Published in final edited form as:

J Ambient Intell Smart Environ. 2016 ; 8(4): 437–451. doi:10.3233/AIS-160385.

Clustering Home Activity Distributions for Automatic Detection of Mild Cognitive Impairment in Older Adults¹

Ahmad Akl^{a,*}, Belkacem Chikhaoui^a, Nora Mattek^b, Jeffrey Kaye^b, Daniel Austin^b, and Alex Mihailidis^a

^aInstitute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Canada

^bOregon Centre for Aging and Technology, Oregon Health and Science University, Portland, Oregon, USA

Abstract

The public health implications of growing numbers of older adults at risk for dementia places pressure on identifying dementia at its earliest stages so as to develop proactive management plans. The prodromal dementia phase commonly identified as mild cognitive impairment is an important target for this early detection of impending dementia amenable to treatment. In this paper, we propose a method for home-based automatic detection of mild cognitive impairment in older adults through continuous monitoring via unobtrusive sensing technologies. Our method is composed of two main stages: a training stage and a test stage. For training, room activity distributions are estimated for each subject using a time frame of ω weeks, and then affinity propagation is employed to cluster the activity distributions and to extract exemplars to represent the different emerging clusters. For testing, room activity distributions belonging to a test subject with unknown cognitive status are compared to the extracted exemplars and get assigned the labels of the exemplars that result in the smallest normalized Kullback–Leibler divergence. The labels of the activity distributions are then used to determine the cognitive status of the test subject. Using the sensor and clinical data pertaining to 85 homes with single occupants, we were able to automatically detect mild cognitive impairment in older adults with an $F_{0.5}$ score of 0.856. Also, we were able to detect the non-amnesic sub-type of mild cognitive impairment in older adults with an $F_{0.5}$ score of 0.958.

Keywords

Mild Cognitive Impairment; Generalized Linear Models; Room Activity Distributions; Unobtrusive Sensing Technologies; Clustering

¹This work was supported by P30AG024978 Oregon Roybal Center for Translational Research on Aging, P30AG008017 Oregon Alzheimer's Disease Center, R01AG024059 BRP, ISAAC Intelligent Systems for Assessing Aging Changes, and Intel Corporation BRP.

*Corresponding author. ahmad.akl@utoronto.ca.

1. Introduction

As of 2015, 47.5 million people worldwide are living with dementia [1]. In 2011, 747,000 Canadians and 5.1 million Americans were living with Alzheimer's disease and other dementias [1] [2]. As the "baby boomer" generation ages, these figures are projected to increase dramatically thus posing a serious challenge to the health-care infrastructure. Consequently, detection of the cognitive decline that precedes Alzheimer's disease, or dementia in general, becomes vital so as to appropriately apply current best practices for clinical evaluations and management as well as to be available for rationale administration of anticipated new therapies.

However, this capability of early detection is very challenging with the contemporary detection processes in the form of conventional clinician visits which are typically not sensitive to detecting mild or subtle cognitive or functional decline. This corroborates what has been reported in some studies that in more than 50% of detected cases it was the family members who served as the primary source of recognition and not the family doctors [3]. This delay in detecting cognitive decline can be detrimental especially to older adults with reversible forms of dementia, who form up to 11% of the cognitively impaired population [4]. Early detection of remediable causes of cognitive impairment such as medication complications or nutritional deficiencies facilitates timely intervention, possibly increasing the chances of reversing the condition. Even for subjects with irreversible decline, early recognition still provides them and their families with an opportunity to proactively plan for their future by seeking the appropriate interventions and to avoid being forced into crisis management [5].

Mild Cognitive Impairment (MCI) is a concept that describes the transitional period between normal aging with intact cognition and early Alzheimer's disease or other dementia. Individuals with MCI have measurable changes in their memory and thinking abilities but with minor impact on their ability to carry out activities of daily living. Although not all subjects with MCI progress to develop Alzheimer's or other dementias, the proposed criteria and guidelines for diagnosis of Alzheimer's disease published in 2011 have documented an accelerated rate of progression to dementia and Alzheimer's disease in MCI subjects [6]. Therefore, detecting MCI serves the objective of detecting cognitive decline early enough for subjects to seek the right intervention and treatment. Different subtypes of MCI exist such as amnesic MCI (a-MCI), where the memory of an individual is compromised, and non-amnesic MCI (na-MCI), where other cognitive domains of an individual are impaired [7].

Several studies have demonstrated that changes in motor capabilities precede and may be indicative of cognitive impairment [8], and that changes in walking speed could be good predictors of progression to dementia [9]. Accordingly, and with the advancement in technology and the proliferation of smart systems, a good alternative to the traditional clinical paradigm is to bring assessment into the daily activity of a person in their home environment via unobtrusive sensors and systems. In this paper, we propose an approach for home-based automatic detection of MCI in older adults through continuous monitoring using unobtrusive sensing technologies. Our approach employs a clustering algorithm to

group room activity distributions belonging to different states of cognition into different clusters, each represented by an “exemplar”. These exemplars are then used to classify room activity distributions pertaining to a test subject, and then these classifications are used to determine the cognitive status of the test subject.

The rest of the paper is designed as follows: Section 2 summarizes related work, lists research questions, and presents contributions. Section 3 describes the inclusion criteria for subject recruitment, the data acquired, and the labeling protocol implemented. Section 4 defines the problem and describes the methodology we implemented to automatically detect MCI in older adults. Section 5 presents and discusses results obtained. Section 6 addresses the limitations of the current work and proposes directions for future work. Finally, Section 7 concludes the paper.

2. Related Work

The literature is rich with smart systems and studies designed to monitor the health and well-being of older adults and to support their independence such as the Microsoft’s EasyLiving project [10], the GATOR Tech Smart House [11], the AWARE home at Georgia Tech [12], and the GE QuietCare System [13]. In addition, in one study by the Center for Advanced Studies in Adaptive Systems (CASAS), Dawadi et al. implemented a machine learning approach to discriminate cognitively impaired older adults from their cognitively intact counterparts based on their ability to complete a ‘Day Out Task’ that consisted of a number of interwoven Instrumental Activities of Daily Living (IADL) [14]. However, all results reported by these systems and studies were based on data acquired in a laboratory environment and not in a real world setting. An approach that would be more reflective of the subjects’ actual performance would be to continuously monitor the subjects, perhaps in their own homes where they are more likely to exhibit their true performance.

In another study, Sixsmith performed a trial to test the ability of an intelligent monitoring system, composed of sensors in the home, to identify emergencies by detecting deviations from normal activity patterns [15]. The sensing technologies consisted of passive infra-red sensors, contact sensors, and temperature sensors. Of the total 61 alerts that were recorded in a period of 3 months, 46 alerts were classified as false alerts and only 15 alerts were classified as genuine. In other words, the approach suffered from a high rate of false positives, approximately 75%. In a similar study, Skubic et al. implemented a simple alert algorithm to generate health alerts to clinicians via continuous monitoring of the subjects in their homes using unobtrusive sensing technologies such as passive infrared motion sensors, bed sensors, and temperature sensors [16]. Similarly, the alert algorithm generated 183 false alerts out of a total of 299 alerts in a period of 9 months. This translates into an accuracy of 38.8% and a false positive rate of 61.2%. In addition, both of these studies monitored the general health of the occupants and were not designed to specifically detect cognitive changes.

The ORegon Center for Aging and TEChnology (ORCATECH) recruited more than 300 cognitively healthy older adults and monitored them for an average period of 5 years using unobtrusive sensing technologies, resulting in a large database of sensor data and clinical

data [17]. Recently, Dodge et al. presented trajectories of home-based daily walking speeds and their variability of recruited subjects and compared the resulting trajectories between the cognitively intact subjects and those with non-amnesic mild cognitive impairment (na-MCI) [18] [19]. Participating older adults with na-MCI were characterized by a slowing of walking speed and exhibited the highest and lowest variability in their walking speeds in comparison with their cognitively intact counterparts.

Building on the work by ORCATECH and Dodge et al., we explored the feasibility of autonomously discriminating older adults with MCI from their cognitively intact counterparts using a number of predefined measures associated with their in-home walking speed such as the median weekly walking speed, the coefficient of variation of the weekly walking speed, the coefficient of variation of the morning walking speed, and other measures [20]. Different window sizes were used to extract features from these predefined measures which were then used to train and test two machine learning algorithms, namely support vector machines and random forests. We were able to detect MCI in older adults with an area under the ROC curve and an area under the precision-recall curve of 97% and 93%, respectively, using a time frame of 24 weeks.

We also explored using a number of predefined measures associated with the subjects' home activity. Activity was measured as the total number of sensor firings. We experimented with measures such as the coefficient of variation of total activity, coefficient of variation of morning activity, and coefficient of variation of evening activity. However, we discovered that such measures did not carry sufficient discriminative information to discriminate older adults with MCI from their cognitively healthy counterparts. This motivated us to seek an alternative approach especially after Hayes et al. reported preliminary results on potential variability in the day-to-day pattern of activities of subjects with MCI when compared to those who were cognitively intact [21].

Accordingly, we designed a better approach to detect MCI in older adults based on their home activity. We built statistical models of the subjects' home activity by using inhomogenous Poisson processes to model the presence of subjects within different rooms throughout the day [22] [23]. The resulting generalized linear models (GLMs) of the subjects' home activity provided intuitive statistical analysis, and using the Kullbak–Leibler (KL)-divergence measure, we demonstrated an existing statistical difference between activity models pertaining to subjects when cognitively intact and when experiencing symptoms of MCI. Building on our work on estimating GLMs of the subjects' home activity, we explore the following research questions:

1. Can we use GLMs of home activity, in the form of individual room activity distributions, to automatically detect MCI in older using unobtrusive sensing technologies?
2. What size of time frame results in the highest $F_{0.5}$ score when detecting MCI in older adults?
3. How do home activity distributions of older adults when cognitively intact compare with their activity distributions when exhibiting symptoms of amnesic MCI (a-MCI) or non-amnesic MCI (na-MCI)?

Using the ORCATECH dataset, we were able to answer these questions using sensor and clinical data pertaining to 85 subjects collected over an average period of 3 years via unobtrusive sensing technologies installed in the subjects' homes. This paper makes the following contributions:

1. We propose a clustering-based method to automatically detect MCI in older adults using estimated GLMs of their home activity. Exemplars are created to represent clusters of the room activity distributions belonging to different states of cognition using a combination of k -means and affinity propagation (AP). These exemplars are then used to classify room activity distributions belonging to a test subject. These classifications are then employed to determine the cognitive status of the test subject.
2. We were able to detect MCI in older adults with an $F_{0.5}$ score of 0.856 using a time frame of 20 weeks, which is approximately 5 months. Accordingly, a time frame of 20 weeks is empirically reported as the duration that generates room activity distributions that are most conducive to detecting MCI in older adults.
3. Interestingly, when including subjects exhibiting symptoms of na-MCI only, we were able to detect MCI in older adults with an $F_{0.5}$ score of 0.958. On the other hand, we were only able to detect significant differences in the bedroom activity distributions of older adults when cognitively healthy and when experiencing the other sub-type of MCI, namely a-MCI.

3. Data Acquisition & Labeling

In this analytic study, we focused on 85 ORCATECH homes with single occupants with the requisite sensor and clinical data needed for the subsequent analyses.

3.1. Participants and Data Acquisition

Participants were recruited from the Portland, Oregon, metropolitan area. The eligibility criteria included [17]:

1. aged 70 years or older;
2. living independently in a larger than one-room "studio" apartment;
3. cognitively healthy (Clinical Dementia Rating (CDR) score < 0.5 ; Mini-Mental State Examination (MMSE) score > 24); and,
4. in average health for age (well-controlled chronic diseases and co-morbidities or none at all).

Data were acquired by installing sensing technologies in the homes of the recruited subjects. In order to detect movement and general activity, passive infrared motion sensors were installed in rooms frequently visited by the participating subjects. Visitors and absences from the home were tracked through wireless contact switches placed on the exit doors of the home. All sensor firings were sent wirelessly to a transceiver, time-stamped and

recorded. In addition to the sensing technologies, recruited subjects were requested to complete a weekly online questionnaire, where subjects reported any visitors during the week, days spent away from the home, any change in health or medication, admittance to ER, and a number of other queries. For full details on data acquisition, the reader is referred to [17].

3.2. Labeling of Data

Participants were assessed in-home at baseline, and during annual in-home visits by research personnel who administered a standardized battery of tests, including six neuropsychological tests which are considered to be representative of five cognitive domains. The six neuropsychological test are summarized in Table 1

Diagnosis of MCI was made using the Petersen criteria [7] operationalized as absence of dementia, intact functional activities, normal general cognitive function (MMSE score ≥ 24), and objective impairment on one or more of the neuropsychological tests listed in Table 1. Impairment on neuropsychological testing was defined as a score of 1.5 SD or more below the predicted mean value for subject-specific age, sex, and educational levels, based on normative values generated in the past by the Alzheimer's Disease Centers (ADC) program [30]. We also investigated different sub-types of MCI, namely amnesic (a-MCI), and non-amnesic MCI (na-MCI). Accordingly,

- a-MCI was defined as impairment on the neuropsychological testing in the memory domain only.
- na-MCI was defined as impairment on the neuropsychological testing in any of the four domains - executive function, language, attention, and visuospatial but not in the memory domain.

Since subjects were assessed annually, data labels fell into three categories:

1. cognitively intact: "CIN",
2. transitioning to mild cognitive impairment: "TR", and
3. having mild cognitive impairment: "MCI".

Figure 1 illustrates how these labels were assigned by depicting a subject who was monitored for at least 3 years and was administered three annual assessments besides baseline. The subject was assessed as cognitively intact at baseline, but was assessed as exhibiting symptoms of MCI on the 2nd and 3rd year assessments. Therefore, the data from baseline up to the 1st year assessment were assigned the label "CIN" and the data from the 2nd year assessment onward were assigned the label "MCI". The data between the 1st year and the 2nd year assessments were assigned the label "TR" for transitioning to MCI. This is because the conversion to cognitive impairment is not an instantaneous event but a gradual process. Accordingly, the subject's cognitive status would be in flux between years 1 and 2 and would not belong to either cognitive intactness or MCI. In order to ensure consistency in labeling subjects' data based on their annual assessment scores, we developed and implemented the following protocol:

1. An impairment in a cognitive domain should appear on at least two consecutive annual assessments to be considered; otherwise it was ignored.
2. If a subject exhibited impairment in one cognitive domain on an annual assessment, and this impairment disappeared on the following assessment, but then reappeared in the same cognitive domain on all the subsequent assessments, then the subject was considered impaired in this cognitive domain from the first assessment onward.
3. If a subject exhibited impairment in one cognitive domain on an annual assessment, and this impairment disappeared on the following assessment, but then reappeared in a different cognitive domain on all subsequent assessments but resulted in the same sub-type of MCI, either na-MCI or a-MCI, then the impairments were combined and the subject was again considered impaired from the first assessment onward. Otherwise, each impairment was treated separately.

4. Problem Definition & Methodology

Suppose that a database consists of N subjects, each residing in a living unit of R rooms, who were monitored in their homes for a few years using unobtrusive sensing technologies. For each subject, a sliding time window of size ω weeks is used to build a dataset of room activity distributions. These distributions estimate the probability of the subject being present in room r during a time interval t of the day, where $r = 1, \dots, R$, and $1 \leq t \leq K$, where K is the number of intervals per day. Datasets from all subjects are used to create a database of room activity distributions. Given this database, we are interested in autonomously classifying subjects as belonging to subjects who are cognitively intact or having MCI. In other words, we are formulating the problem as a classification problem with two classes: ‘cognitively intact’ corresponding to the “CIN” label in (3.2), and ‘MCI’ corresponding to the labels “TR” and “MCI” in (3.2).

In this work, we represent variables by lower case letters, e.g. d , vectors by bold lower case letters, e.g. $\mathbf{d}$, matrices by bold upper case letters, e.g. $\mathbf{D}$, and sets by calligraphic upper case letters, $\mathcal{D}$. Mathematically, the cognitive status recognition problem can be formulated as follows: The database consists of N subjects, each residing in a living unit of R rooms, tabulated as the following sets,

$$\mathcal{D}_1 = \{\mathbf{D}_{1,1}, \mathbf{D}_{1,2}, \dots, \mathbf{D}_{1,R}\}, \mathcal{D}_2 = \{\mathbf{D}_{2,1}, \mathbf{D}_{2,2}, \dots, \mathbf{D}_{2,R}\}, \dots, \mathcal{D}_N = \{\mathbf{D}_{N,1}, \mathbf{D}_{N,2}, \dots, \mathbf{D}_{N,R}\}.$$

(1)

Each $\mathbf{D}_{n,r}$ is a $s_n \times K$ matrix, where each row represents an estimated distribution of room r pertaining to subject n , and s_n is the total number of distributions estimated for each room for subject n . Each of the estimated distribution served as a data instance. Note that s_n is

different across subjects because subjects could have been monitored for different periods of time. Therefore, subjects do not necessarily have the same number of distributions for all rooms. Hence, the subscript n .

The database of room distributions (1) is used to supply distributions that best represent this database, referred to as “exemplars” using an off-line procedure. This constitutes the training stage. In a test stage, for a subject with R room activity distributions, the objective of the cognitive status recognition system is to assign a label to each room activity distribution and then use these labels to determine the cognitive status of the subject. Fig. 2 depicts the general overview of the proposed approach in assigning labels to room activity distributions corresponding to one room r , where $r = 1, \dots, R$. Note that the block diagram represents a two-stage system. Fig. 2a) represents the first stage being the training stage whereas Fig. 2b) represents the second stage being the test stage.

The training stage comprises two parts. For each subject, a sliding time window of size ω weeks is used to divide the sensor data for room r into frames each of size ω weeks. Then for each frame, the ‘room activity distribution estimator’ block estimates the corresponding room activity distribution. This results in a dataset of room activity distributions for room r per subject. A training set for room r is formed by combining all datasets of room activity distributions from all subjects. This training set is then processed through a clustering stage that consists of two sub-blocks. The first clustering sub-block performs high level clustering using k -means to deal with possible mislabeling of the distributions. Since subjects were assessed annually, then it was not possible to determine when exactly the transition to MCI occurred. Accordingly, as per our label assigning procedure, some distributions could potentially be mislabeled. For examples, some distributions could be labeled “CIN” when they should actually be labeled “TR” or the other way around. The relabeled distributions are then fed into the affinity propagation (AP) sub-block to decompose the training set into multiple clusters. Clustering in essence represents the core of the training stage. Each cluster is represented by one of its members, referred to as the “exemplar”. The output of the clustering stage, and in turn the training stage, is a set of exemplars $\mathbf{E}$, each exemplar representing a cluster of room activity distributions, and each cluster containing distributions belonging to one state of cognition only.

As for the test stage, a room activity distribution coming from a subject with unknown cognitive status is compared to the set of exemplars by computing the normalized KL-divergence between the unknown distribution and the set of exemplars. The distribution is then assigned the label of the exemplar that yields the smallest normalized KL-divergence. The cognitive status of the test subject based on room r activity distribution is then determined as the class to which the label of the activity distribution belongs, either ‘cognitively intact’ or ‘MCI’. Note that the overview portrayed by the block-diagram in Fig. 2 corresponds to one room only. Therefore, the training described above should be repeated R times. The details of each sub-block are discussed in the following subsections.

4.1. Room Activity Distribution Estimator

The first step in the cognitive status recognition process is creating a training set, $\mathbf{D}_r$, for each of the R rooms using the datasets of room activity distributions built for each subject. This is equivalent to taking a vertical slice of the database defined in (1). In other words,

$$\mathbf{D}_r = \begin{bmatrix} \mathbf{D}_{1,r} \\ \mathbf{D}_{2,r} \\ \vdots \\ \mathbf{D}_{N,r} \end{bmatrix}. \quad (2)$$

The process of building a dataset of room activity distributions for room r for subject n using a sliding window of size ω weeks is summarized in Fig. 3. Each frame of the sensor data, extracted using a sliding window of size ω weeks, is used to build a $K \times 1$ distribution denoted by $\mathbf{d}^{(c)}$ where $c \in \{-1, 1\}$ with ‘-1’ representing the ‘cognitively intact’ class, and ‘1’ representing the ‘MCI’ class.

The estimated room distribution provides the probability of the subject being present in room r during some time interval during the day. An inhomogeneous Poisson process can well-model this problem since a Poisson distribution models the number of occurrences of an event in a fixed time interval. If we define a variable y_r^t to capture the number of times a subject visits room r during time interval t , the probability of the subject being present in room r can then be given by

$$p(y_r) = \frac{e^{-\lambda} \lambda^{y_r^t}}{y_r^t!} \quad (3)$$

where λ is the Poisson distribution parameter. As discussed in [22] [23], because we do not expect the probability to monotonically increase or decrease throughout the day, we divide the day into K intervals and a λ_t is defined for each interval, where $1 \leq t \leq K$. Therefore, a day is represented by a matrix $\mathbf{X}$, defined as follows:

$$\mathbf{X} = [\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(t)}, \dots, \mathbf{x}^{(K)}] \quad (4)$$

where,

$$\mathbf{x}^{(t)} = x_j \text{ for } 1 \leq j \leq K \text{ s.t. } \begin{cases} x_j = 1 & t = j \\ x_j = 0 & t \neq j. \end{cases} \quad (5)$$

In other words, $\mathbf{X}$ takes the form of an identity matrix of size $K \times K$, where each column represents a time interval, and each interval is associated with a λ_t .

For each frame of size ω weeks, the resulting distribution is estimated as follows. The input space consists of a total number of $z = (7 \times \omega \times K)$ vectors or time intervals, and each of the z time intervals has a corresponding label y_r indicating the number of times the subject visits room r . The problem is then formulated as estimating the probability of the subject being present in room r given a time interval $\mathbf{x}^{(t)}$ as input. To estimate this probability, the λ_{rt} 's corresponding to each interval for room r need to be found, denoted by the vector $\boldsymbol{\lambda}_r$. Given this dataset of z time intervals along with the corresponding labels represented by the vector $\mathbf{y}_r = \{y_r^{(v)}\}_{v=1}^z$, the goal is to find $\boldsymbol{\lambda}_r$ that maximizes the likelihood function $L(\boldsymbol{\lambda}_r)$,

$$L(\boldsymbol{\lambda}_r) = p(\mathbf{y}_r | \mathbf{X}; \boldsymbol{\lambda}_r) = \prod_{v=1}^z \frac{e^{-\lambda_{rv}} \lambda_{rv}^{y_r^{(v)}}}{y_r^{(v)}!}. \quad (6)$$

In order to address the inhomogeneity in the problem, we define $\boldsymbol{\lambda}_r$ as a non-linear mapping of the input vectors. In other words, $\boldsymbol{\lambda}_r = \exp(\mathbf{X}^T \mathbf{w}_r)$, where $\mathbf{w}_r$ is a $K \times 1$ vector of weights. The problem then becomes finding this vector of weights, $\mathbf{w}_r$. By substituting this definition of $\boldsymbol{\lambda}_r$ in (6), maximizing the log of the resulting equation is equivalent to minimizing the negative of it, i.e. minimizing

$$\ell'(\mathbf{w}_r) = -\ell(\mathbf{w}_r) = \sum_{v=1}^z e^{\mathbf{x}^{(v)T} \mathbf{w}_r} - \sum_{v=1}^z y_r^{(v)} \mathbf{x}^{(v)T} \mathbf{w}_r + \sum_{v=1}^z \log(y_r^{(v)}!) \quad (7)$$

whose derivative is given by

$$\frac{d\ell'(\mathbf{w}_r)}{d\mathbf{w}_r} = \sum_{v=1}^z \mathbf{x}_t^{(v)} e^{\mathbf{x}^{(v)T} \mathbf{w}_r} - \sum_{v=1}^z y_r^{(v)} \mathbf{x}_t^{(v)}. \quad (8)$$

Once $\mathbf{w}_r$ is found, $\boldsymbol{\lambda}_r = [\lambda_{r1}, \lambda_{r2}, \dots, \lambda_{rK}]$ is computed as $\lambda_{rt} = e^{\mathbf{w}_r^T \mathbf{x}_t}$. The individual probabilities of the resulting distribution $\mathbf{d} = [d_1, d_2, \dots, d_K]$ are then estimated as,

$$d_t = e^{-\lambda_{rt}} \lambda_{rt} \text{ for } 1 \leq t \leq K. \quad (9)$$

Note that the estimated room activity distribution corresponds to one frame of sensor data only. Therefore, the above process of estimating the room activity distribution should be repeated to every frame as depicted in Fig. 3. The resulting dataset, $\mathbf{D}_{n,r}$ of activity distributions for room r pertaining to subject n is created as,

$$\mathbf{D}_{n,r} = \begin{bmatrix} \leftarrow & \mathbf{d}_{n,1}^{(-1)} & \rightarrow \\ \leftarrow & \mathbf{d}_{n,2}^{(-1)} & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{d}_{n,i-3}^{(-1)} & \rightarrow \\ \leftarrow & \mathbf{d}_{n,i-2}^{(-1)} & \rightarrow \\ \leftarrow & \mathbf{d}_{n,i-1}^{(-1)} & \rightarrow \\ \leftarrow & \mathbf{d}_{n,i}^{(1)} & \rightarrow \\ \leftarrow & \mathbf{d}_{n,i+1}^{(1)} & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{d}_{n,s_n}^{(1)} & \rightarrow \end{bmatrix}. \quad (10)$$

A subject who transitioned to MCI during the monitoring period would have $(i - 1)$ distributions belonging to the ‘cognitively intact’ class and $(s_n - i + 1)$ distributions belonging to the ‘MCI’ class. A subject who did not transition to MCI would have all s_n distributions belonging to the ‘cognitively intact’ class.

4.2. Relabeling Using k-means

Because subjects were assessed annually for MCI, it was not possible to determine at what point exactly the transition took place. Consequently, for subjects who transitioned to MCI, some distributions could potentially be mislabeled. This was demonstrated in our previous work [22] [23], where we demonstrated that the majority of the subjects started transitioning to MCI several months prior to the start of the “TR” period, defined in Fig. 1. Therefore, several distributions labeled “CIN” should be actually labeled “TR”. As for subjects who started transitioning after the start of the transitioning period, some of the distributions labeled “TR” should most likely be labeled “CIN”.

In order to address this issue, we added a preprocessing step of high-level clustering using k -means. We are not much concerned about the distributions bordering the “TR” period and the “MCI” period since we are formulating the problem as a classification problem and both of these periods are lumped into one class, the ‘MCI’ class. Therefore, for each subject who transitioned to MCI, the room activity distributions for each room are clustered using k -means into two clusters: one cluster for the ‘cognitively intact’ class, and another cluster for the ‘MCI’ class. As a result, for subjects who started transitioning prior to the start of the “TR” period will have some of their “CIN” distributions assigned the label 1, indicating that they belong to the ‘MCI’ class. This change is reflected in the datasets of room activity distributions for each subject. For subjects who did not transition, no change occurs in their datasets of activity distributions. For subjects who transitioned prior to the start of the “TR” period, possible changes are shown in (11).

$$\mathbf{D}_{1,r} = \begin{bmatrix} \leftarrow d_{n,1}^{(-1)} \rightarrow \\ \leftarrow d_{n,2}^{(-1)} \rightarrow \\ \vdots \\ \leftarrow d_{n,i-4}^{(-1)} \rightarrow \\ \leftarrow d_{n,i-3}^{(1)} \rightarrow \\ \leftarrow d_{n,i-2}^{(1)} \rightarrow \\ \leftarrow d_{n,i-1}^{(1)} \rightarrow \\ \leftarrow d_{n,i}^{(1)} \rightarrow \\ \leftarrow d_{n,i+1}^{(1)} \rightarrow \\ \vdots \\ \leftarrow d_{n,s_n}^{(1)} \rightarrow \end{bmatrix} \quad (11)$$

By comparing (11) to (10), we see that distributions bordering the “CIN” and the “TR” periods such $(i-3)$, $(i-2)$, and $(i-1)$ which were previously labeled ‘-1’, are now labeled ‘1’, meaning that they belong to the ‘MCI’ class. Similarly, for subjects who started transitioning after the “TR” period started, in comparison with (10), distributions such as (i) , $(i+1)$, and $(i+2)$ would be labeled ‘-1’ instead of ‘1’, meaning that they belong to the ‘cognitively intact’ class instead.

4.3. Kullback–Leibler Divergence Measure

Since we are dealing with probability distributions, a good measure of similarity between the different distributions is the Kullback-Leibler (KL)-divergence [31]. The KL-divergence of an activity distribution, $\mathbf{q}$ from an activity distribution $\mathbf{p}$, is computed as,

$$D_{KL}(\mathbf{p}, \mathbf{q}) = \sum_t^K \mathbf{p}(t) \ln \frac{\mathbf{p}(t)}{\mathbf{q}(t)} \quad (12)$$

where t and K , as defined in (4) and (5), are the interval number and the number of intervals per day, respectively. A KL-divergence of 0 implies that the two activity distributions are identical. The higher the KL-divergence measure the greater the difference between the two activity distributions. In order to limit the values of KL-divergence to be between 0 and 1, we define a normalized version of the KL-divergence as

$$D_{KL}^{\text{norm}}(\mathbf{p}, \mathbf{q}) = 1 - \exp(-D_{KL}(\mathbf{p}, \mathbf{q})). \quad (13)$$

One useful property of the KL-divergence is asymmetry. So $D_{KL}(\mathbf{p}, \mathbf{q}) \neq D_{KL}(\mathbf{q}, \mathbf{p})$. Accordingly, we are able to quantify the divergence of distributions pertaining to a cognitive state from the distributions pertaining to the preceding cognitive state. For example, we can quantify the divergences of the “TR” distributions from the “CIN” distributions, the

divergence of the “MCI” distributions from the “CIN” distributions, and the divergence of the ‘MCI’ distributions from the “TR” distributions.

4.4. Affinity Propagation

The next sub-block in the clustering step is using affinity propagation (AP) [32] to divide the training set distributions into different clusters and to appoint a representative to each cluster, referred to as the “exemplar”. AP is an algorithm that simultaneously considers all distributions as potential exemplars and recursively exchanges real-valued messages among the distributions until a good set of exemplars and clusters emerges. Clustering is based on the exchange of two types of messages: the “responsibility” message to decide which distributions are exemplars, and the “availability” message to decide to which cluster a distribution belongs. The responsibility message is given by

$$r(i, j) = s(i, j) - \max_{j' \text{ s.t. } j' \neq j} \{a(i, j') + s(i, j')\} \quad (14)$$

where i, j , $s(i, j)$ is the pairwise similarity that indicates how well the distribution $\mathbf{d}_j$ is suited to be the exemplar for the distribution $\mathbf{d}_i$, and is defined as

$$s(i, j) = D_{KL}^{\text{norm}}(\mathbf{d}_i, \mathbf{d}_j) \forall i, j \in \{1, 2, \dots, M\} \quad (15)$$

where M is the total number of distributions in the training set for each room, and the availability message is given by

$$a(i, j) = \min \left\{ 0, r(i, j) + \sum_{i' \text{ s.t. } i' \neq i, j} \max\{0, r(i', j)\} \right\}. \quad (16)$$

In addition to the measure of similarity, AP takes as input a set of real numbers, known as self-similarity or *preference* (p) for each activity distribution, so that distributions with larger values of p are more likely to be chosen as exemplars. For the proposed cognitive status recognition system, the self-similarity p is proportional to the median of the input similarities, that is

$$p = \beta * \text{median}\{s(i, j), \forall i, j \in \{1, 2, \dots, M\}\} \quad (17)$$

where β is a constant that controls the number of clusters to be generated in an inversely proportional manner. In other words, as the value of β decreases, more clusters will be generated.

AP is chosen as the clustering technique because it handles better the fact that activity distributions belonging to the same class should not be forced to have only one exemplar or center as is the case with k -means for example. Ideally, it would be perfect if we are able to

generate only one cluster for ‘cognitively intact’ class and another cluster for the ‘MCI’ class. However, practically this is not the case given that different individuals are more likely to exhibit different patterns of daily activity. Accordingly, with AP, we are able to divide the training set of distributions into different clusters, each represented by an exemplar while ensuring that members of each cluster belong to the same class.

The output of AP is a matrix of exemplars, E_r , represented as,

$$E_r = \{e_1, e_2, \dots, e_j, \dots, e_{m_r}\} \quad (18)$$

along with a vector of exemplar labels, c_r , represented as,

$$c_r = \{c_1, c_2, \dots, c_j, \dots, c_{m_r}\} \quad (19)$$

for each room r . Each column, e_j , of the matrix E_r represents an exemplar, and each label, c_j , in c_r indicates whether the corresponding exemplar represents a cluster of ‘cognitively intact’ distributions or a cluster of ‘MCI’ distributions, and $c_j \in \{-1, 1\}$ with ‘-1’ corresponding to the ‘cognitively intact’ class and ‘1’ corresponding to the ‘MCI’ class, where $1 \leq j \leq m_r$. Note that the number of exemplars m_r is room dependent because different subjects can potentially exhibit more variability in their activity in some rooms than other rooms. For example, subjects, or older adults in general, normally have a high probability of being in the bedroom early in the morning and late at night. You would expect to see more variability in other rooms such as the living room. Therefore, it would be intuitive to expect fewer exemplars to represent bedroom distributions than living room distributions. Also note that clustering is done per room. Therefore, clustering should be done R times, one for each room. A summary of our proposed cognitive status recognition methodology is presented in Algorithm 1.

5. Results & Discussion

The cognitive status of the 85 subjects based on annual neuropsychological assessments during the monitoring period were as follows:

1. 59 subjects remained cognitively intact throughout the monitoring period.
2. 11 subjects had a-MCI at baseline or transitioned to a-MCI during the monitoring period.
3. 15 subjects had na-MCI at baseline or transitioned to na-MCI during the monitoring period.

Baseline here corresponds to the beginning of the data that we received from ORCATECH, which for some subjects did not necessarily represent the actual subjects’ baseline - when the subjects were recruited. Subjects were monitored for different periods of time for reasons such as subjects passing away, moving out of the metropolitan area, or feeling overwhelmed

by study procedures. Table 2 presents statistics associated with the monitoring periods (in weeks) of the 85 subjects.

Sensing technologies were deployed in the subjects' living units. Subjects lived in a variety of settings from apartments in organized retirement communities to single-family detached homes. In order to standardize comparison among the recruited subjects, we decided to base our analysis on four rooms: main bedroom, main bathroom, kitchen, and living room. In other words, R in (1), was equal to 4.

Algorithm 1

Cognitive status recognition algorithm using clustering of room activity distributions

---

Input:

- Sets of estimated room activity distributions $\{\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_N\}$ for R rooms for all subjects

Output:

- Cognitive status of an older adult: 'cognitively intact' or 'MCI'

Training Stage

- 1: **for** room $r = 1, \dots, R$ **do**
- 2: For subjects who transitioned to MCI, relabel distributions using k -means
- 3: Run Affinity Propagation to extract a set of exemplars, $\mathbf{E}_r$, and vector of exemplar labels, $\mathbf{c}_r$
- 4: **end for**

- Output of Training Stage: $\{\mathbf{E}_1, \mathbf{c}_1, \mathbf{E}_2, \mathbf{c}_2, \dots, \mathbf{E}_R, \mathbf{c}_R\}$

Classification Stage

For an older adult with room activity distributions $\{d_1^{test}, d_2^{test}, \dots, d_R^{test}\}$:

- 1: **for** each d_r^{test} **do**
- 2: compute normalized KL-divergences $\{D_{KL_1}^{norm}, D_{KL_2}^{norm}, \dots, D_{KL_{m_r}}^{norm}\}$ of d_r^{test} from exemplars in $\mathbf{E}_r$
- 3: d_r^{test} label ($c_{d_r^{test}}$) = label of exemplar corresponding to $\min \{D_{KL_1}^{norm}, D_{KL_2}^{norm}, \dots, D_{KL_{m_r}}^{norm}\}$
- 4: **end for**

- Output of Classification Stage: $\{c_{d_1^{test}}, c_{d_2^{test}}, \dots, c_{d_R^{test}}\}$

- Return cognitive status of older adult as the class that corresponds to the majority of $\{c_1, c_2, \dots, c_R\}$

5.1. Data Cleaning

Because we focused on homes with single occupants, before we could start estimating the rooms activity distributions, it was necessary to clean the data. The weekly questionnaires that the subjects were asked to fill out were very helpful for the cleaning process. The cleaning process involved discarding days on which subjects had their in-home annual assessments, days on which subjects had any visitors over, days which subjects spent away from the home, days spent in ER, days on which subjects had maintenance people over, or

days on which people reported health problems that limited their activity, and days on which sensors failed to fire due to a dead battery or other malfunction.

5.2. Evaluation Score

Although the majority of clinical studies use sensitivity and specificity scores as evaluations metrics, next we justify why these scores can be misleading especially with imbalanced databases like the one at hand. Suppose we have a database of 100 subjects, 80 of which are cognitively intact and the remaining 20 have MCI. Let's assume we are able to detect MCI in this dataset with equal sensitivity and specificity scores of 0.8. At first glance, one is tempted to believe that the system is performing very well. However, if we analyze these numbers, we will discover that the system is suffering from a rate of false positive as high as 50%. Recall that,

$$\text{sensitivity}=\text{recall}=\frac{TP}{TP+FN}, \quad (20)$$

and,

$$\text{specificity}=\frac{TN}{TN+FP}, \quad (21)$$

where TP stands for true positives, FN stands for false negatives, TN stands for true negatives, and FP stands for false positives. A sensitivity score of 0.8 means that we are able to correctly classify 16 of the 20 subjects with MCI, i.e., $TP = 16$. Similarly, a specificity score of 0.8 means that we are able to correctly classify 64 of the 80 cognitively intact subjects, i.e., $TN = 64$, but which also means that $FP = 16$. In other words, in order to correctly classify 16 of the 20 MCI subjects, we incorrectly classified as many subjects to be experiencing symptoms of MCI. However, because the database is imbalanced - having many more cognitively intact subjects than subjects with MCI, this high rate of false positives did not significantly impact the sensitivity and specificity scores. On the other hand, if we compute the precision score using the given TP and FP,

$$\text{precision}=\frac{TP}{TP+FP}, \quad (22)$$

we find that $\text{precision} = 0.5$. The F-score, defined as,

$$F=2 \cdot \frac{\text{precision} \times \text{recall}}{\text{precision}+\text{recall}} \quad (23)$$

is equal to 0.62. These scores better reflect the true performance of systems with imbalanced databases.

Accordingly, we use the F score instead to evaluate the performance of our approach, and to put more emphasis on precision than recall, i.e. making sure we penalize the system for false positives, we report the $F_{0.5}$, defined as

$$F_{0.5} = (1 + 0.25) \frac{\text{precision} \times \text{recall}}{(0.25 \times \text{precision} + \text{recall})} \quad (24)$$

which, using the sensitivity and specificity scores given in the example above, is equal to 0.54. Note that a high $F_{0.5}$ means high F , sensitivity, and specificity scores.

5.3. 80% Training Data and 20% Test Data

Since not all the subjects transitioned to MCI during the monitoring period, it was not possible to test our approach using a leave-one-subject-out evaluation process by including all 85 subjects. Alternatively, we randomly divided our database of subjects into a training set and a test set of subjects, such that the training set contained 80% of the distributions belonging to each cognitive state - "CIN", "TR", and "MCI" and each MCI subtype - "a-MCI", and "na-MCI", and the test set contained 20% of the distributions belonging to each cognitive state and MCI subtype. Note that the creation of these sets was based on full subject data in order to ensure that no distributions pertaining to the same subject existed in both the training and the test sets.

We experimented with different time frames ranging from 4 – 24 weeks as well as different slide sizes ranging from 2 – 4 weeks. We empirically found that a value of β , defined in (17), of 0.5 resulted in the best performance. The results obtained using a time frame of 4 weeks, i.e. $\omega = 4$, and slide sizes of 2 weeks, 3 weeks, and 4 weeks are summarized in Table 3. The 'Bedroom', 'Bathroom', 'Kitchen', and 'Living Room' $F_{0.5}$ scores reported in Table 3 corresponded to detecting MCI using the room distributions separately. The 'Combined' $F_{0.5}$ scores corresponded to detecting MCI using the majority voting of the four rooms. At least three rooms should have declared that the subject was having MCI for the subject to belong to the 'MCI' class. If a tie existed, where two rooms declared the subject was experiencing symptoms of MCI and the other two rooms declared that the subject was still cognitively intact, the subject was classified as belonging to the 'MCI' class.

According to Table 3, we were able to detect MCI in older adults with an $F_{0.5}$ score of 0.625 based on the bedroom distributions only, an $F_{0.5}$ score of 0.573 based on the bathroom distributions only, an $F_{0.5}$ score of 0.615 based on the kitchen distributions only, an $F_{0.5}$ score of 0.439 based on the living room distributions only, and an $F_{0.5}$ score of 0.760 based on majority voting of the four rooms, as described in Algorithm 1, all using a slide size of 2 weeks. Interestingly, using majority voting of the four rooms, a 13.5% improvement in performance was achieved compared to the best performance achieved using individual room distributions, which was using the bedroom in this case. As the slide size increased, we continued to observe that the majority voting of the four rooms resulted in an improvement of approximately 6% compared to using the bedroom distributions only. Another interesting observation is that, using individual room distributions only, although higher $F_{0.5}$ scores

were achieved as the slide size increased, we did not note a significant change in the $F_{0.5}$ scores using majority voting of the four rooms.

Similarly, Table 4 presents the $F_{0.5}$ scores using a time frame of 8 weeks and slide sizes of 2 weeks, 3 weeks, and 4 weeks. Interestingly, using the individual room distributions, we were able to achieve a performance that was comparable to the majority voting performance that was obtained using a time frame of 4 weeks. This was true for the bedroom, bathroom, and living room distributions but not the kitchen. Despite the low $F_{0.5}$ scores using the kitchen distributions, using majority voting a 2% improvement in performance was recorded compared to time frame of 4 weeks. According to Table 4, for a time frame of 8 weeks, we found that a slide size of 2 weeks resulted in the best performance.

As the window size continued increasing, better performance and higher $F_{0.5}$ scores were noted until a time frame of 20 weeks, after which the performance started dropping, as shown in Tables 5 and 6. The best performance was achieved using a time frame of 20 weeks and a slide size of 4 weeks as shown in Table 5. An $F_{0.5}$ score of 0.815 was achieved using the bedroom distributions only, an $F_{0.5}$ score of 0.789 was achieved using the bathroom distributions only, an $F_{0.5}$ score of 0.791 was achieved using the kitchen distributions only, an $F_{0.5}$ score of 0.747 was achieved using the living room distributions only, and an $F_{0.5}$ score of 0.789 was achieved using majority voting of the four rooms.

One reason that could explain the worsening performance for windows bigger than 20 weeks is that as the window size increased, the number of estimated distributions per room for each subject decreased. Many subjects who transitioned to MCI during the monitoring period also had a long period of being cognitively intact. This was in addition to subjects who remained cognitively intact throughout the monitoring period. As a result, the database clearly contained more distributions belonging to the ‘cognitively intact’ class than the ‘MCI’ class, which in turn meant that the set of exemplars created from the training stage most likely contained more exemplars representing distributions that belonged to the ‘cognitively intact’ class. If exemplars representing one class dominate the set of exemplars, then there is a higher probability for a room activity distribution belonging to a test subject to be labeled as belonging to the class with the higher number of exemplars. In this case, it is the ‘cognitively intact’ class.

5.4. Leave-One-Subject-Out Cross Validation

Out of the 26 subjects who were labeled as having MCI, 22 of them were cognitively intact for some period and then transitioned to either a-MCI or na-MCI. So in order to test the validity of the results obtained using the 80% – 20% division of the database, using the features that resulted in the best performance, i.e., $\omega = 20$ and slide size = 4 weeks, we completed a leave-one-subject-out cross-validation process using the 22 subjects who transitioned to MCI during the monitoring period. The corresponding $F_{0.5}$ scores achieved are summarized in Table 7.

As Table 7 shows, comparable performance was noted which confirmed the results obtained using the 80% – 20% division of the database. The majority voting $F_{0.5}$ score was higher by approximately 6%. This was mostly likely because the distributions pertaining to the 22

subjects were more balanced. So there was not a big difference between the number of room distributions belonging to the ‘cognitively intact’ class and the number of distributions belonging to the ‘MCI’ class as was the situation with the 80% – 20% case. As a result, the number of exemplars representing each class was most likely balanced as well.

5.5. a-MCI & na-MCI Sub-Types

We were also interested in investigating the different sub-types of MCI and how the room activity distributions of subjects experiencing symptoms of each of the sub-types differed from their room distributions when they were cognitively intact. Of the 15 subjects who were labeled as having na-MCI, 14 of them transitioned to na-MCI during the monitoring period. Table 8 summarizes the $F_{0.5}$ scores corresponding to the data pertaining to these 14 subjects obtained using leave-one-subject-out cross-validation.

The performance of the algorithm was impressive; the $F_{0.5}$ scores were all over 0.9 using the individual room distributions, and with majority voting the $F_{0.5}$ score was brought up to 0.958. This result was very interesting since it indicates that subjects experiencing symptoms of na-MCI have room activity distributions that were significantly different from their room activity distributions when they were cognitively intact. According to the definition of na-MCI, it is a subtype of MCI that impacts the thinking skills other than memory of an individual, so an individual experiencing symptoms of na-MCI is less able to make sound decisions, or to judge the sequence of steps needed to complete a complex task [7]. Accordingly, it is expected for changes to exist in the way older adults start completing certain activities as they start experiencing symptoms of na-MCI, which would be reflected as changes in their room activity distributions.

As for a-MCI, of the 11 subjects who were labeled as having a-MCI, 8 of them transitioned to a-MCI during the monitoring period. Table 9 summarizes the $F_{0.5}$ scores corresponding to the data pertaining to these 8 subjects obtained using leave-one-subject-out cross-validation. The observations based on these results were very interesting as well. First, the $F_{0.5}$ scores dropped drastically except for the bedroom, even majority voting yielded an $F_{0.5}$ score of 0.539, which was slightly better than random classification. This indicated that there were no significant differences between the subjects’ room activity distributions when they were cognitively intact and when they were experiencing symptoms of a-MCI.

According to the definition of a-MCI, it is a sub-type of MCI that primarily affects the memory of an individual, so an individual experiencing symptoms of a-MCI is less able to remember information such as appointments, conversations, or recent events [7]. In other words, subjects with MCI still have their problem-solving skills intact. Therefore, individuals with a-MCI most likely continue carrying out activities in the same sequence and fashion as when they were cognitively intact since their ability to judge the sequence of steps needed to complete a complex task is not compromised. Consequently, changes in their room activity distributions are less likely to occur.

However, the majority of the subjects did have significant differences in their bedroom activity distributions after they transitioned to a-MCI. This is mostly attributed to disturbed sleep patterns, mainly in the form of waking up at night. This finding is supported by a

recent study by McGee et al., who concluded that sleep disturbance is common in subjects with a-MCI, and that many individuals with a-MCI have more than one symptom of sleep disturbance [33] [34]. Besides waking up at night, other symptoms include initial insomnia, snoring, breathing difficulties, nightmares, or excessive daytime sleepiness. In another study, Hita-Yañez et al. investigated the existence of sleep disturbances in MCI subjects via polysomnographic recordings in cognitively healthy older adults and in patients with MCI [35]. Hita-Yañez et al. concluded that sleep disruptions were evident in MCI patients. These differences in results between aMCI and naMCI may be related to differences in the pathology underlying these MCI sub-types.

6. Future Work

One limitation of the proposed work is the prerequisite of having sufficient data to build a database of room activity distributions in order to extract exemplars that would represent the different states of cognition. Accordingly, an alternative method has to be designed in order to address cases where such a database of room activity distributions cannot be created. One direction for future work is to investigate the feasibility of using changepoint detection for automatic detection of MCI in older adults. If feasible, such an approach would not require a database of room activity distributions. Instead, it would track each individual separately and perform intra-individual comparison by comparing each subject's activity distributions to their past distributions, and assess their cognition accordingly.

Another limitation of this work is the small population size of older adults who transitioned to MCI during the monitoring period. We believe it is very important to test our proposed approach on a larger population size in order to corroborate these results and confirm their validity.

7. Conclusion

In conclusion, we proposed a home-based clustering technique for automatic detection of MCI in older adults. We formulated the problem as a classification problem with two classes: 'cognitively intact' class and 'MCI' class. A sliding window was used to create a database of room activity distributions from the sensor data pertaining to 85 subjects. Affinity Propagation was then applied to cluster the room distributions in the database and to extract exemplars to represent the distributions belonging to the different cognitive states. This represented the training stage. In the test stage, the room distributions belonging to a test subject were compared to the exemplars and were assigned the labels of the exemplars that yielded the smallest normalized KL-divergence. The labels of the room distributions were then used to determine to which class the test subject belonged. This answered our first research question.

By varying the size of the time frame, mainly from 4 weeks to 24 weeks, and the slide size from 2 weeks to 4 weeks, we observed an overall improvement in performance, reflected as higher $F_{0.5}$ scores. We did not really detect significant differences between the different slide sizes. The best performance was obtained using a time frame of 20 weeks and a slide size of 4 weeks. An $F_{0.5}$ score of 0.815 was achieved using the bedroom distributions only, an $F_{0.5}$

score of 0.789 was achieved using the bathroom distributions only, an $F_{0.5}$ score of 0.791 was achieved using the kitchen distributions only, an $F_{0.5}$ score of 0.747 was achieved using the living room distributions only, and an $F_{0.5}$ score of 0.789 was achieved using majority voting of the the four rooms. This answered our second research question.

Finally, we investigated the different sub-types of MCI, namely a-MCI and na-MCI. Using leave-one-subject-out cross-validation, we found that subjects experiencing symptoms of na-MCI had room activity distributions that were significantly different from their room activity distributions when there were cognitively intact. This was reflected as very high $F_{0.5}$ scores, with a majority voting $F_{0.5}$ score of 0.958. As for a-MCI, we did not find significant differences between the subjects room activity distributions when experiencing symptoms of a-MCI and when cognitively intact, except for the bedroom distributions. This interesting finding was most likely attributed to disturbed sleep patterns which have been shown to be prevalent in subject with a-MCI. This answered our third research question.

References

1. Alzheimer Society of Canada. [Online]. Available: <http://www.alzheimer.ca/en/About-dementia/What-is-dementia/Dementia-numbers>
2. Centers for Disease Control and Prevention. [Online]. Available: http://www.cdc.gov/aging/pdf/cognitive_impairment/cogimp_poilicy_final.pdf
3. Fortinsky RH, Wasson JH. How do physicians diagnose dementia? Evidence from clinical vignette responses. *American Journal of Alzheimer's Disease and Other Dementias*. 1997; 12(2):51–61.
4. Clarfield A. The reversible dementias: Do they reverse? *Annals of Internal Medicine*. 1988; 109(6): 476–486. [PubMed: 3046450]
5. Boise L, Camicioli R, Morgan DL, Rose JH, Congleton L. Diagnosing dementia: Perspectives of primary care physicians. *The Gerontologist*. 1999; 39(4):457–464. [PubMed: 10495584]
6. Jack CR, Albert MS, Knopman DS, McKhann GM, Sperling RA, Carrillo MC, Thies B, Phelps CH. Introduction to the recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's & Dementia*. 2011; 7(3): 257–262.
7. Petersen RC. Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*. 2004; 256(3):183–194. [PubMed: 15324362]
8. Camicioli R, Howieson D, Oken B, Sexton G, Kaye J. Motor slowing precedes cognitive impairment in the oldest old. *Neurology*. 1998; 50(5):1496–1498. [PubMed: 9596020]
9. Aggarwal NT, Wilson RS, Beck TL, Bienias JL, Bennet DA. Motor dysfunction in mild cognitive impairment and the risk of incident alzheimer disease. *Arch Neurolology*. 2006; 63(12):1763–1769.
10. Brumitt, B.; Meyers, B.; Krumm, J.; Kern, A.; Shafer, SA. *Proc. International Symposium on Handheld and Ubiquitous Computing*. London, UK: Springer-Verlag; 2000. Easyliving: Technologies for intelligent environments; p. 12-29.
11. Helal S, Mann W, El-Zabadani H, King J, Kaddoura Y, Jansen E. The gator tech smart house: a programmable pervasive space. *Computer*. 2005; 38(3):50–60.
12. Abowd, GD.; Atkeson, CG.; Bobick, AE.; Essa, IA.; Elizabeth, BM. *Proc. Conf. Human Factors in Computing Systems (CHI)*. ACM Press; 2000. Living laboratories: The future computing environments group at the Georgia Institute of Technology; p. 215-216.
13. GE Healthcare. [Online]. Available: http://www.careinnovations.com/Data/Downloads/Quietcare/QuietCare_Brochure.pdf
14. Dawadi PN, Cook DJ, Schmitter-Edgecombe M. Automated cognitive health assessment using smart home monitoring of complex tasks. *IEEE Transactions on Systems, Man, and Cybernetics*. 2013; 43(6):1302–1313.

15. Sixsmith AJ. An evaluation of an intelligent home monitoring system. *Journal of Telemedicine and Telecare*. 2000; 6(2):63–72. [PubMed: 10824373]
16. Skubic, M.; Guevara, RD.; Rantz, M. Proceedings of the 10th International Smart Homes and Health Telematics Conference on Impact Analysis of Solutions for Chronic Disease Prevention and Management, ICOST'12. Berlin, Heidelberg: Springer-Verlag; 2012. Testing classifiers for embedded health assessment; p. 198–205.
17. Kaye JA, Maxwell SA, Mattek N, Hayes TL, Dodge H, Pavel M, Jimison HB, Wild K, Boise L, Zitzelberger TA. Intelligent systems for assessing aging changes: Home-based, unobtrusive, and continuous assessment of aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2011; 66B(suppl 1):i180–i190.
18. Austin, D.; Leen, T.; Hayes, TL.; Kaye, J.; Jimison, H.; Mattek, N.; Pavel, M. 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2010. IEEE; 2010. Model-based inference of cognitive processes from unobtrusive gait velocity measurements; p. 5230–5233.
19. Dodge HH, Mattek NC, Austin D, Hayes TL, Kaye JA. Inhome walking speeds and variability trajectories associated with mild cognitive impairment. *Neurology*. 2012; 78(24):1946–1952. [PubMed: 22689734]
20. Akl A, Taati B, Mihailidis A. Autonomous unobtrusive detection of mild cognitive impairment in older adults. *IEEE Transactions on Biomedical Engineering*. 2015 May; 62(5):1383–1394. [PubMed: 25585407]
21. Hayes TL, Abendroth F, Adami A, Pavel M, Zitzelberger TA, Kaye JA. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. *Alzheimer's and Dementia*. 2008; 4(6):395–405.
22. Akl, A.; Snoek, J.; Mihailidis, A. 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2014. IEEE; 2014. Generalized linear models of home activity for automatic detection of mild cognitive impairment in older adults; p. 680–683.
23. Akl A, Snoek J, Mihailidis A. Unobtrusive detection of mild cognitive impairment in older adults through home monitoring. revisions submitted to the *IEEE Journal of Biomedical and Health Informatics* on. 2015 Aug 19.
24. Wechsler, D. Wechsler Memory Scale - Revised. San Antonio, TX: The Psychological Corporation; 1987.
25. Lezak, MD. Neuropsychological Assessment. New York: Oxford University Press; 1995.
26. Reitan, RM.; Wolfson, D. The Halstead-Reitan Neuropsychological Test Battery. Tempe, AZ: Neuropsychology Press; 1985.
27. Wechsler, D. Wechsler Memory Scale (WMS-III). San Antonio, TX: The Psychological Corporation; 1997.
28. Kaplan, EF.; Goodglass, H.; Weintraub, S. The Boston Naming Test. Philadelphia: Lea & Febiger; 1983.
29. Wechsler, D. Wechsler Adult Intelligence Scale - Revised Manual. New York: Psychological Corporation; 1981.
30. Weintraub S, Salmon D, Mercaldo N, Ferris S, Graff-Radford NR, Chui H, Cummings J, DeCarli C, Foster NL, Galasko D, Peskind E, Dietrich W, Beekly DL, Kukull WA, Morris JC. The alzheimer's disease centers' uniform data set (UDS): The neuropsychologic test battery. *Alzheimer Disease and Associated Disorders*. 2009; 23(2):91–101.
31. Pinto, D.; Benedí, J-M.; Rosso, P. Clustering narrow-domain short texts by using the kullback-leibler distance. In: Gelbukh, A., editor. *Computational Linguistics and Intelligent Text Processing*, ser. Lecture Notes in Computer Science. Vol. 4394. Berlin Heidelberg: Springer; 2007. p. 611–622.
32. Frey BJ, Dueck D. Clustering by passing messages between data points. *Science*. 2007; 315(5814):972–976. [PubMed: 17218491]
33. McGee JS, Darby E, Biernot J, Doody R. Amnestic MCI, sleep disturbance, and cognitive functioning. *Journal of the International Neuropsychological Society*. 2012; 18(Suppl S1):277. [PubMed: 22300593]
34. Hayes TL, Riley T, Mattek N, Pavel M, Kaye JA. Sleep Habits in Mild Cognitive Impairment. *Alzheimer Disease and Associated Disorders*. 2014; 28(2):145–150. [PubMed: 24145694]

35. Hita-Yañez E, Atienza M, Gil-Neciga E, Cantero JL. Disturbed sleep patterns in elders with mild cognitive impairment: the role of memory decline and apoe e4 genotype. *Current Alzheimer Research*. 2012; 9(3):290–297. [PubMed: 22211488]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1.
Example of a subject who was assessed as having MCI on the 2nd year assessment onward.

Fig. 2. General overview of the cognitive status recognition process using distributions corresponding to room r . (a) Training Stage. (b) Test Stage.

Fig. 3.

The process of building a dataset of room activity distributions for room r for subject n

Table 1

Six neuropsychological tests used to assess the subjects cognitive capacity.

Cognitive Domain	Test
Memory	Logical Memory Delayed Recall [24]
Executive Function	Category Fluency Animals [25] or Trail Making B [26]
Attention	Wechsler Adult Intelligence Scale (WAIS) Digit Symbol [27]
Language	Boston Naming Test [28]
Visuospatial Function	WAIS Revised Block Design [29]

Table 2

Statistics of monitoring periods of subjects

monitoring period (in weeks)				
minimum	maximum	mean	median	std. deviation
42.5	257	172.6	184.2	60.1

Table 3

$F_{0.5}$ scores with $\omega = 4$ weeks and slide sizes of 2 weeks, 3 weeks, and 4 weeks.

Slide Size (in weeks)	$F_{0.5}$ Scores				
	Bedroom	Bathroom	Kitchen	Living Room	Combined
2	0.625	0.573	0.615	0.439	0.760
3	0.696	0.567	0.477	0.570	0.762
4	0.683	0.596	0.441	0.525	0.762

Table 4

$F_{0.5}$ scores with $\omega = 8$ weeks and slide sizes of 2 weeks, 3 weeks, and 4 weeks.

Slide Size (in weeks)	$F_{0.5}$ Scores				
	Bedroom	Bathroom	Kitchen	Living Room	Combined
2	0.799	0.788	0.564	0.717	0.780
3	0.794	0.780	0.422	0.789	0.778
4	0.726	0.785	0.472	0.744	0.776

Table 5

$F_{0.5}$ scores with $\omega = 20$ weeks and slide sizes of 2 weeks, 3 weeks, and 4 weeks.

Slide Size (in weeks)	$F_{0.5}$ Scores				
	Bedroom	Bathroom	Kitchen	Living Room	Combined
2	0.817	0.781	0.792	0.747	0.788
3	0.815	0.774	0.789	0.727	0.785
4	0.815	0.789	0.791	0.747	0.789

Table 6

$F_{0.5}$ scores with $\omega = 24$ weeks and slide sizes of 2 weeks, 3 weeks, and 4 weeks.

Slide Size (in weeks)	$F_{0.5}$ Scores				
	Bedroom	Bathroom	Kitchen	Living Room	Combined
2	0.812	0.788	0.490	0.747	0.780
3	0.813	0.758	0.492	0.744	0.783
4	0.806	0.783	0.533	0.705	0.777

Table 7

$F_{0.5}$ scores window of size $\omega = 20$ weeks and slide size = 4 weeks associated with leave-one-subject-out cross-validation.

$F_{0.5}$ Scores				
Bedroom	Bathroom	Kitchen	Living Room	Combined
0.842	0.829	0.813	0.826	0.856

Table 8

$F_{0.5}$ scores window of size $\omega = 20$ weeks and slide size = 4 weeks with naMCI distributions only.

$F_{0.5}$ Scores				
Bedroom	Bathroom	Kitchen	Living Room	Combined
0.941	0.932	0.926	0.913	0.958

Table 9

$F_{0.5}$ scores for $\omega = 20$ weeks and slide size = 4 weeks with aMCI distributions only.

$F_{0.5}$ Scores				
Bedroom	Bathroom	Kitchen	Living Room	Combined
0.784	0.524	0.498	0.574	0.539