Vocal-Diary : A Voice Command based Ground Truth Collection System for Activity Recognition

Enamul Hoque
Center for Wireless Health
University of Virginia
Charlottesville, Virginia, USA
eh6p@virginia.edu

Robert F. Dickerson
Center for Wireless Health
University of Virginia
Charlottesville, Virginia, USA
rfdickerson@gmail.com

John A. Stankovic
Center for Wireless Health
University of Virginia
Charlottesville, Virginia, USA
stankovic@cs.virginia.edu

ABSTRACT

We present *Vocal-Diary*, a voice command based ground truth collection system that uses grammar based commands from residents to log start and end of activities. *Vocal-Diary* ensures robustness in the presence of sounds from different environmental noise and day-to-day conversation by using two-way acknowledgement and integrating speaker recognition in the pipeline. *Vocal-Diary* also utilizes the sensor data produced by the underlying activity recognition system to query residents periodically to check if they forgot to log any activity. Evaluation shows that *Vocal-Diary* increases precision by at least 40% and recall by at least 10% compared to a system that uses voice command recognition without any acknowledgement and speaker recognition.

Keywords

Ground Truth; Activity Recognition; Deployment;

Categories and Subject Descriptors

H.1.2 [**Human information processing**]

1. INTRODUCTION

Activity recognition systems based on in-home sensors are used for different applications such as home health care, energy monitoring, and security. Specially, in-home monitoring of elderly people is one of the most fascinating promises of wireless health domain, and inferring daily activities based on in-home sensors enables such remote monitoring. The goal of these monitoring applications is to provide the daily activities details of the residents to caregivers or family members so that any health related problems may be detected as soon as possible. There is a variety of sensors that are currently used for activity recognition including motion sensors, door sensors, contact sensors in objects of daily use, bed / chair sensors, and on-body sensors. The goal of the activity recognition systems is to learn the relationship among different daily activities and different sensors that are installed in home based on training data.

To train activity recognition systems, we need labeled activity episodes (i.e., the starting and finishing times of each daily activity) during the training period. These accurate annotations of daily activities are called ground truth. Accuracy of activity recognition systems depends on sufficient amount of ground truth for training. Ground truth collection requires active participation from residents which is challenging. The challenges include ensuring comfort and privacy of residents, maintaining accuracy in ground truth, considering different real home scenarios that may cause errors in activity labeling, and addressing the reality that residents may forget to log activities. In this paper, we address this technical problem of collecting ground truth for daily activities for training activity recognition systems by considering the above mentioned challenges in real home settings. Our proposed solution enables proper training of activity recognition systems which will facilitate many wireless health applications.

Existing ground truth collection systems suffer from different shortcomings such as use of cameras [17, 15] which is a privacy concern and very tedious, real-time logging by the user [2, 6] which is not suitable in real home settings for sufficiently long training duration, manual annotation of activity labels from sensor data based on some rules [12, 3] which is time-consuming and may not be always accurate, and necessity of wearing a microphone that may be uncomfortable for users [9, 8]. Also, none of the existing ground truth collection systems address the fact that residents may forget to log the begin and / or end of activities occasionally.

We present *Vocal-Diary*, a voice command based ground truth collection system where residents log activities by specific voice commands. *Vocal-Diary* is privacy-aware, and robust to different environmental noise in the home and day-to-day conversations among the residents. To increase robustness, *Vocal-Diary* uses a speaker recognition system that is trained with the voice segments of all the residents in a home to ensure that the voice commands are spoken by the residents. Also, *Vocal-Diary* utilizes the sensor data produced by the underlying activity recognition system to query a resident periodically to check if he / she forgot to log any activity.

The main contributions of *Vocal-Diary* are:

1) A novel ground truth collection system to collect accurate activity labels by listening to specific voice commands from residents with help of one / more microphone(s) in the home. The system will be made publicly available so that other research groups can use it.

2) The novelty of *Vocal-Diary* includes the use of two-way acknowledgement for listening to voice commands from a resident and integration of speaker identification in the pipeline for robustness. *Vocal-Diary* is privacy-aware, because it only processes voice segments from residents that have a predefined structure (e.g., <sys-

Wireless Health '14 Oct 29-31, 2014, Bethesda, MD, USA

tem 'activity_name start'>) and raw voice of residents is never recorded. *Vocal-Diary* is also comfortable for a resident to use as it is not necessary to carry a microphone (in each room, one microphone is placed in a suitable place), and also because a resident does not need to manually turn the microphone on / off when giving voice commands.

3) Another novel feature of *Vocal-Diary* is to query residents periodically to check if they forgot to log any activity by voice commands with the help of the sensor data produced by the underlying activity recognition system.

4) We evaluate *Vocal-Diary* by deploying in three homes (two single-resident, one double-resident) for one month each. Results show that *Vocal-Diary* increases precision by at least 40% and recall by at least 10% compared to a one-way voice command recognition based system.

The rest of the paper is organized as follows. Section 2 presents a summary of related works and their shortcomings. Section 3 details different components of the *Vocal-Diary* system. Section 5 presents the advantages of *Vocal-Diary* . Section 6 describes the details of experiments and evaluation results. We conclude in Section 7.

2. RELATED WORK

A widely used method for collecting ground truth is using cameras and annotating activity episodes from the recording [16, 17, 15]. However, cameras suffer from privacy concerns, and it is very time consuming to go through the recordings to identify and annotate activities. Another common approach [3, 12, 5] is to go through the sensor firings during the training period and annotate activity labels based on some rules. This approach is also very time consuming and there is a magnitude of error which is not acceptable in ground truth. Authors in [2, 6] collect ground truth in real time as participants do different activities in a controlled lab setting. However, doing so in real deployments for a sufficiently long training period is impractical.

Another way to collect ground truth requires users to log each activity manually. However, this technique is not very accurate as pointed out in [10, 14]. The main reason is that over long periods of data collection, there are inaccurate timestamping from retroactive self reporting and also missing logged events. Logan et al. [11] use an experience sampling method to collect ground truth. In this method, periodic queries are sent to the smart phone of residents asking what activity they are doing. Authors in [7] use a combination of three methods that include experience sampling, a hand-written log where the resident records different activity labels at different times, and a collection of snap shots that the resident takes using the camera of smart phone. All these methods require the resident to actively interact with a smart phone / laptop / diary to log. Such high level interaction in a real home for a sufficiently long period is uncomfortable for residents.

We believe that the most comfortable way to collect ground truth is through voice commands. Interacting by voice has been proved acceptable in many applications (e.g., voice-based search, navigation, email). In the system developed by Kasteren et al. [9, 8], a resident wears a bluetooth headset and presses a button to give specific voice commands (implemented using Microsoft Speech API (SAPI [1])). However, this system has the discomfort that residents have to wear a headset before giving voice commands. Also, they may forget to press the switch. From our experiments, we find that if we always keep the microphone open and / or if the microphone is far from the user, SAPI erroneously records many voice commands when the user did not speak. This is due to noise in the environment or other sounds (e.g., TV, day-to-day conversation). An ideal system should perform accurately even in such scenarios without a resident needing to turn on / off a switch. This is what *Vocal-Diary* accomplishes.

Also, none of the above systems address the issue that residents may occasionally forget to log start and / or end of activities. This may result in incomplete ground truth which can affect the training of activity recognition systems.

3. SYSTEM DESCRIPTION

Here we describe the two key components of *Vocal-Diary*: voice command recognition and speaker recognition. Following that we explain how these two components are used together to build the end-to-end system of *Vocal-Diary*.

3.1 Voice Command Recognition

We implement the voice command recognition program using Microsoft Speech API (SAPI [1]) in C# using the .NET environment. We use a recognition grammar that employs the following usage pattern: a resident begins an activity saying "system <activity name> start". *Vocal-Diary* then plays back a pre-recorded audio file that asks for confirmation of the same activity being started such as "you are beginning to <activity name>". If the system understood correctly, the resident then says "system yes" as a two way confirmation.

We find from tests that two-way acknowledgement is necessary since the microphones are often moderately far from residents and there may be different environmental noise, so the accuracy of voice command recognition is not always perfect. When the resident finishes the activity, he / she needs to say "system <activity name> end", and the same confirmation above is used. For start / end of any activity, the activity name, timestamp, and start / end status are logged. The recognition grammar consists of a fixed vocabulary of activities.

3.2 Speaker Recognition

Despite using the voice command recognition with two-way acknowledgement, sometimes different environmental noise or daily conversation may be erroneously recognized as commands from residents. To address this, we implement a speaker recognition program to classify such noise as not being spoken by the residents. We use the open-source MARF framework [13] to implement the speaker recognition program in Java. We convert the speaker recognition program in Java to a .jar executable and invoke it from the voice command recognition program in C#.

The speaker recognition program needs training. At the start of a deployment, each resident speaks the voice segments "system <activity name> start" and "system <activity name> end" for each activity, and the two segments "system yes" and "system no". These voice segments from each resident along with few segments that contain different sources of noise (sounds from kitchen appliances, footsteps, opening and closing of doors, washing machine, and TV) construct the training data. We have seen from our experiments that ambient noises such as these are often detected as voice commands by the Microsoft speech API.

The noise sounds are recorded at the start of deployment in each home. In our experiments, residents give voice commands when within $1-5$ feet of a microphone (during both training and testing). Note that we do not consider the scenario when a resident may give a command when there is ambient noise in the environment (e.g., when the TV is on). However, we believe *Vocal-Diary* can also address such scenarios by training with such recordings (i.e., giving commands when TV or microwave is on). Currently the utility of the speaker recognition system is in identifying the cases when sounds from the environment are wrongly detected as commands

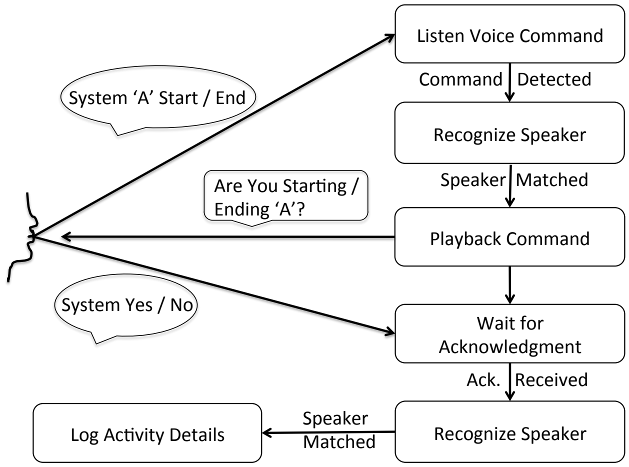

Figure 1: *Voice-Diary* End-to-End System

by the voice recognition module.

The MARF framework provides options for different pre-processing (normalization, end-pointing, high, low, and band pass filters), feature selection (e.g., fast fourier transform, linear predictive coding), and classification (e.g., neural networks, nearest neighbors) techniques. We train with different combinations and use the combination that performs the best on the training data. This combination may be different for different homes. The trained program is used for speaker recognition in real time.

3.3 End-to-End System

Figure 1 shows the end-to-end system. 'Listen Voice Command' module waits for a voice command in the specific format described above. When it listens to a command (correctly or incorrectly), it invokes the 'Recognize Speaker' program to check if the voice belongs to any of the residents. If it does not, then *Vocal-Diary* logs this error in a separate debug file. If it does, then the 'Playback Command' finds the pre-recorded audio file for that specific activity's start / end. Note that for each activity there are two pre-recorded audio files in the system that plays back the query whether the resident actually started / finished a particular activity. After playing the correct audio file, 'Wait for Acknowledgment' module waits for acknowledgement.

After listening to the audio played back, the resident acknowledges by saying "system yes" if it was the correct command. Else the resident says "system no". If the resident says 'no' or does not say anything, *Vocal-Diary* does not write anything in the activity log, but it logs this error in a separate debug file. If the resident acknowledges yes, *Vocal-Diary* once again verifies if this voice segment belongs to the same resident of the home. If it does, then finally *Vocal-Diary* writes the timestamp, activity name and start / end status in the activity log.

Ideally, there should be one microphone and one speaker in a suitable place in each room of the home. *Vocal-Diary* logs voice commands collected by each microphone in separate files and combines all of them offline. We can use wireless microphones and speakers in each room that send data to a laptop in the home. Another approach can be using a USB microphone and a USB speaker connected to a Beaglebone with flash memory in each room as used in [4]. *Vocal-Diary* is totally privacy-aware. Therefore, residents should be comfortable with the presence of a microphone in each room. However, if a resident does not want any microphone in a particular room / rooms, the voice commands have to be given in a room where there is a microphone.

4. WHAT IF RESIDENTS FORGET

Vocal-Diary , or any other ground truth collection system requires a resident to inform the system (by voice commands in case of our system) before starting and ending each activity. However, in reality a resident may forget to do so occasionally. Therefore, *Vocal-Diary* also reminds a resident periodically to give voice commands.

Vocal-Diary utilizes the sensors that the underlying activity recognition systems use (e.g., motion, contact, door, bed sensors). If *Vocal-Diary* can access such data, then it works in the following way to remind a resident:

a) After entering each room, if a resident uses one or more objects in the room (this can be detected from the sensor firings) but does not log any activity with voice command within $ENTRY\_THR$ minutes, then *Vocal-Diary* queries the resident "Have you forgotten to log an activity?". In reply, a resident can ignore the query or log an activity.

b) If a resident ignores a query, then *Vocal-Diary* does not query again within $REPEAT\_THR$ minutes which is configurable. This is to ensure that the queries do not become a nuisance for residents. If the resident does not want to listen any such query at all, *Vocal-Diary* is configured accordingly.

Note that this feature is dependent on the availability of the sensor data generated by the underlying activity recognition system which is true in one of our deployments.

5. ADVANTAGES

5.1 Privacy-Aware

Vocal-Diary is privacy-aware, as it does not record day-to-day conversation. It only listens to voice commands in specific format. Moreover, whenever a resident gives voice commands, their raw voice is not recorded. Only the timestamps, activity names and start / end status are logged. The audio files recorded during training contain the voice of a resident. However, we delete them as they are not used after training.

5.2 Robust

Vocal-Diary is robust in the presence of different environment noise that may arise in a real home settings by ensuring that such noise are not recorded as voice commands. Note that, *Vocal-Diary* has not been trained and tested for scenarios where residents may give voice commands at the same time when any such noise source is active. However, with adequate training, *Vocal-Diary* can also be used for such scenarios. In this paper, we show that with its two way acknowledgement and speaker recognition features, *Vocal-Diary* identifies the cases where such noise can be identified as voice commands (when residents actually give no voice commands) and removes such cases. Such noise may include, but not limited to sound from TV, music player, dish washer, washing machine, microwave, coffee maker, footsteps, opening / closing of doors. Such noise can be recognized as voice commands by Microsoft Speech API SAPI [1]. However the use of speaker recognition and two-way acknowledgement by *Vocal-Diary* makes sure that they are not logged as ground truth. Also, residents can have regular conversation with each other, with visitors, or with someone over phone.

5.3 Ease of Use

Unlike [9, 8], *Vocal-Diary* does not require a resident to wear any headset. A microphone and a speaker can be placed in any suitable location in the room because *Vocal-Diary* is robust even in the presence of different environmental noise. Deploying microphones and speakers in each room takes little time in each home, and the time to disassemble the system after the deployments is also minimal. In our experiments, residents give voice commands when within $1-5$ feet of a microphone and when there is no ambient noise from the environment (e.g., TV). Therefore, the microphones and speakers can be placed anywhere in the home as long as the residents are comfortable in giving commands when within $1-5$ feet of the microphones.

We believe with adequate training, *Vocal-Diary* can also work in the scenarios when residents give commands in the presence of noise. In such cases, the training set has to include samples when the residents give voice commands in the presence of different noise sources in the environment. Another advantage of *Vocal-Diary* is that residents do not need to turn on / off the microphone before speaking each command as the microphone is always on. However, it only listens to specific commands and does not record any other noise or conversation.

5.4 Supports Multi-Resident Homes

In case of multi-resident homes, it may be important to collect ground truth of activities from both residents. Also, for different residents, the set of activities that need to be monitored may be different. If there are multiple residents in a home, *Vocal-Diary* is initially trained with each of the resident's voice for the corresponding sets of activities. After training, all residents can interact with *Vocal-Diary* in the same way. *Vocal-Diary* identifies each voice based on the speaker recognition program and adds the identifier along with other information in the log file.

6. EVALUATION

The evaluation consists of three parts. First, we evaluate how accurately *Vocal-Diary* recognizes voice commands. Then we evaluate the effectiveness of querying the residents. Finally, we investigate the feasibility of voice commands as a ground truth collection mechanism.

We deployed *Vocal-Diary* in two single-resident and one two-resident homes for one month each. In each deployment, we used wireless microphones and speakers in each room. The residents agreed to participate in the study voluntarily, and no monetary compensation was provided to them. In one of the single-resident homes ('Single Resident Home 1'), sensors were also deployed for recognizing activities. Therefore, for this home, by deploying *Vocal-Diary* for additional days / months, we were able to evaluate the effectiveness of querying the residents (Section 6.2) and the feasibility of using *Vocal-Diary* for ground truth (Section 6.3) collection. However, no sensors (except microphones) were deployed in the other two homes.

For evaluation, raw audio of any voice command detected by *Vocal-Diary* (correctly or incorrectly) is saved as a .wav file so that we can listen offline to verify if it is a voice command. Note that the files are only saved for evaluation purpose. We listen to the recordings to calculate how many activity logs were actually voice commands, i.e., number of true positives (TP), how many voice commands were not logged as activities, i.e., number of false negatives (FN), and how many recordings containing noise or other conversation were logged as activities, i.e., number of false positives (FP). Note that, our baseline for comparison is the voice commands detected by Microsoft SAPI. If there are instances where a resident gives a voice command, but Microsoft SAPI does not detect it at all, then *Vocal-Diary* also fails to recognize such instances and they are not reflected in the false negatives (FN) we calculated. However,

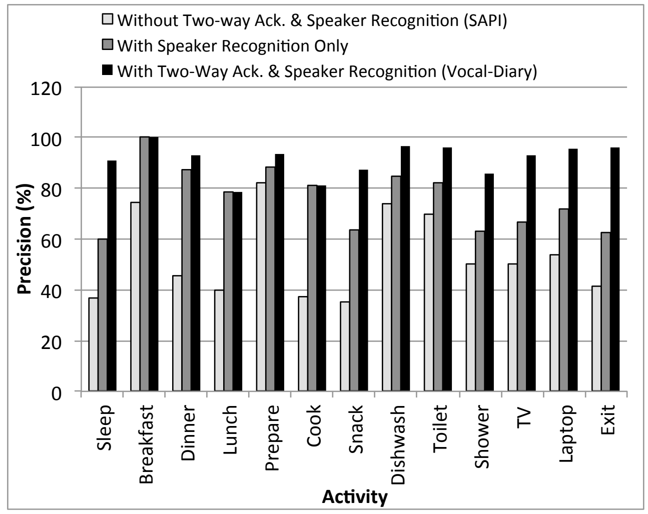

Figure 2: Precision values for single-resident home 1. Adding speaker recognition increases the precision for all activities, and adding two-way acknowledgement increases it more.

in our tests, we never found any such case where a voice command is entirely ignored by SAPI. A voice command may be detected as the wrong one, or noise sound may be detected as voice commands. *Vocal-Diary* aims to reduce such errors, and our evaluation is to show the accuracy of that.

As evaluation metrics, we use precision and recall which are defined in Equation 1. We got for IRB approval for such experiments. The residents knew their voice recordings are recorded for evaluation. Note that, the raw voice recordings are only recorded for evaluating the performance of *Vocal-Diary*.

$$\begin{aligned} precision &= \frac{TP}{TP+FP} \\ recall &= \frac{TP}{TP+FN} \end{aligned} \tag{1}$$

6.1 Voice Command Accuracy

Figure 2 shows precision values for single-resident home '1'. For all activities, the basic SAPI based system (which has no two-way acknowledgement and no speaker recognition) has very low precision values. This is mainly due to different sounds generated by environmental noise. Using speaker recognition increases the precision values significantly for all activities as it helps in removing the false positives. False positives may also occur due to an actual voice command being detected as a different voice command and / or other day-to-day conversations by residents being detected as voice commands. Integrating two-way acknowledgement system, which ensures that *Vocal-Diary* does not log any detected command without acknowledgement, helps in removing such false positives. Still there may exist some false positives, because either sometimes noise is detected as "system yes" and/or error in speaker recognition.

Figure 3 shows the recall values for the same home. The SAPI based system has false negatives. This is because sometimes a voice command is detected as another one. This is due to the variation in the ways different people utter the same word. Because SAPI does not train per speaker, such errors are not surprising. Adding the speaker recognition feature cannot remove all such er-

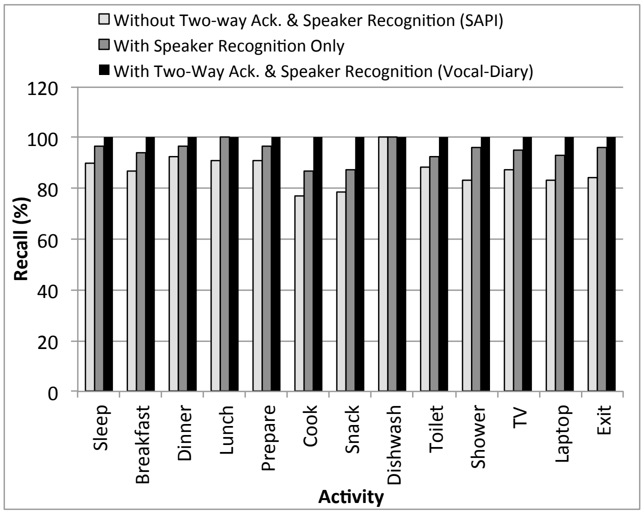

Figure 3: Recall values for single-resident home 1. *Vocal-Diary* achieves 100% recall for all activities with the help of speaker recognition and two-way acknowledgement features.

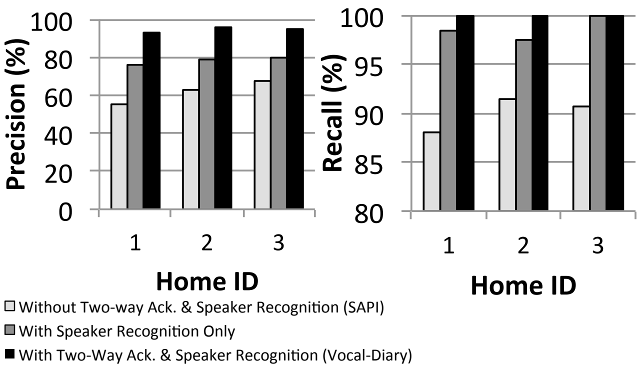

Figure 4: Avg. Precision and Recall over all activities. Home 3 has two residents; Home 1 and 2 have single resident.

rors. However, *Vocal-Diary* removes such false negatives with the help of two-way acknowledgments and achieves 100% recall for this single-resident home.

Figure 4 shows average precision and recall values for all activities in the three homes; home 3 is the double-resident home. For all three homes, *Vocal-Diary* increases precision values significantly. Also, *Vocal-Diary* removes false negatives with the help of two-way acknowledgments and achieves 100% recall for all homes. *Vocal-Diary* increases precision by at least 40% and recall by at least 10% compared to a SAPI based system without two-way acknowledgement and speaker recognition.

In calculating precision and recall for the double-resident home, *Vocal-Diary* is considered accurate if it correctly differentiates sounds caused by environmental noise from voice of residents. Whether it can assign a voice command to the correct resident is not considered. However, if we consider that, the average precision and recall values for all activities drops to 88% and 95%, respectively. MARF framework does not support advanced features (e.g., MFCC) and classifiers (e.g., support vector machine). Implementing a speaker recognition program using these would increase accuracy.

6.2 Effectiveness of Querying Residents

To evaluate the effectiveness of querying the residents, we deploy the system in single-resident home '1' for 15 days. During this time, *Vocal-Diary* had access to all the sensor data from the activity recognition system that include motion sensors in each room, door sensors, contact sensors in various objects of daily use (e.g., microwave, freezer, sink, toilet, shower), and pressure pads in bed and chair. We set the value of $ENTRY\_THR$ as 1 minute and the value of $REPEAT\_THR$ as 5 minutes. During the 15 days, the resident did not log activity start / end times by voice commands on purpose for 25 times; so the experiments here are in a controlled setting. *Vocal-Diary* accurately detected all the 25 instances and queried the resident for the ongoing activity status. This shows that *Vocal-Diary* can help in logging activities by querying the residents when they forget to do so.

However, *Vocal-Diary* queried the resident a total of 39 times, out of which 25 queries were correct as mentioned above. The remaining 14 instances, *Vocal-Diary* queried the resident when no activity was being started / ended. The reason is that sometimes the resident moves around in a room (which causes the motion sensors to fire) or uses some objects that are not part of any specific activity. Most of these false positives were caused by the motion sensors. If we do not consider the motion sensors in the querying system, the number of false positives reduces from 14 to 5. Still, all the 25 correct cases are identified. Therefore we modified our system so that the querying system does not use data from motion sensors if there are other sensors available in a room.

6.3 Feasibility of Voice Commands to Collect Ground Truth

So far we have shown results from evaluation of how accurate *Vocal-Diary* is in recognizing voice commands and in querying residents if they forget to give voice commands. However, we also need to evaluate the feasibility of using such a system for a long period. If the residents do not give voice commands or do not reply to the queries, those scenarios will not be reflected in the above experiments. Therefore, we evaluate *Vocal-Diary* by deploying it along with an underlying activity recognition system (that uses in-home sensors) for three months continuously in single-resident home '1'. During this deployment, the feature of querying the resident was not used.

Here, we need ground truth to evaluate our ground truth collection system. From the sensor firings, all activity segments were manually annotated offline as in [3] with feedback from the resident who also logged activities using *Vocal-Diary*. Results show that there were 992 total activity instances during the three months of which 59 activity instances were not logged by voice commands. Therefore, the resident used *Vocal-Diary* to log activities in more than 94% of the cases. Given the effectiveness of querying residents discussed above, we hypothesize that most if not all of the 6% of missed activities would have been captured if the querying was used. In the future we plan to compare *Vocal-Diary* to ground truth detection based on offline logging and cameras.

7. CONCLUSIONS

Vocal-Diary is a privacy-aware, easy-to-use and robust ground truth collection system based on voice commands which shows high accuracy in evaluation. The accuracy in voice command recognition is achieved by two-way acknowledgement and speaker recognition. The ease of use, robustness, and high accuracy come at the cost of additional training and microphones in each room. However, the microphones are inexpensive (with Beaglebones [4]) and the training effort is minimal (few minutes per resident). Comprehensive evaluation is necessary for the effectiveness of querying residents.

The novelties presented in this paper in voice command recognition based on microphones placed in home are also applicable for other home heath care applications that need to record voice samples in specific format from the residents such as collecting survey responses from the residents (e.g., PHQ-9 questionnaire for depression patients), and controlling home / medical appliances using voice commands. In future, we plan to evaluate *Vocal-Diary* with more deployments and longer duration for each deployment with emphasis on evaluating the effectiveness of querying the residents, and comparison of *Vocal-Diary* with other ground truth collection systems in long-term data.

Acknowledgments

This paper was supported, in part, by NSF Grants CNS-1319302 and CNS-1239483, and a gift from PARC, Palo Alto.

8. REFERENCES

[1] Microsoft speech api (sapi). `http://msdn.microsoft.com/en-us/library/ms723627(v=vs.85).aspx`, 2014.

[2] M. Buettner, R. Prasad, M. Philipose, and D. Wetherall. Recognizing daily activities with rfid-based sensors. In *UbiComp*, 2009.

[3] D. Cook and M. Schmitter-Edgecombe. Assessing the quality of activities in a smart environment. *Methods of Information in Medicine*, 48(5), 2009.

[4] R. Dickerson, E. Hoque, P. Asare, , S. Nirjon, and J. Stankovic. Resonate: Improving speech classification in home environments using reverberant environment simulation. In *IPSN*, 2014.

[5] T. Gu, S. Chen, X. Tao, and J. Lu. An unsupervised approach to activity recognition and segmentation based on object-use fingerprints. *Data Knowl. Eng.*, 69(6), 2010.

[6] M. R. Hodges, N. L. Kirsch, M. W. Newman, and M. E. Pollack. Automatic assessment of cognitive impairment through electronic observation of object usage. In *Pervasive*, 2010.

[7] T. Huynh, M. Fritz, and B. Schiele. Discovery of activity patterns using topic models. In *UbiComp*, 2008.

[8] T. Kasteren, G. Englebienne, and B. J. A. Kröse. Activity recognition using semi-markov models on real world smart home datasets. *J. Ambient Intell. Smart Environ.*, 2(3), 2010.

[9] T. Kasteren, A. Noulas, G. Englebienne, and B. Kröse. Accurate activity recognition in a home setting. In *UbiComp*, 2008.

[10] P. Klasnja, B. L. Harrison, L. LeGrand, A. LaMarca, J. Froehlich, and S. E. Hudson. Using wearable sensors and real time inference to understand human recall of routine activities, 2008.

[11] B. Logan, J. Healey, M. Philipose, E. M. Tapia, and S. Intille. A long-term evaluation of sensing modalities for activity recognition. In *UbiComp*, 2007.

[12] D. Lymberopoulos, A. Bamis, and A. Savvides. Extracting spatiotemporal human activity patterns in assisted living using a home sensor network. In *PETRA*, 2008.

[13] S. A. Mokhov. Evolution of marf and its nlp framework. In *Proceedings of the Third C* Conference on Computer Science and Software Engineering*, C3S2E '10, 2010.

[14] A. Möller, M. Kranz, B. Schmid, L. Roalter, and S. Diewald. Investigating self-reporting behavior in long-term studies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013.

[15] T. Mori, A. Takada, Y. lwamura, and T. Sato. Automatic human life summarization system in sensory living space. In *International Conference on Systems, Man and Cybernetics*, 2004.

[16] E. M. Tapia, T. Choudhury, and M. Philipose. Building reliable activity models using hierarchical shrinkage and mined ontology. In *Pervasive*, 2006.

[17] J. Wu, A. Osuntogun, T. Choudhury, M. Philipose, and J. M. Rehg. A scalable approach to activity recognition based on object use. In *ICCV*, 2007.