in this case, are worse than the previous ones since the distribution is flatter and distributed towards higher values of the number of clusters. In fact, the mean value across all the users is 6.73 homes (Std 9.83).

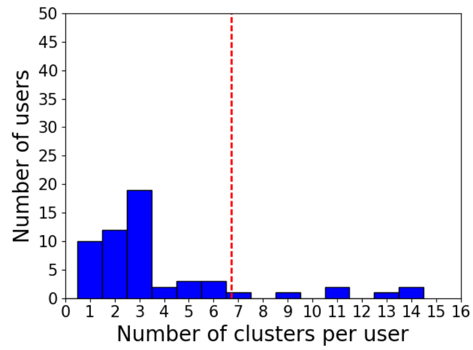


Fig. 5: Distribution of the number of clusters for students' homes calculated in an unsupervised approach.

The results of the evaluation process are summarized in Table IV. What emerges is that for the situation we addressed, the unsupervised technique provided the worst results and this motivates the need for the human-machine collaboration. Accounting for users to annotate their data is needed and improves the results. In fact, when considering the human unreliability by filtering out the unreliable annotations, the results of the context recognition process improve even more, especially concerning the cognitive bias.

TABLE IV: Comparison of the results of the 4 approaches for determining the user Home location context.

Approach			**Mean**	**Std**
Unsupervised			6.73	9.83
Supervised	Biased		3.47	2.47
	Unbiased	Δ_{QA}	2.81	1.92
		$\Delta_{A(1,3)}$	2.38	1.47

IV. Related Work

Active learning, where a learning algorithm is able to interactively query the user to obtain the desired outputs at new data points [18], is used to alleviate the task of labeling data in supervised approaches, which rely on expert users [22]. This means that the process iterates as the data is collected, unlike traditional supervised machine learning, where the labeling is done offline [21]. In the area of activity recognition, active learning is being used especially in smart homes, given the limited range of activities to be recognized. For instance, in a very recent work [14], a new model that discovers unseen new activities and includes the new activity class in the supervised learning model is proposed. The main advantage is that the model itself is dynamic in finding new activities, whereas the standard ones only consider a set of pre-defined activities.

Other areas are focused on real-life scenarios, thus dealing with non-expert users; for instance, participatory sensing [15]. The main idea is to have users collect and share sensed data from their surroundings using their mobile phones. One increasingly important line of research has been understanding the best approaches to elicit not only user data but their annotations as well. [3] analyze three approaches for the data collection, i.e., *Participatory* (PART), *Context-Triggered In Situ* (SITU), and *Context-Triggered Post Hoc* (POST). Participatory refers to users actively collecting data for the whole experiment. Context triggered in situ and post hoc refer to obtaining annotations from user data during the experiment either when a specific condition is triggered or to prompt users afterwards to obtain retrospective annotations, respectively. These approaches were experimented on 37 users that had to record and annotate their travelling habits. The results suggest the PART approach is the most effective one since it produces a larger amount of activity data and with less noise, although SITU and POST leads to more activity recordings.

The main issue with current approaches is that they ignore the reliability of users' annotations. In participatory sensing, the main measure of quality is the amount of activities reported and data collected on a specific aspect of the daily life of users, e.g., travelling [3]. In the case of supervised machine learning, expert may be enough for correct annotations; however, these approaches cannot scale outside of controlled environments.

Our definition of context is quite different from that usually found the pervasive computing community (see [1] for a survey). To us, context is an intermediate representation layer which bridges the human and machine representations of the world, rather than an aggregation mechanism exploited by machines to reason about sensor data [1]. As such, our notion of context does not need to represent uncertainty. We implement reasoning by representing context as an ontology and by exploiting efficient Modal/Description Logic inference engines [10], [6].

V. Conclusions

In this paper, we proposed a parametrization of behavioral biases to improve human-machine collaboration for context recognition. Humans are needed in the context recognition loop because of the different way the machine represents the world; without them, the machine cannot generate meaningful information. However, users in real life scenarios cannot be assumed to be entirely reliable in the annotation task. We can quantify the user label reliability and account for it in the context recognition process to improve the final accuracy. We evaluated our approach on the SmartUnitn One dataset, proving that an unsupervised technique for detecting the user's home location shows poor results. Instead, a supervised approach with user annotations improves this results, showing that the user feedback is needed. Finally, when accounting for the user unreliability during the analysis, the results show an increased accuracy and follow our expectations.

Acknowledgment

The project leading to this application has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 732194.

References

[1] C. Bettini, O. Brdiczka, K. Henricksen, 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.

[2] P. Bouquet and F. Giunchiglia, "Reasoning about theory adequacy. A new solution to the qualification problem," *Fundamenta Informaticae*, vol. 23, no. 2, 3, 4, pp. 247–262, 1995.

[3] Y.-J. Chang, G. Paruthi, H.-Y. Wu, H.-Y. Lin, and M. W. Newman, "An investigation of using mobile and situated crowdsourcing to collect annotated travel activity data in real-word settings," *International Journal of Human-Computer Studies*, vol. 102, pp. 81–102, 2017.

[4] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise." in *Kdd*, vol. 96, no. 34, 1996, pp. 226–231.

[5] V. A. Freedman, J. Broome, F. Conrad, and J. C. Cornman, "Interviewer and respondent interactions and quality assessments in a time diary study," *Electronic international journal of time use research*, vol. 10, no. 1, p. 55, 2013.

[6] E. Giunchiglia, F. Giunchiglia, R. Sebastiani, and A. Tacchella, "More evaluation of decision procedures for modal logics," *KR*, vol. 98, pp. 626–635, 1998.

[7] F. Giunchiglia, "Contextual reasoning," *Epistemologia, special issue on I Linguaggi e le Macchine*, vol. 16, pp. 345–364, 1993.

[8] F. Giunchiglia, E. Bignotti, and M. Zeni, "Personal context modelling and annotation," in *Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on*. IEEE, 2017, pp. 117–122.

[9] F. Giunchiglia and M. Fumagalli, "Concepts as (recognition) abilities," in *Formal Ontology in Information Systems: Proceedings of the 9th International Conference (FOIS 2016)*, vol. 283. IOS Press, 2016, p. 153.

[10] F. Giunchiglia and R. Sebastiani, "Building decision procedures for modal logics from propositional decision procedures: The case study of modal K (m)," *Information and Computation*, vol. 162, no. 1-2, pp. 158–178, 2000.

[11] F. Giunchiglia, M. Zeni, E. Gobbi, E. Bignotti, and I. Bison, "Mobile social media and academic performance," in *International Conference on Social Informatics*. Springer, 2017, pp. 3–13.

[12] M. Hellgren, "Extracting more knowledge from time diaries?" *Social Indicators Research*, vol. 119, no. 3, pp. 1517–1534, 2014.

[13] E. Hoque and J. Stankovic, "AALO: Activity recognition in smart homes using active learning in the presence of overlapped activities," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on*. IEEE, 2012, pp. 139–146.

[14] 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.

[15] S. S. Kanhere, "Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces," in *Mobile Data Management (MDM), 2011 12th IEEE International Conference on*, vol. 2. IEEE, 2011, pp. 3–6.

[16] J. Krumm, "Inference attacks on location tracks," *Pervasive computing*, pp. 127–143, 2007.

[17] 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.

[18] B. Settles, "Active learning literature survey," *University of Wisconsin, Madison*, vol. 52, no. 55-66, p. 11, 2010.

[19] K. J. Shelley, "Developing the American time use survey activity classification system," *Monthly Lab. Rev.*, vol. 128, p. 3, 2005.

[20] R. Tourangeau, L. J. Rips, and K. Rasinski, *The psychology of survey response*. Cambridge University Press, 2000.

[21] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 2013.

[22] A. R. Webb, *Statistical pattern recognition*. John Wiley & Sons, 2003.

[23] B. T. West and J. Sinibaldi, "The quality of paradata: A literature review," *Improving Surveys with Paradata*, pp. 339–359, 2013.

[24] M. Zeni, I. Zaihrayeu, and F. Giunchiglia, "Multi-device activity logging," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 2014, pp. 299–302.