An Approach to Cognitive Assessment in Smart Home

Prafulla Dawadi, Diane Cook
School of Electrical Engineering and Computer Science,

Carolyn Parsey, Maureen Schmitter-Edgecombe,
Department of Psychology,

Washington State University, USA
{pdawadi , cook@eecs.wsu.edu},{cparsey,schmitter-e@wsu.edu}

Miya Schneider
Department of Computer Science
Brown University, USA
miya_schneider@brown.edu

ABSTRACT
In this paper, we describe an approach to developing an ecologically valid framework for performing automated cognitive assessment. To automate assessment, we use a machine learning approach that builds a model of cognitive health based on observations of activity performance and uses lab-based assessment to provide ground truth for training and testing the learning algorithm. To evaluate our approach, we recruited older adults to perform a set of activities in our smart home test-bed. While participants perform activities, sensors placed in the smart home unobtrusively capture the progress of the activity. During analysis, we extract features that indicate how well participants perform the activities. Our machine-learning algorithm accepts these features as input and outputs the cognitive status of the participants as belonging to one of two groups: Cognitively healthy or Dementia. We conclude that machine-learning algorithms can distinguish between cognitively healthy older adults and older adults with dementia given adequate features that represent how well they have performed the activity.

Categories and Subject Descriptors
I.2.6 [Artificial Intelligence]: Learning

General Terms
Algorithms

Keywords
Cognitive Assessment, Ensemble Learning, Smart Home, Activities of Daily Living

1. Introduction
Alzheimer's is the sixth-leading cause of death in the USA [1]. An older adult having such cognitive impairment loses independence in life and may require institutionalization for care and support [2]. Thus, a huge amount of money is being spent for the care-giving and support of patients experiencing such impairment. Detection of such cognitive impairment is difficult [3]. Research shows that 75% of dementia and early dementia cases go unnoticed [3] and many such cases are only diagnosed when such impairment reaches moderate or advanced stage.

The majority of current diagnosis methods for cognitive impairment are based on caregiver report and laboratory assessment [4]. Current clinical approaches to assessing change in everyday functioning are limited by numerous factors, including restricted behavior sampling [14] and collection of data outside the home environment in a laboratory or clinic [15]. By tracking trends in everyday activity performance in a person's home environment on a daily basis and comparing the findings with clinical data, we may be able to identify new markers that signal acute health care changes as well contribute a missing, fundamental element to our knowledge concerning the natural history of functional change between healthy aging and dementia. Such contributions could lead to new innovative, proactive (as opposed to reactive) health care interventions. These contributions could also lead to the development of more ecologically valid measures of everyday functional status. Current clinical methods which are commonly used as proxies for real-world functioning (e.g., self report, informant report, assessment of cognitive impairment, and performance-based measures) all have limitations [16]. Early detection of cognitive status also provides a patient with advantages because it provides them with time to prepare for the future and benefit from early treatment [1,5,6].

Researchers such as Marcotte et al. and Aretouli and Brandt [5,6] argue that assessment of everyday functioning is an ecologically valid way to measure cognitive decline while others, including Willis et al [17]., have investigated the relationship between cognitive decline in older adults and the ability to carry out activities of daily living. Smart home technologies facilitate quantitative analysis of one's ability to carry out tasks and everyday functioning at home. Various sensors placed in the smart home capture events related to the behavior of its residents. Our algorithms analyze sensor data collected in smart homes to perform cognitive assessment of the resident. Some of the advantages of smart home based assessment are smart home technology facilitates continuous assessment of residents. This assessment method is ecologically more valid because assessments are performed in an individual's own home as they perform everyday tasks. Unlike lab-based assessment, our approach does not require the presence of a trained physician.

3. Data Collection

To generate data, we recruited 79 older adults (mean age = 66.69 with a range of 50 to 87 years) belonging to one of two cognitive groups: cognitively healthy (n=65, mean age = 64.53, with a range of 50 to 87) and dementia (n=14, mean age= 76.71 with age range or 59 to 85). All participants diagnosed with dementia met DSM-IV criteria [21]. Out of 79 older adults, 29 were male and 50 were female. We asked participants to perform selected Instrumental Activities of Daily Living (IADLs) in our smart home. While they perform activities, various sensors placed in the smart home capture the progress of the activity to its completion.

3.1 Smart Home Test Bed

The CASAS smart home test-bed is a two-story apartment located on the Washington State University campus. The apartment contains a living room, a dining area and kitchen on the first floor and three bedrooms and a bathroom on the second floor. The rooms in the smart home are equipped with motion sensors on the ceiling, door sensors on cabinets and apartment doors and item sensors on selected kitchen items. The test bed also contains temperature sensors in each room and a power meter to measure electricity consumption. These generated sensor data is collected by CASAS middleware. Each sensor event is represented by four fields (Date, Time, Sensor ID, and Message). A separate research team member (who neither runs the experiment nor performs analysis of the algorithms) annotates the sensor data to relate the sensor event to the label of an activity that was performed while the sensor events were generated.

3.2 Activity List

Data is collected in collaboration with the Department of Psychology at Washington State University. The following is the list of activities that we requested participants to perform:

1. Sweeping the kitchen and dusting the living room.
2. Obtaining medicine containers and a weekly medicine dispenser, filling the dispenser according to the directions.
3. Writing a birthday card, enclosing a check and writing the address on an envelope.
4. Finding the appropriate DVD and watching the corresponding news clip.
5. Obtaining a watering can and watering all plants in the living space.
6. Answering the phone and responding to questions.
7. Preparing a cup of soup using the microwave.
8. Picking a complete outfit for an interview from a selection of clothing.

The above set of eight activities represents what is often known as Instrumental Activities of Daily Living (IADLs) [9]. In order for a person to live independently, a person needs to successfully complete these IADLs. However, completion of IADLs requires a high level of cognitive and functional ability. Further, researchers have shown that declining ability to perform IADLs is often related to some form of decline in cognitive ability, including difficulties with memory and executive functioning [4,15]. Because sensors can monitor the progress of IADLs, the data gathered from such sensors could also assess for changes in the functional and cognitive ability of a person.

3.3 Ground Truth

Ground truth data for a participant is generated from a comprehensive clinical assessment, which included neuropsychological testing data, interview with a knowledgeable informant, completion of the Clinical Dementia Rating [19,20] ,the Telephone Interview of Cognitive Status(TICS) [18] and review of medical records. All participants in this study completed a battery of standardized cognitive tests before initiating the smart home testing sessions. The lab-based neuropsychological tests measure a participant's cognitive abilities, including memory, language, attention, speeded processing and executive skills, relative to standardized norms for individuals of the same age and education level. Participants in the dementia group met DSM-IV criteria for dementia, which includes the presence of multiple cognitive deficits that negatively affect everyday functioning and represent a decline from a prior level of functioning. Dementia participants had a score of 0.5 (very mild dementia) or higher on the CDR. In addition, scores on the TICS, a cognitive status exam, ranged from 18 (moderately to severely impaired) to 29 (ambiguous range); M = 24 (mildly impaired), std = 3.71. Cognitively healthy older adults scored within 1.5 standard deviations of standardized test norms on all tests administered, had a CDR score of 0 (normal), and performed within the non-impaired range (> 32) on the TICS.

4. Machine Learning Framework

In this section, we describe our algorithmic approach to predicting the cognitive status of an individual given their performance of everyday activities in a smart home.

4.1 Feature Generation

In the first step of our approach, we extract various features from the smart home sensor data that represent how well participants performed each activity. Table 1 summarizes the 38 activity features that we defined as input for the machine-learning problem. These features in essence quantify behavioral characteristics of cognitively impaired people. A participant suffering from dementia may wander, get confused, commit frequent mistakes or may not complete the activity at all, among other errors. For instance, features such as activity duration and the time delay before the participant initiates the task may indicate slowness in completing the activity or difficulty/confusion in initiating the activity etc.

In the table, *a major sensor* is defined as the most frequent sensor that is triggered by all participants while performing the activity. A *related sensor* for an activity is most widely used sensor by participants while performing activities. All these features are either dependent on time or on sensor counts, and are independent of the features derived from the clinical analysis of the psychological tests. The training set for the learning algorithm has 38 such extracted features, which includes the number of times each of the 26 motion sensors was triggered.

4.2 Machine Learning Model

As our next step, we build learning models that predict the cognitive status of the person given a set of input features that represents the activity performance of the individual. Initially, we extract same set of features for each activity as explained in

Table 1. Feature descriptions for a single activity

Feature Name	**Feature Description**
Duration	Time taken to complete the activity
Sensor frequency	Total number of times each sensor was triggered
Age	Participant age
Unique sensors count	The number of unique sensors (out of 26) that were used for this activity
Instruction given	An indicator that experimenter help was given so as to clarify what was expected in order to complete the task.
Time to initiate task	The time delay between experimenter instructions and the beginning of activity performance
All Activities Completed	A Boolean feature that represents whether the participant was able to complete all eight activities
M01-M25,M051	A vector representing the number of times each motion sensor was triggered (there are 26 motion features)
Major sensor off	The most frequently-used sensor was triggered during this activity
Count unrelated sensor	Number of unrelated sensors that were triggered
Count unrelated sensor value	Number of unrelated sensor events
Door sensor Count	Number of interactions with various door sensors
Item sensor Count	Number of interactions with item sensors
Status	Status of the patient (Cognitively Healthy or Dementia)

Section 4.1 and train a machine-learning algorithm separately for each activity. This means that a dementia participant who completes all eight activities generates features for eight 'dementia-activities' that are used to learn eight separate predictive models. To construct a reliable predictor, we combine results from the individual learning algorithms for each activity. In addition, as seen from the list of activities that an individual performs, activities are fundamentally different from each other. For instance, the activity "preparing a cup of soup" is different from the activity "outfit selection". In this case, a classifier learns a mapping from the eight individual outputs to a label indicating the person's cognitive health (Cognitively Healthy or Dementia). Further, we use a meta-classifier Cost Sensitive Classifier to train our learning models as training examples have imbalanced class distribution. A cost sensitive classifier makes its base classifier cost sensitive based on a cost matrix. It takes misclassification costs for a class in account and reweights the training examples. A cost matrix provides such cost based on a confusion matrix.

5. Experimental Results

All of the experiments were performed with WEKA API [13] using leave-one-out cross fold validation and a cost sensitive learning algorithm [11]. We chose commonly used learning algorithms: Naïve Bayes, Decision Tree (J48), Sequential Minimal Optimization (SMO), and Neural Network. We chose cost sensitive versions of these algorithms as the class are imbalanced [12]. The main objective was to see if *learning algorithms could learn the class boundary between two classes cognitively healthy and dementia provided only with the sensor data*. The activity-sensor dataset contains sensor events for 65 cognitively healthy participants and 14 participants with dementia performing eight activities in a smart home test bed. Our learning algorithm was trained using ground truth obtained from the neuropsychological analysis (see Section 3.3). When results are combined with averaging or voting [7], the performance of the learning algorithm is better than when the algorithms are trained with a single training set and trained separately with different activity training data. As the result in Table 2 and 3 indicate, this represents a good start given the limited amount of training data. We can also see that the F-value for the cognitively healthy group is high as compared to the dementia group. The reason is likely the small number of training examples for people with dementia. We conclude that we can predict cognitive status from the sensor data given adequate features and training points that can measure how well an activity is performed.

The objective of this paper was to see if we can predict the cognitive status from sensor data with a machine learning model. We will consider possible enhancements such as using different learning algorithms for each activity and employing alternative cost matrices in our future work.

6. Discussion and Conclusions

In this work, we have discussed a framework to perform automated cognitive assessment of an individual by analyzing the individual's performance on activities of daily living in a smart home. We hypothesize that learning algorithms can identify features that represent situations such as mistakes, confusions, and wanderings that an individual having cognitive impairment frequently commits while performing activities. The results from our experiment suggest that learning algorithms can indeed differentiate between task performance for individuals in these groups, although differences in how tasks are performed Furthermore, we need to note that activities are fundamentally different from each other and have varying levels of difficulty. Thus, performance of a participant on an activity represents a score in one dimension and performance on eight different activities gives us eight distinct types of "scores." In addition, participants may perform a few activities in a "normal" way while they commit frequent mistakes on other activities, whereas, some participants fail to complete some activities at all. Thus, while our current algorithm gives equal weight to all eight activities, future work needs to address the question of how prediction for an activity can be weighted. The current problem is formulated as a supervised learning problem in which the task is to predict the cognitive status of a person from activity sensor data based on the ground truth obtained from a psychological analysis.

Table 2: Voting -Results by combining the Prediction from Individual Classifier

Learning Algorithm	AUC	F-value		GMean	Acc(%)
		Class-CH	Class-Dem		
Naïve Bayes	0.73	0.93	0.60	0.69	88.63
J48	0.80	0.91	0.64	0.79	86.07
SMO	0.77	0.91	0.62	0.76	86.07
Neural Network	0.73	0.93	0.60	0.69	88.60

*Dem = Dementia *CH= Cognitively Healthy

Table 3: Averaging Results by combining the Prediction from Individual Activity

Learning Algorithm	AUC	F-value		GMean	Acc(%)
		Class-CH	Class-Dem		
Naïve Bayes	0.86	0.93	0.64	0.73	88.60
J48	0.80	0.91	0.64	0.79	86.07
SMO	0.84	0.89	0.60	0.78	83.54
Neural Network	0.86	0.93	0.64	0.73	88.60

*****Dem = Dementia *CH= Cognitively Healthy

In future work, we would like to determine if an unsupervised learning algorithm can find patterns for different cognitive groups that can be validated with this ground truth. We also face the challenge of adhering to a ground truth that may itself be error prone. We can in the future design techniques to assess the consistency of ground truth labels. Finally, this is our preliminary work in introducing functional assessment techniques for smart homes. We have not addressed subtle issues, such as the sub-activities and step errors, among others and we are only limited to resolution from a motion sensor, and item sensor. Further, we will also explore functional assessment techniques by introducing new activities that are more complicated, interweaved and concurrent. An additional expansion to our work would be to introduce real time assessment in streaming data. We want to detect when the cognitive health of an older adult declines and detect early symptoms of dementia and Alzheimer's disease.

8. Acknowledgements

We want to thank Chad Sanders, Alyssa Hulbert and Jennifer Walker for their assistance in coordinating data collection.

This material is based upon work supported by the National Science Foundation under Grant No. 0852172, by NIBIB under Grant No. R01EB009675, and by the Life Sciences Discovery Fund.

References

[1] http://www.alz.org/alzheimers_disease_facts_and_figures.asp

[2] Gaugler J. E., Kane R. L., Kane R. A. , Newcomer R. 2005. Early community based service utilization and its effects on institutionalization in dementia caregiving. *In Gerontologist*, 45:177-185.

[3] Hodges M. R.,Kirsch K., Newman M. W., Pollack M. E. 2010. Automatic Assessment of Cognitive Impairment through Electronic Observation of Object Usage. *Pervasive 2010*: 192-209.

[4] Folstein M.F., Folstein S.E., McHugh P.R..1975. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatry Res.*

[5] Marcotte T. D., Scott J. C., Kamat R., and Heaton R. K., 2010. Neuropsychology and the prediction of everyday functioning in *The Guilford Press*, New York, 2010.

[6] Aretouli E. and Brandt J. 2010.Everyday functioning in mild cognitive impairment and its relationship with executive cognition. *In International Journal of Geriatric Psychiatry*, vol. 25, pp. 224-233.

[7] Dietterich T. G. 2000. Ensemble Methods in Machine Learning. In *Proceedings of the First International Workshop on Multiple Classifier Systems* (MCS '00), Josef Kittler and Fabio Roli (Eds.). Springer-Verlag, London, UK, 1-15.

[8] Liu X. Y. and Zhou Z. H.,2006, The Influence of Class Imbalance on Cost-Sensitive Learning: An Empiri*cal Study, Proc. Int'l Conf. Data Mining*, pp. 970-974

[9] Lawton M.P., Brody EM.1969.The Instrumental Activities of Daily Living Scale. In *The Gerontologist* 1969.

[10] Willis S. and Marsiske M. 1993. Manual for the Everyday Problems Test, Pennsylvania State University Department of Human Development and Family Studies, University Park, Pennsylvania.

[11] Ethem A. 2005. Introduction to Machine Learning by, MIT Press.

[12] Yanmin Sun 2007. Cost-sensitive boosting for classification of imbalanced data, University of Waterloo, Waterloo, Ont., Canada.

[13] Mark H., Eibe F., Geoffrey H., Bernhard P., Peter R., Witten I. H. (2009); The WEKA Data Mining Software: An Update; *SIGKDD Explorations*, Volume 11, Issue 1.

[14] Marcotte, T. D, Scott J. C., Kamat R., and Heaton, R. K.2010. Neuropsychology and the prediction of everyday functioning. In *Neuropsychology of everyday functioning*, pages 5-38. New York: The Guilford Press

[15] Diehl M.,Marsiske M., Horgas A., Rosenberg A., Saczynski J., and Willis S., 2005. The revised observed tasks of daily living: A performance-based assessment of everyday problem solving in older adults. *Journal of Applied Gerontology*, 24:211-230

[16] B. Wilson. 1993. Ecological validity of neuropsychological assessment: Do neuropsychological indexes predict performance in everyday activities? *Applied and Preventive Psychology*, 2(4):209-215.

[17] Willis S. L., Jay G. M., Diehl M., and Marsiske M., 1992, Longitudinal change and prediction of everyday task competence in the elderly. *Research on Aging*, vol. 14, pp. 68-91.

[18] Brandt, J. and Folstein, M. 2003. Telephone Interview for Cognitive Status. Lutz, Florida: Psychological Assessment Resources, Inc.

[19] Hughes, C. P., Berg, L., Danzinger, W. L., Coben, L. A., & Martin, R. L. 1982. A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, 140, 566-572.

[20] Morris, J. C. 1993. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, 43, 2412-2414.

[21] American Psychiatric Association. 2000. Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington,DC.