SARRIMA: Smart ADL Recognizer and Resident Identifier in Multi-resident Accommodations

Ifat Afrin Emi
Dept. of Computer Science
University of Virginia
Charlottesville, VA USA
Email: iae4qb@virginia.edu

John A. Stankovic
Dept. of Computer Science
University of Virginia
Charlottesville, VA USA
Email: stankovic@cs.virginia.edu

***Abstract*— Systems for measuring Activities of Daily Livings (ADL) play a significant role in home health-care. The ability of performing ADLs successfully is used as an important factor in deciding treatments and services for patients and elderly citizens. However, most of these systems are designed for single-resident homes. The presence of multiple people creates higher numbers of parallel and overlapping activities, and introduces additional complexities in defining and recognizing activity instances. We present SARRIMA, a system that recognizes activity instances and assigns those activities to a person in 2-resident homes using only passive sensors. We evaluate the efficiency of SARRIMA in two different public datasets (data from real homes) with multiple residents. On the average SARRIMA detects more than 97% of the activity instances. We also show how the person assignment accuracy varies as a function of the similarity of behavior of the 2 people living together and of the types of passive sensors installed.**

***Index Terms*—Activity of Daily Living, Multiple Residents, ADL Recognition, Person Identification, Wireless Sensor Network.**

I. Introduction

Activities of Daily Living (ADL) refer to the daily self-care activities performed by an individual; example includes eating, sleeping, showering, and toileting. On the other hand, Instrumental ADL (IADL) refer to more complex set of activities that are not fundamental, but very important for independent living - such as preparing dinner, cleaning house, talking in phone, and managing finance. The ability to perform ADL and IADL successfully is considered as a major criterion to access the condition of stroke patients and patients suffering from depression, Alzheimer, orthopedic, neurological or sensory deficits [19], [17], [9]. Besides, in the case of older citizens, these criteria have been found to be significant predictors of admission to a nursing home, use of paid home care, use of hospital services, living arrangements, use of physician services, insurance coverage, and mortality [23]. According to a government profile of older Americans [1], on average 40% of the non-institutionalized and 92% of the institutionalized older citizens have difficulty in performing one or more ADLs and this percentage becomes higher as they grow older. Figure 11 (Appendix) shows example ADLs and statistics of institutionalized Americans having difficulties performing different ADLs. Therefore, detecting and recognizing ADLs are important for detecting early symptoms of disease, the improvement of access to prescribed medication, providing exact medical history to physicians, and as an important preliminary step in systems for assisting ADLs.

ADL detection systems are commonly designed for single-user residences. Although 28% of older American citizens live in single-resident homes, most of the older citizens (57%) live with their spouse [1]. Therefore, most homes require ADL systems that can perform accurately in presence of more than one individual. However, the presence of multiple people creates additional complexities. The amount of overlapping and parallel activities increase [14] as number of people in a home increases; it makes detecting activities from raw sensors more difficult, since a sensor can be triggered by multiple activities. Again, difficulty in recognizing activities arises because different persons perform an activity in different ways [4]. The cost of scaling an existing ADL recognition system might grow exponentially if each individual has to be dealt separately. Identifying people without using privacy invasive device is also extremely challenging.

One way of tackling the challenges of multiple people scenario is to use RFID technology and wearable sensors for activity detection and consider each person separately [5], [16]. However, the expectation of elderly people or patients carrying additional devices while performing all the activities is often unreasonable. Moreover, this approach makes user uncomfortable and does not work if the user forgets to wear/use the device. This approach cannot detect visitors and requires additional equipment each time a new user enters the system. Another way of handling multiple users is using a camera [18] for both ADL recognition and user identification. However, in this approach the main problems are limited coverage area, obstacles, complexity of recognition due to users angular variation while performing activity, higher cost, requirement of huge amount of data processing, and the violation of privacy - since most ADLs are private. In addition, the works in both vision and wearable sensors are mainly focused on physical activities or gestures from which the ADLs are inferred. Thus, activities that have similar physical movements require more nuance analysis and more computational resources.

In this paper, we present SARRIMA, a system that recognizes ADLs from passive wireless sensors installed in multi-resident homes. It uses semi-supervised algorithms for detecting ADLs in order to minimize the trade-off between data labeling and training time [13]. The system uses location and temporal information to detect ADLs and achieves average accuracy as high as 97% in all the tested datasets. The user is identified by considering differences in performing activities, co-relating activities of the same user, and using information from non-wearable sensors. The performance results for each of these aspects of the algorithm are shown. We also evaluate the person assignment solution as a function of the numbers and types of passive sensors. These results show not only the difficulty of performing person activity assignment with a small set of sensors, but also that one can achieve a high accuracy by appropriately adding other passive sensor modalities. We evaluate the performance of SARRIMA using different datasets [7], [3] with real homes having multiple residents. Although the basic machine learning algorithms used in SARRIMA are not novel, we have preprocessed the initial data and included additional temporal features to make SARRIMA applicable to multi-resident homes. The person identification module is also different than the state-of the art solutions. The ADL recognition module and person identification module exchange information to identify a person and to verify the choice of the recognized ADL.

One of the major limitations of the work is that it is tested only in two-residence homes due to the lack of available data. However, we believe that the system will work in homes with more residence, although the accuracy might be lower. Nonetheless, there are not many homes where there is more than 2 residents and where ADL recognition is required.

The rest of the paper is organized as follows. Section II provides a high-level system description of SARRIMA. Section III, IV, and V discuss the details of the major system modules. The evaluation is shown in Section VII, followed by discussion in Section VIII, related work in Section IX, and finally Section X concludes the paper.

II. System Description

SARRIMA operates based on the assumption that *particular activities of daily living are usually performed in some specific room area and generally occur at the same time of day*. Therefore, if sensors are positioned around the place where a particular activity takes place, then those specific set of sensors will be triggered whenever the activity is performed. This approach has already been applied on single-person residences [13], but the presence of multiple people introduces randomness and therefore complexity in defining the Activity Classes. Moreover, additional problems occur in recognizing activities that can take place in any room; examples include talking on cell phones, having conversation, using Internet, and cleaning the house. However, since the assumption works for most ADLs - we have excluded the special cases from the system scope. Figure 1 shows the overall architecture of SARRIMA.

Sensing Layer: This layer consists of all the sensors placed for activity recognition. The type and number of sensors may vary from house to house. However, all the sensors

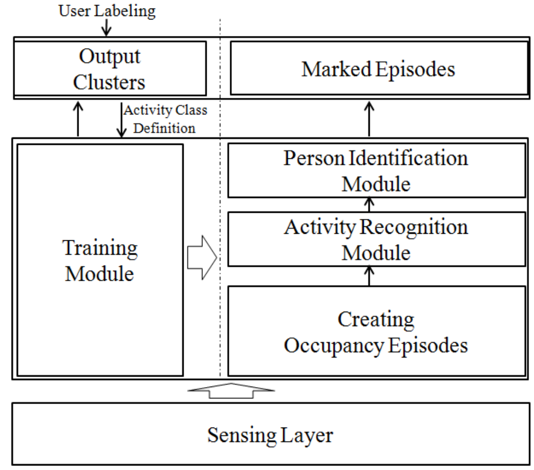


Fig. 1: Overall system architecture of SARRIMA

are non-intrusive wireless sensors. Examples include (but not limited to) contact sensors, item sensors, motion sensors, binary pressure pads, and temperature sensors.

Training Module: The training module processes the data and creates clusters corresponding to different activity classes. An activity class is defined as $A_i = (usedSensors_i, meanStartTime_i, meanDuration_i)$. Here, $usedSensors_i$ is the set of sensors that fire when any instance of A_i is performed. The features $meanStartTime_i$ and $meanDuration_i$ are used to differentiate activity classes that use the same sensor set. The clusters are labeled by users.

Activity Recognition Module: SARRIMA divides the time period of each room into small time durations called occupancy episodes. The occupancy episodes represent the time duration when someone is present in the room. Then, the activity recognition module determines what activities are performed in each of the occupancy episodes, based on the definition of activity class.

Person Identification Module: SARRIMA identifies the user of an activity instance in following ways:

- Personalized Activity Classes: An activity class can be associated to a particular user due to -
  Behavior differences: If different users trigger different sensor sets for a particular activity class, or have different start times or durations performing the activity, then SARRIMA produces separate clusters where each of them is associated with the corresponding user.
  History: History includes the statistics of performing a particular Activity by different users.
- Specialized Sensors: Microphone or non-binary pressure pads can identify users based on voice or weight difference. These type of sensors enforce a user difference by using identifiable human traits. The placement of these sensors are very important in homes where definitions of most or all activity classes are generalized. SARRIMA uses the data collected from these sensors for identification purposes.

Fig. 2: Training Framework for Defining Activity Classes

- Linking Occupancy Episodes: Activity instances are linked based on rooms, time, and activity class properties where all the linked activities are performed by the same user. However, this step is dependent on the previous steps, since at least one activity instance on a particular ADL link has to be marked in order to apply it to recognizing the other linked activities.

Output: The output of SARRIMA gives a list of marked occupancy episodes in the form ($roomId$, $\langle startTime$, $duration\rangle$) $\rightarrow$ List { $\langle activityID$, $PersonID\rangle$ }.

III. Training Module

In this section, we discuss the training framework of SARRIMA (Figure 2). We describe the steps of the framework and briefly elaborate how the randomness created by multiple people is tackled by the framework.

Input: SARRIMA takes raw sensor data as input. The raw data basically contains the status of each sensor at different time intervals throughout the data collection period. The system also requires the information of room ID associated with each sensor.

Pre-Processing: The input data is processed to create a sequence of pairs of the form (s_i, t_{1i}, t_{2i}, v_i, r_n) where sensor s_i is deployed in room r_n, and s_i is on from timestamp t_{1i} to timestamp t_{2i} with value v_i.

Room-wise Separation: In this step, SARRIMA separates the sensor events of different rooms into different files. Therefore, each file contains a collection of sensor events sequentially listed based on the starting timestamp (t_{1i}) of the corresponding room.

Occupancy Episodes: Occupancy episode of a particular room is defined as the time duration when the room is occupied by someone. It is represented in the form ($roomId$, $startTime$, $duration$, $usedSensors$) where $usedSensors$ is the set of sensors that fired during the episode. Now, the presence of a person would cause sensor firings, and logically no sensor should fire when the person leaves the room and no activity is being performed. In a single-user home, there exists only one occupancy segment at any time and, therefore, the occupancy duration is unambiguous. On the other hand, in a multiple resident home, concurrent sensor firings from different rooms make the actual leaving and entering moment unrecognizable. However, SARRIMA assumes that a user is unlikely to leave while performing an activity and if no sensor fires for a certain amount of time, then the room is empty. This time limit is defined as $timeThreshold$. If a sensor fires after that time, then a new occupancy episode starts. Therefore, the time difference between two consecutive sensor firings in an occupancy episode is always less than or equal to the $timeThreshold$ of that room.

Frequency Item Sets: The main purpose of the training framework is to find what set of sensors are associated with each activity class. In this step, SARRIMA determines the unique combination of sensor sets (FI_k) that occur in the occupancy episodes of a particular room. It calculates the number of times a particular set of sensor is present in all the occupancy segments, i.e., how many times $FI_k \subseteq usedSensors_i$ for all i. If $\frac{count\ of\ FI_k}{total\ number\ of\ days} \geq ADL_{support\_threshold}$, then it assumes FI_k to be associated with some particular ADL. $ADL_{support\_threshold}$ defines the minimum frequency for an activity class to be considered as the ADL.

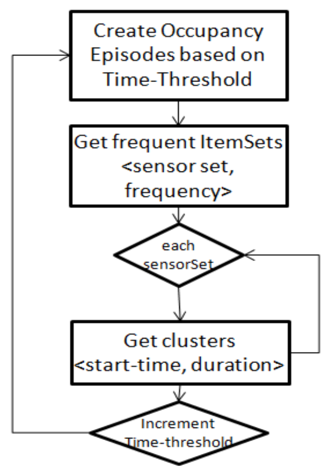


Fig. 3: Finding the sensor clusters of a particular room (getActivityClusters)

Clustering: Clustering differentiates activity classes that use the same set of sensors. For example, *'preparing breakfast'* and *'preparing dinner'* happens at different times of the day, but are likely to use the same set of sensors. Therefore, in this step (Figure 3), SARRIMA clusters the instances of each frequent itemset based on their temporal characteristics i.e., start times and durations. This step also helps to define personalized activity classes when the temporal features are user specific. SARRIMA applies the density based clustering algorithm DBSCAN [10] which does not need to specify the number clusters in advance. This is important since the numbers of different activity classes in each room are not predefined. Now, for each frequent itemset FI_i of a room, SARRIMA runs DBSCAN separately on the set of tuples ($startTime_{ik}$, $duration_{ik}$); here k = (1, 2, ... , $number\ of\ occupancy\ episodes$)

where FI_i occurs. Each attribute of each tuple is normalized before clustering. In DBSCAN, the number of clusters depends on the threshold parameter. SARRIMA calculates and uses the lowest threshold parameter that gives maximum number of clusters but minimum number of unrecognized instances for each FI_i. Each of the clusters signifies a particular activity class.

Output and Labeling: The overall framework outputs a number of clusters for each room. Each cluster C_i is represented by the tuple ($usedSensors_i$, $meanStartTime_i$, $meanDuration_i$, $neighborhoodRadius_i$, $label_i$, $preferredTimeThreshold_i$, $personID_{optional}$). The labeling is done by the user based on the clustering properties. The value of mean start time, mean duration, and number of occurrences of activity instances depends on what $timeThreshold$ is used to create the occupancy episodes. Therefore, the user also chooses $preferredTimeThreshold$ based on the other parameters. So, after this step SARRIMA defines each activity class by associating it with one or more clusters. $personID_{optional}$ indicates the user (if any) associated with the cluster. If the cluster is general for all users, then $personID_{optional}$ is null.

IV. Recognizing Activities of Daily Living

After training, each Activity class is defined in SARRIMA in terms of used sensor sets and temporal parameters. SARRIMA uses the following steps to recognize subsequent activity instances:

- First, it defines the $roomTimeThreshold_r$ =min$\{preferredTimeThreshold_i\}$ where i=(1,2, ... $number\ of\ clusters$) associated with room r. For each room r, the system creates the occupancy episodes using the value of $roomTimeThreshold_r$ as time threshold.
- Then, for each occupancy episode SARRIMA takes the set of used sensors u_e and finds the set of clusters $\{C_j\}$ such that $usedSensors_j \subseteq u_e$ (j $\in$ { 1, 2, 3, ... , $numberofClusters$}). There can be multiple such clusters. This happens because the user may do multiple activities in the same occupancy episode, or multiple user doing different activities can be present in the same occupancy episode. Moreover, even with the same set of used sensors, there may be multiple activities based on the temporal characteristics. This module marks all the matched activity classes as the possible activities that occurred during the episode.
- Finally, if more than one activity is recognized for the same occupancy episode, then the system compares the start time of the episode to the $meanStartTime$ of each of the probable activity classes and assigns it to the closest match. If still unassigned, then the system uses the $meanDuration$ feature.

Note that the duration feature has given less importance since the duration of occupancy episodes varies based on the value of $timeThreshold$ used for calculating it.

V. Person Identification

In home-health care systems, identifying the user of each activity instance is necessary in order to collect the accurate behaviors of the patient. It can be done in following ways:

- Sophisticated sensors can be used to identify and track each person. However, the issue of privacy invasion comes up with this approach. Moreover, the research of identifying people from sensors is very broad and out of scope of this paper.
- The user can label the performer of each of the identified activity instances. This method is really cumbersome and contradicts the sole purpose of recognizing activities automatically.

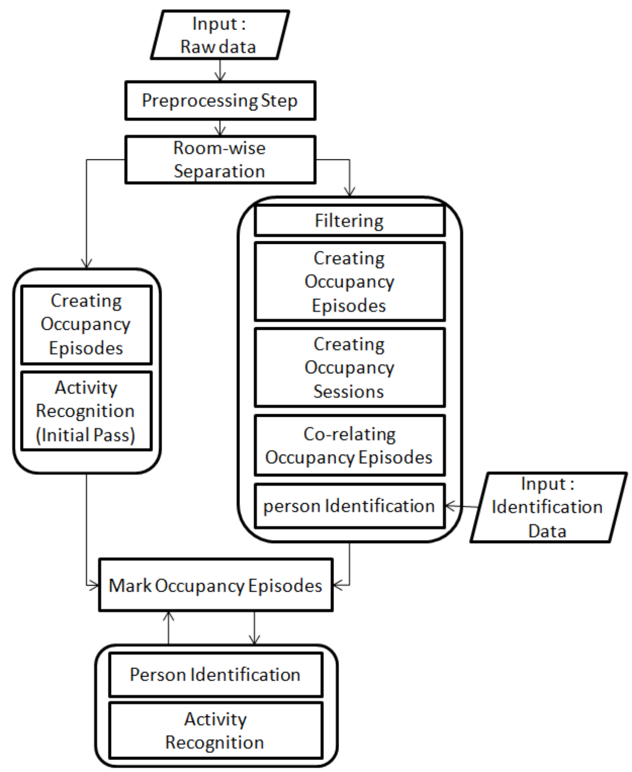


Fig. 4: Person Identification

Therefore, SARRIMA identifies user first by identifying the difference in performing activity and then uses the information from identifying sensors. In this paper, we chose pressure pads and microphones as identifying sensor because they can be installed passively and are less privacy invasive. Since the sensors are placed only in places where the user feels comfortable and the users are given the option to turn-off microphone at any time, SARRIMA only gets partial personal identification data in this process. The system uses the partial information to iteratively determine who was present in the corresponding occupancy episodes and repeats until all occupancy episodes are labeled or no further improvement can be made. The complete process is shown in figure 4. The first three steps are similar to the steps of the Activity training Framework (Section III), and therefore the description is omitted here.

A. Phase: Activity Recognition

Activity instances are recognized based on the steps described in Section IV.

B. Phase: Person Identification

Filtering: Some sensors (Temperature sensors and Infrared sensors) fire even when nobody is present in the room. These sensor values are important in determining activity instances, but misleading in determining the presence of a user. Therefore, SARRIMA ignores the data of these sensors in person identification step (but not in Activity recognition step). As a result, the remaining data ensures that sensor firing occurs only when someone is present in the room.

Creating Occupancy Episodes: SARRIMA creates occupancy segments based on the modified data. The minimum value of $roomTimeThreshold_{roomID}$ for $roomID$ = 1,2, ... N_r is used as the $timeThreshold$ for all the rooms.

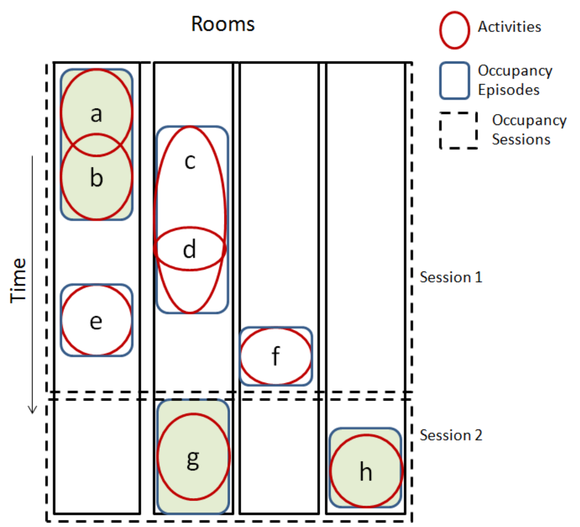


Fig. 5: creating Occupancy Sessions

Creating Occupancy Sessions: An occupancy session is a collection of consecutive and overlapping occupancy episodes from all the rooms of the house. The main idea behind creating occupancy session is to separate the episodes of different users and thereby correlate the occupancy episodes of a particular user within a smaller time frame (occupancy session). The start time of an Occupancy Session is the start time of the earliest occupancy episode and the end time is the end time of the occupancy episode that finishes after all episodes in that session has finished. Now, as long as a new occupancy episode overlaps with one or more episodes occurring in a different room, it can be separated from those and therefore included in the same occupancy session. Whenever SARRIMA finds an occupancy episode that starts after all the previous episodes of the session have finished, it increments the current session number and adds the occupancy episode in a new session. The process continues until all the occupancy episodes are placed under some occupancy session.

Co-relating Occupancy Episodes The occupancy episodes of a particular session can be separated or co-related based on their spatial and timing parameters. The rules for co-relating occupancy episodes of a particular session is described below. The description assumes a two-person setting for simplicity.

- *Occupancy episodes in different room:* The occupancy episodes of a particular person should never overlap, since he/she cannot be in multiple rooms at the same time. Therefore, if the system detects two occupancy episodes occurring at the same time in two different rooms, then it places the two occupancy episodes in seperate groups. For example, in figure 5, occupancy episode with activity b is placed in a different group than activity c, c in a different group than e, and e in a different group than f. Therefore, in a two user setting, the occupancy episodes with a,b, and e are co-related and the activity instances performed by the same user, whereas the occupancy episodes with c and f are co-related and performed by the remaining user.
- *Occupancy episodes in same room:* All activities detected in a particular occupancy episode is assumed to be done by the user present in that episode. However, if the system detects only a single occupancy episode at a particular time, then most likely all the residents are in the same room. In that case, it is assumed that the activity can be done by any user and thus no assignments are made. For example, the user of activity d in Figure 5 cannot be determined with certainty.

The set of co-relating occupancy episodes in the same occupancy session is defined as an activity group. All the occupancy episodes of a particular session can be determined by detecting the user of any occupancy episode within that session.

Person Identification As mentioned earlier, SARRIMA uses two sensor types for the identification purpose.

- *Pressure Pad:* Non-binary pressure pads or weight sensors reports the weight of the the user when someone sits on it and therefore provides information to differentiate between two persons based on their weight difference.
- *Microphone:* SARRIMA uses state-of-the-art speaker identification tool (RESONATE [8]) for identifying speaker from voice data.

Mark Occupancy Episodes The overall goal of the system is to find the activity instances performed in each occupancy episode and associate the instances with some user.

- After completing the previous steps. SARRIMA gets a list of occupancy episodes, marked with suggested activity instances. Some of the instances are marked with associated user.
- Now, in a Occupancy session, if SARRIMA finds at least one occupancy episode associated with a particular user, theoretically all other episodes in that session can be marked with the correct user. However, if the initial information is wrong than all the episodes will be incorrectly marked. Therefore, SARRIMA marks an entire session

if it has several marked occupancy episodes that do not contradict each other. Otherwise, no assignment is done for the unmarked episodes of that session.

C. Iterative Phases of Activity Recognition and Person Identification

Merging Occupancy Episodes: SARRIMA merges consecutive occupancy episodes of the same user if all the episodes suggest the same activity. In this way, the actual duration of the performed activity instance can be determined instead of representing it as several segmented instances. This step is important in determining and verifying the instances of activity classes that are differentiated by activity duration.

Merging Occupancy Sessions:

Similar activities: For certain activity classes, if the existence of activity instances from the same class is detected within a certain time in two different groups, then the activities are considered to be performed by two different users. For example, A and B are groups in $occupancy\ session$ 1 and C and D are groups in $occupancy\ session$ 2. If the system detects brushing teeth in both groups A and D, then SARRIMA merges the two sessions where groups A and C are merged into one larger group and groups B and D are merged into another larger group.

Complementing activities: Some activities instances frequently occur together. For example, people go to toilet after waking up in the morning or before going to bed. Similarly, using a wardrobe can be detected before taking a shower. If these complementing activity instances are detected within certain time periods, then SARRIMA considers these activities to be performed by the same user.

VI. Parallel and Interleaved Activities

Parallel and overlapping activities: If two or more activity instances occur at the same time, then they are called parallel activities. In a multi-user scenario, it is very common and the number of such instances increases as the number of user increases. SARRIMA detects parallel activities by detecting all possible activity instances in a particular time frame (occupancy episode). It assigns all activities to a user if he/she is the only one present. However, if multiple users are detected in the same episode, the system refrains from making any user assignment (unless it matches personalized activity class definition) and considers all the detected persons as a probable performer of the activity

Interleaved activity: A person can leave in the middle of activity and later come back and finish the activity. These types of activity instances are called interleaved activity. Our system can recognize interleaved activities. When the system recognizes same activity type performed by the same person in consecutive occupancy episodes, it checks the time difference between the occupancy episodes and merges the episodes if the time difference is less than some threshold value. This threshold value depends on activity type and is determined empirically based on statistical probability.

VII. Evaluation

There are two publicly available datasets for recognizing ADLs from passive wireless sensors which have homes with more than one person [3], [7]. We tested SARRIMA in these datasets to detect activities and then to identify the person who performed that activity. Each of the datasets has data from multiple homes and labeled ground truth for different sets of activities. The homes have different floor-plans, different types of sensors, and different demographics (Appendix figure 12). The algorithms of SARRIMA are implemented on MATLAB.

A. Datasets

CASAS

The CASAS Smart Home project [7] has collection of data from several houses where there are more than one resident. However, not all of the houses have annotated data. Moreover, in some cases the project aim is different than ours and therefore annotated data could not be used to evaluate SARRIMA. For example, data was collected to differentiate among individual activities, group activities, and co-operative activities and thus the identification information of the user was not necessary. In this paper, we included the results of performance of SARRIMA in the *"CASAS Spring 2009 multiperson dataset"* and the *CASAS Summer 2010 multiperson dataset"*. In these datasets, data was collected from a two-story apartment that housed two residents and they performed their normal daily activities. The datasets annotates several ADLs including sleeping, personal hygiene, preparing meal, work, study, and watching TV. Seventy-two sensors were deployed in the house which includes motion sensors, item sensors, door/contact sensors, and temperature sensors.

ARAS

The ARAS dataset [3] has a total of two months of data; collected from two real homes - House A and House B . The residents of one house were husband and wife, whereas the residents of other house were two graduate students. The floor plans of these two houses were different and had different sensor settings. However, both the houses had a total 20 sensors. The sensor types included contact sensors, force sensors, photocells, pressure mats, distance sensors, sonar distance, IR, and temparature sensors. All the activities were annotated by the users.

B. Results: Activity Detection

SARRIMA assigns one or more activity instances to each occupancy episode based on the triggered sensor set and the temporal features of that episode. To minimize the number of false positives, occupancy episodes with very small durations (length$\leq$10s) are filtered before making the assignment.

Figure 6 shows the percentage of accuracy of recognizing instances of different ADL classes in four different houses. The missing columns in figure 6 indicates that no ground truth labeling is found for the corresponding activity class in that particular dataset. We see that SARRIMA achieved 97.34% and 98.15% accuracy on average respectively on CASAS Spring and Summer dataset. The accuracy is higher

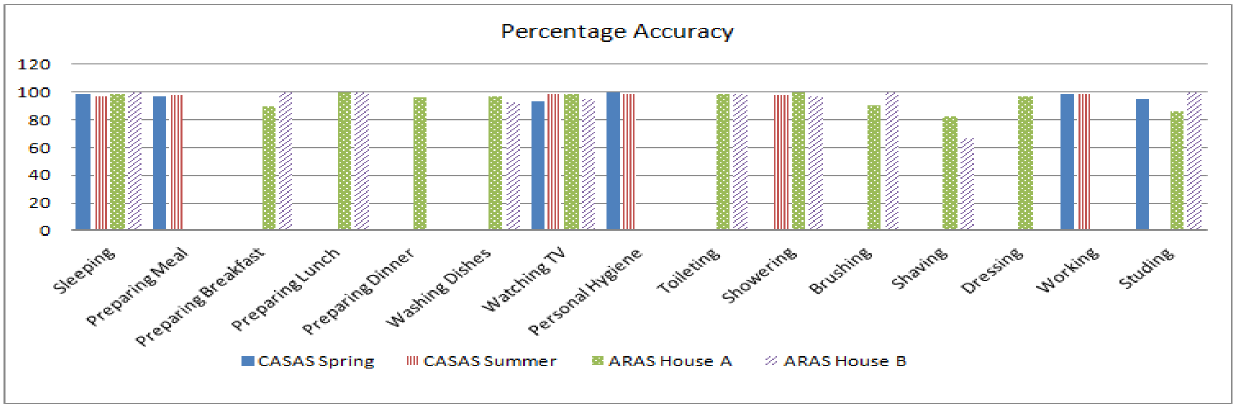


Fig. 6: Percentage of activity instances of Activity Classes recognized correctly in CASAS Spring 2009, CASAS Summer 2010, ARAS House A, and ARAS House B (time_threshold = 2 minutes)

than HMM and SemiHMM (92% on average) applied on the same dataset [7]. In ARAS houses, the activity classes have more fine grained definition. For example, ‘preparing breakfast‘, ‘preparing lunch‘, and ‘preparing dinner‘ instead of ‘preparing meal‘ or ‘toileting‘, ‘brushing‘, and ‘shaving‘ instead of just ‘personal hygiene‘. Therefore, relatively lower accuracy is achieved when considering these activity class definitions (average accuracy of 87% for House A and 95.3% for House B). For example, activity instance of ‘brushing teeth‘ and ‘shaving‘ have lower accuracy since both of them uses similar item sets and generally occur around the same time of day. Again, in House A, the ‘Preparing lunch‘ time of Person A often overlaps the ‘preparing dinner‘ time of person B and results lower accuracy. However, the average accuracy of activity detection is more than 98% in both houses when general ADL class definition (similar to CASAS houses) is considered.

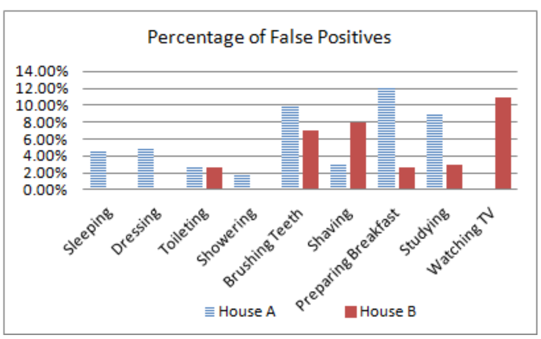


Fig. 7: False positive of reported activity instances ARAS (House B) (time_threshold = 2 minutes)

CASAS had less than 2% false positives for any activity instances. Whereas, higher number of false positives was observed in the ARAS dataset due to more fine grained definition of each of the activities (figure 7). We notice that more false positives occurs in House A for all activity classes except for ’Watching TV’ for which the system reports zero false positives. The reason behind is that in House A ‘Watching TV‘ is detected from Infrared sensor but in House B ‘Watching TV‘ is detected from binary pressure pads in living room. Therefore, in House B ‘Watching TV‘ is often reported when the actual activity is ‘Study‘ or ‘Using internet‘. The number of false positive in bathroom and kitchen reduces when higher $timeThreshold$ is considered for creating occupancy episodes. However, in that case occupancy episodes have longer duration and therefore longer activity duration is reported for some activity instances.

C. Results: Person Identification

CASAS annotates the activity instances of sleep, work, and bed_to_toilet for each user. However, the rest of the Activity Class instances do not have user specific ground truth. The two users in CASAS always performed the sleep, work, and bed_to_toilet in separate rooms. Therefore, SARRIMA recognizes the user of annotated user specific instances just by checking room location. Consequently, we achieve a 100% accuracy for this simple home living situation.

ARAS has more complex activity class definitions and all the activities of each user is labeled. Therefore, the following evaluation will focus only on ARAS datasets.

User Assignment from Personalized Activity Classes

SARRIMA uses user behavioral difference to define a personalized activity class. For example, in House A the two people sleep in two different room where as in HouseB Person1 always sleeps on the right side of the bed. In House B, ‘preparing meal‘ is always done by person 1; whereas in House A the residents had different sleeping patterns and therefore different times for using the toilet in the morning and at night.

Figure 8 shows the percentage of activity instances assigned correctly to a user based on only one behavior difference. We can see that in some cases, the recognition is almost 100% correct. For example, this is true for sleeping and dressing. However, detection of ’Watching TV’ in House B shows low accuracy due to lower accuracy in detecting the activity instances of that class. On the other hand, in House A detection of ’Watching TV’ for person 1 gives low accuracy because that person does not always view TV from the same location or position. Here, we want to note that the user behavior will not be same for every resident. For some resident there might not be any identifiable difference between two users performing the same activity. However, if differences exist in term of the sensor activated or the time of the activity, then the system recognizes the differences. The labeling of the classes helps to identify the differences of user behavior. The labeling is done only once after the framework generates the cluster. Therefore, not much extra user effort is required for this step.

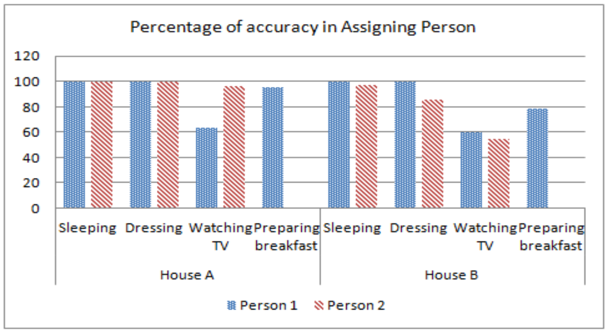


Fig. 8: Identifying accuracy of User from Personalized Activity Classes for selected Activity classes in ARAS Houses

Linking Occupancy Episodes Behavioral differences do not differentiate users for all activity classes. Therefore, SARRIMA also correlates occupancy episodes based on the overlapping episodes of different rooms. Table II shows the episode information of a randomly chosen day from Aras (House B). For a few problematic activities the first column shows the percentage of occupancy episodes correctly identified by SARRIMA based on behavior alone. The second column shows the percentage of episodes where a user was correctly identified by using the SARRIMA feature of linking the episodes with already identified ones. This shows the value of this feature of SARRIMA.

TABLE I: Effect of user identification using behavioral information and episode linking (threshold = 3 min) for a single day in ARAS House B

	User identified correctly in percent of episodes (Behavioral)	(Correlating)
Bed	100%	-
Bath	0	47.3%
Living	37.5%	31.8%
Kitchen	0	42.85%
Hall	0	40%

Specialized Sensors

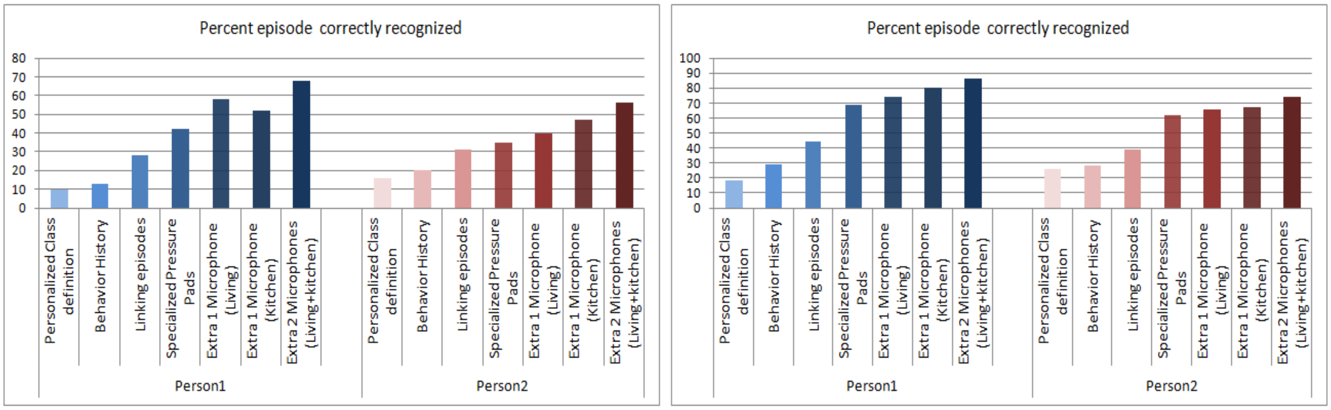
In this paper, we use the term specialized sensor for the sensors which are capable of differentiating among multiple persons; examples include weight sensors, height sensors, and microphones. These types of sensors provide the system with intermittent information about a users identity that can be used for further improvement to assignment of activities. Use of specialized sensors is more significant in the homes where activity instances of the same activity classes cannot be differentiated based on behavioral differences or the history of the user. For example, if a husband and wife do not sleep on their own side of the bed, then weight monitoring pressure pads on the bed will help differentiating who is who.

Weight sensor: The ARAS dataset has binary pressure pads placed in the bedroom, living room, and dining, and annotates the activity of each user. Therefore, we have simulated the data using actual data and user annotation information, where the pressure pad shows actual user weight instead of a binary value. The pressure pads in the bedroom did not provide additional information since bedroom activities are already included in the definition of the personalized Activity class. However, the pressure pads on the dining room chair identifies the user who is eating a meal and the pressure pads in the living room identify the user of activities ‘Watching TV‘, and ‘study‘.

Microphones: A microphone is a passive sensor that can be placed anywhere in the room. It can identify users uniquely by recognizing voice when they are speaking. The voice recognition only uses extracted features and discards the actual conversation. Moreover, a user can turn off the device at any time he/she wants. However, the problem with microphone is that there is no certainty whether the user will talk or not. The accuracy varies on the length of speech sample and distance of user is from the microphone. Also, human voice from television or radio creates confusion. SARRIMA does not have algorithms for speaker identification, but uses external tool for this purpose. For example, RESONATE [8] is a speaker identification tool that uses SVM classifier for identifying a person from a single microphone deployed in a room and achieves accuracy higher than 86% in real home environment with 5 speakers when utterance length is 2 secs. The accuracy is much higher (almost 100%) when number of speaker decreases and the utterance length is longer. Therefore, we simulated voice data in occupancy episodes with long duration with the assumption that the longer an user stays in a room, it is more likely he/she will talk.

Figure 9 shows the percentage of activities correctly assigned to each person of House A and House B. The accuracy is not high enough if only binary sensors are considered, however it increases significantly just by placing a few extra passive sensors that are able to differentiate person from one of their traits. In the case of the microphone, SARRIMA considers the episodes that have duration higher than five minutes assuming that a person would talk if he/she stays longer than this in a room. We can observe from the figure that the effect of adding specialized sensors on accuracy varies from person to person and that it often helps increase the accuracy to near 100%.

The ARAS dataset have been used in other research work for ADL and person identification. In the paper [3], the researchers used a hidden Markov model (HMM) and obtained an average accuracy of 61.5% for House A (with min=46.3% and max=88.4%) and 76.2% (with min=31.1% and max=96.7%) for House B. In paper [2], the researches applied an incremental decision tree model and obtained



(a) (b)

Fig. 9: Person Identification Accurracy (threshold=2 min) (a) House A (b) House B

classification accuracy 40% for House A and 82% for House B. The low accuracy in these previous works demonstrates the significant difficulties in person assignment. SARRIMA performs significantly better and we also show that with adding specialized passive sensors we can achieve very high accuracy for many activities.

VIII. Discussions and Future Work

We want to share some observations -

Result variation among houses: Although the activity detection module worked very well in all the houses; the accuracy of person identification varied. The main reason can be identified as the effect of personalized activity classes. In the CASAS dataset there was not much sesnsors that can differentiate each user. Moreover, in ARAS House A linking episodes based on overlapping was more effective than in House B. The reason was because in House B the users were in the same room more often than in House A.

False positive: Figure 7 shows high false positive for some activities. These false positive occurs where the sensors are not definitive for particular activities. for example, motion sensors that trigger both when brushing teeth and toileting, or when people just pass by. We can reduce the number of false positive by filtering episodes with short duration; however it causes the accuracy to decrease as well.

This work can be extended in different important direction:

- *Access the Quality of ADLs for long term monitoring:* If the time for an activity is gradually increasing, it indicates that the person is having more difficulty in performing the activity. Similarly, the number of time an activity is performed over day/week can be changed due to medical condition. For example, going toilet more, skipping meals often, and watching TV all day. Although, quality monitoring is closely related to anomaly detection; it is different in cases where the anomaly detection system considers a consistent change as a new pattern for regularity.
- *Sensor importance in activity detection vs. person identification:* From this paper we can see that the sensors deployed for detecting activities (pressure pads) can be helpful in identifying the person as well; however it does not hold true for all type of sensors. It would be interesting to find out by deploying large amount of different type of sensors and analyze which sensors are effective for activity detection, which ones for identification, and which ones work for both.
- *Effect of emotion (anger, happiness, sadness) on ADL quality:* There are systems that monitor the emotional state of patients having sensitive diseases. It would be interesting to find how emotional state affects the quality of activity and whether performing activities certain ways can be considered as actions indicating emotional states.

IX. Related Work

There are different commercial systems available for activity monitoring in home, such as the Quiet Care System by Intel and the e-Neighbor by Healthsense. These systems are used only to detect falls of elderly people. Other interesting products are mobile apps, such as Google activity recognition APP. However, it only provides limited information about the user‘s activity, such as whether the user is on foot, in a car, on a bicycle, or still. Many researchers are using mobile phones as a sensing platform for activity recognition. But usually they define activity as the physical activities like walking, running, or being still, and are mostly used to track an individual‘s exercise (e.g. UbiFit [6]) or for fall detection.

A closely related field of activity recognition is gesture recognition. There are research which attempt to recognize activity by combining gestures [15], [11]. Here, the activities are mainly simple physical activities like sitting, waving etc. A common approach in recognizing a person‘s activity is through wearable sensors [21], [20]. But as mentioned before, they are not comfortable; users often forget to wear them, and scaling the system to multiple persons is energy and cost consuming. Another approach is using a single Infrastructure based sensor, such as ElectroSense or HydroSense [12]. The advantages of these sensors are that they are single-point, which reduce both the installation cost and overall system cost. Though, only a limited set of activities can be detected by the sensor, which consumes the resource measured by that particular sensor. A lot of work in Activity Recognition is based on image processing and analyzing data from camera [18]. However, those systems are not preferable to most users in home environments due to their nature of privacy invasiveness.

The existing research of recognizing activities of daily living with ubiquitous sensors uses different statistical and probabilistic approaches [22], [24]. The common algorithms are Hidden Markov Model (HMM), semi HMM, nave Bayes classifier, and conditional random field. Most of these systems do not consider the presence of multiple person and do not provide any comparisons. In the paper STAR [24], the authors propose that activity recognition and tracking people can be done simultaneously. They performed preliminary experiment with limited tracking granularity up to room-level where the accuracy of tracking decreases to less than 6% when numbers of people are increased to four. They considered activity only as the movement of a person and did not define or relate to specific sets of activities. The paper AALO [13], describes an activity recognition platform which uses active learning technique for activity recognition of single person home. They also consider the existence of overlapped and interleaved activities. Our current work is based on the framework of AALO. However, we have modified the framework so that it can be applied in multi-resident home. In addition, we have provided methods for differentiating one user activities from the other.

X. Conclusion

Activity recognition plays a vital role in home Health-Care systems for monitoring elderly people. ADLs are very important in evaluating health conditions and in prescribing correct medication. In this paper, we presented SARRIMA, a system for detecting activities of daily livings in the presence of multiple people. For the given datasets, SARRIMA is capable of detecting 97% of the activities on average which is higher than a HMM (92%), and SARRIMA also reports parallel and interleaved activities. Importantly, SARRIMA identifies the user of the activity without using wearable sensors or RFID tags. The paper shows how SARRIMA is to achieve very high accuracy for person identification by using a history of personal behavior, linking occupancy intervals across rooms, and by including appropriate sensors in good locations. It also discusses the lesson learned from using data from different real homes.

References

[1] A profile of older american : 2013. Technical report.

[2] H. Alerndar, H. Ertan, O. Incel, and C. Ersoy. Aras human activity datasets in multiple homes with multiple residents. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on*, pages 232–235, May 2013.

[3] H. Alerndar, H. Ertan, O. D. Incel, and C. Ersoy. Aras human activity datasets in multiple homes with multiple residents. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on*, pages 232–235. IEEE, 2013.

[4] T. S. Barger, D. E. Brown, and M. Alwan. Health-status monitoring through analysis of behavioral patterns. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 35(1):22–27, 2005.

[5] M. Buettner, R. Prasad, M. Philipose, and D. Wetherall. Recognizing daily activities with rfid-based sensors. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 51–60. ACM, 2009.

[6] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, et al. Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1797–1806. ACM, 2008.

[7] D. J. Cook, M. Schmitter-Edgecombe, et al. Assessing the quality of activities in a smart environment. *Methods Inf Med*, 48(5):480–485, 2009.

[8] R. Dickerson, E. Hoque, P. Asare, S. Nirjon, and J. Stankovic. Resonate: Reverberation environment simulation for improved classification of speech models. In *Information Processing in Sensor Networks, IPSN-14 Proceedings of the 13th International Symposium on*, pages 107–117, April 2014.

[9] R. F. Dickerson, E. I. Gorlin, and J. A. Stankovic. Empath: a continuous remote emotional health monitoring system for depressive illness. In *Proceedings of the 2nd Conference on Wireless Health*, page 5. ACM, 2011.

[10] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.

[11] A. Fathi, Y. Li, and J. M. Rehg. Learning to recognize daily actions using gaze. In *Computer Vision–ECCV 2012*, pages 314–327. Springer, 2012.

[12] J. E. Froehlich, E. Larson, T. Campbell, C. Haggerty, J. Fogarty, and S. N. Patel. Hydrosense: infrastructure-mediated single-point sensing of whole-home water activity. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 235–244. ACM, 2009.

[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*, pages 139–146. IEEE, 2012.

[14] A. L. Lansky. A representation of parallel activity based on events, structure, and causality. In *Reasoning About Actions & PlansProceedings of the 1986 Workshop*, pages 123–160, 1987.

[15] K. Lyons, H. Brashear, T. Westeyn, J. S. Kim, and T. Starner. Gart: The gesture and activity recognition toolkit. In *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*, pages 718–727. Springer, 2007.

[16] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher. Activity recognition and monitoring using multiple sensors on different body positions. In *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, pages 4–pp. IEEE, 2006.

[17] L. Mercier, T. Audet, R. Hébert, A. Rochette, and M.-F. Dubois. Impact of motor, cognitive, and perceptual disorders on ability to perform activities of daily living after stroke. *Stroke*, 32(11):2602–2608, 2001.

[18] B. Ni, G. Wang, and P. Moulin. Rgbd-hudaact: A color-depth video database for human daily activity recognition. In *Consumer Depth Cameras for Computer Vision*, pages 193–208. Springer, 2013.

[19] K. Rockwood. The measuring, meaning and importance of activities of daily living (adls) as an outcome. *International psychogeriatrics*, 19(03):467–482, 2007.

[20] N. E. Sherwood and R. W. Jeffery. The behavioral determinants of exercise: implications for physical activity interventions. *Annual review of nutrition*, 20(1):21–44, 2000.

[21] B. G. Steele, B. Belza, K. Cain, C. Warms, J. Coppersmith, and J. Howard. Bodies in motion: monitoring daily activity and exercise with motion sensors in people with chronic pulmonary disease. *Journal of rehabilitation research and development*, 40(5; SUPP/2):45–58, 2003.

[22] E. M. Tapia, S. S. Intille, and K. Larson. *Activity recognition in the home using simple and ubiquitous sensors*. Springer, 2004.

[23] J. M. Wiener, R. J. Hanley, R. Clark, and J. F. Van Nostrand. Measuring the activities of daily living: Comparisons across national surveys. *Journal of Gerontology*, 45(6):S229–S237, 1990.

[24] D. H. Wilson and C. Atkeson. Simultaneous tracking and activity recognition (star) using many anonymous, binary sensors. In *Pervasive computing*, pages 62–79. Springer, 2005.

APPENDIX

A. Activities of Daily Livings

Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) are clinical terms used to define the set of activities, performing which reflects a persons ability to live safely and independently. Doctors, rehabilitation specialists, geriatric social workers, and others in senior care often assess ADLs and IADLs as part of an older person‘s functional assessment

Fig. 10: Basic Activities of Daily Livings (ADL)

TABLE II: List of Activities of Daily Livings

Basic ADLs	Instrumental ADLs
Sleeping	Working
Eating	Preparing Meals
Toileting	Washing Dishes
Showering	Cleaning House
Dressing	Watching TV
Brushing	Using Telephone
Moving	Study
	Shopping
	Managing Finances
	Commuting

The following figure shows Percentage of Americans in Residential Care-Facilities having difficulties in performing different ADLs based on Technical report ”A profile of older American : 2013”.

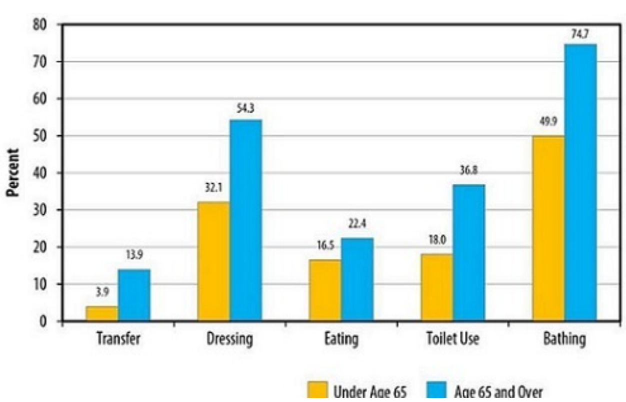
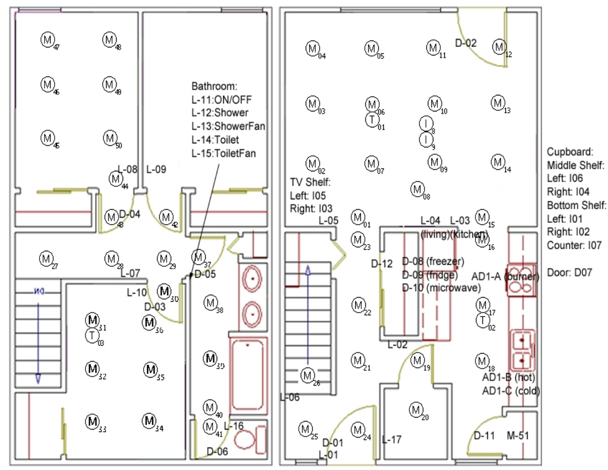


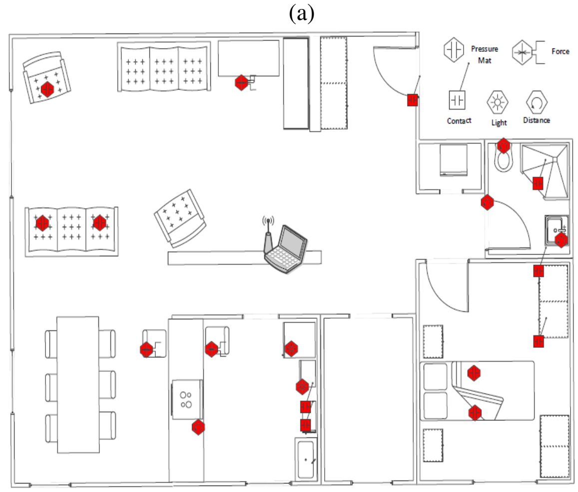
Fig. 11: Percentage of residents in Care-Facilities having problem performing ADLs.

B. Dataset

The CASAS Smart Home project [7] and The ARAS dataset [3] has collection of data from real houses where there are more than one resident. The homes have different floor-plans, different types of sensors, and different demographics. the following figure shows some example layout of floorplan and sensor placement.



(a)



(b)

Fig. 12: Floor plans and sensor layout (a) Casas WSU testbed (b) Aras House B