Identifying Varying Health States in Smart Home Sensor Data:
An Expert-Guided Approach

Roschelle L. Fritz
College of Nursing, Washington State University, Vancouver
14204 NE Salmon Creek Avenue, VLIB210E
Vancouver, WA 98686, USA

and

Diane J. Cook
School of Engineering and Computer Science, Washington State University, Pullman
EME 121 Spokane Street, Box 642752
Pullman, WA 99164-2752, USA

Abstract

The aging population is growing and innovative solutions are needed to address older adults' complex health needs while concurrently extending the reach of the nurse. One emerging solution is the health-assistive smart home. The smart home uses ambient sensors to monitor the movement of older adults and intelligent algorithms to detect changes in health states. Alerts are provided to patients, family and caregivers so older adults can receive timely interventions. Adding a clinician-in-the-loop when training machine learning algorithms may improve the machines ability to accurately identify and predict changes in health states that have clinical relevance. At Washington State University, the CASAS team uses a clinical nurse-expert in a guided approach to machine learning. Here, we describe the expert guided approach, discuss current challenges and offer suggestions for future machine learning research in the area of health-assistive smart homes.

Key Words: Smart Homes, Ambient Motion Sensors, Monitoring, Health, Older Adults

I. Introduction

The older adult population worldwide is increasing at a rapid rate[1] and 80% of adults age 65 and over have been diagnosed with one or more chronic diseases[2]. The complexity of providing excellent care for an older and sicker population while concurrently experiencing a decreasing number of caregivers and nursing professionals is a significant humanitarian concern. Innovative solutions that extend the reach of caregivers and nurses are needed[3]. One such solution is the health-assistive smart home, hereafter referred to as Smart Home.

The Smart Home is designed for use by independent older adults to assist with management of their chronic illness. These older adults may be under the care of a nurse and ambient sensor readings may inform nursing interventions. However, few nurses understand how the Smart Home works and even fewer are involved in its development. Here, we describe the integral role played by the nurse on our multidisciplinary research team. We reveal a step-by-step process for infusing clinical knowledge into machine learning algorithms, which we call a *clinician-in-the-loop approach to machine learning*. We believe this article addresses a major knowledge gap regarding the inclusion of clinical expertise in high-tech health-related machine learning development and design.

A. The Smart Home

In this article, we refer to a specific health-assistive Smart Home, namely Washington State University's (WSU) Center for Advanced Studies in Adaptive Systems (CASAS) Smart Home in a Box. An interdisciplinary team of engineers, computer scientists, psychologists, and nurses are developing this Smart Home to assist older adults with aging in place. The Smart Home has two main components: (a) hardware and (b) software. Hardware components include sensors (motion, temperature, humidity, light, contact), Wi-Fi relays, and computer servers. The sensors detect and transmit their readings. The data is collected by our middleware which adds a sensor identifier and a date/time stamp to each reading. Machine learning algorithms analyze sensor data and identify behavioral patterns. The Smart

Home is currently capable of identifying more than 40 normal activities of daily living (ADLs) with greater than 98% accuracy[4].

There is a growing body of work regarding the relationship between sensor data patterns and specific health states but there is little information regarding how clinical knowledge can be infused into algorithms that learn this relationship. An important role for the Smart Home is accurate detection of health states, otherwise health outcomes may be less than optimal. Here, nurses can be an irreplaceable wealth of knowledge. Including a clinician into the loop of collecting, analyzing, labeling, and learning data patterns can optimize accurate training of machine learning algorithms.
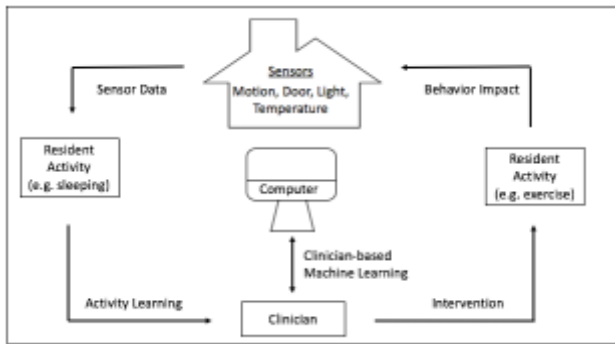

Figure 1. The Clinician-in-the-Loop Smart Environment. Sensors detect movement, activities are labeled by intelligent algorithms, and a clinician reviews raw data and relays interpretations to engineers so accurate clinical interventions can be activated for extending independence.

A major goal is to train the Smart Home to identify baseline health states and to detect changing conditions that have clinical relevance. To explore our ability to detect changes in health states we deployed 5 smart homes to a continuing care retirement community in Washington State. In this ongoing study, we concurrently monitor older adults who have two or more chronic conditions with smart home sensors *and* with weekly clinical nursing health assessments. Our nurse was trained to systematically review raw sensor data looking for sensor data patterns that likely regard clinically relevant chronic illness exacerbations in 5 older adults. To do this, she analyzed information from: (a) participants' medical records, (b) weekly nursing assessments, (c) ambient motion sensor data, and (d) information found in the extant smart home literature.

B. Detecting Health States

In the earliest stages of training, nurses are taught to observe changes in movement as they relate to disease states[5]. This training is an ingrained nursing skill causing nurses to be continuously cognizant of human movement. Almost every disease or condition impacts human motion. This motion is important to overall Smart Home-based health interpretation. The Smart Home team at WSU detected a variety of conditions including insomnia, falls, and a side effect of radiation treatment[6], as well as variations in cognitive decline[7]. Rantz and colleagues also detected changes in health states related to pneumonia, upper respiratory infections, post hospitalization pain and more[8]. For a recent review of the literature on smart homes and monitoring technologies see Liu et al.[9] For a discussion of our quantitative methods used with the qualitative methods discussed here see Sprint et al.[6]

II. Methods
A. Clinician-in-the-Loop Analytic Process
In this section, we introduce our clinician-in-the-loop approaches to training smart homes for detecting health events. The analytic process and interpretation of data relies heavily on our expert's background as both a qualitative methodologist and 25-year careered nurse. In this work, we particularly focus on the chronically ill older adult's lived experience as it relates to moving through space over time, and the capturing of that experience with quantifiable motion sensor data. In this process, we invented a neoteric method of discovery whereby the expert applies qualitative analytic methods to quantitative data in order to identify common patterns and themes of motion across multiple sets of raw motion-sensor data. These data sets are the expert's transcripts. The expert clinician analyzes all of an individual's transcripts and compares multiple individuals' transcripts. We seek common themes for capturing motion as it regards disease. The clinical expert's primary role is to identify, label and describe data that likely represents a change in health state.

Several components of information are needed to analyze sensor data using the following 5-step analytic process. They are: (a) information on baseline health, (b) sensor data that correspond with the dates under investigation, (c) a floor plan for identifying the location and type of sensor by room, and (d) a record of sensor data reliability by date and time so interpretations are only inferred from reliable data.

Step 1. Organize health information. We collect as much health information as possible on the participant including: the medical record with history and physical, nursing physical assessment(s), recent

history, list of all medications, functional status (e.g., use of assistive devices), biometrics, and record of mental health.

Step 2. Identify measures in literature. For each participant, our expert conducts a review of the sensor literature looking for previously identified sensor measures associated with the identified diagnosis. Examples of measures found in the literature are walking speed, sleep/wake cycles, entering/exiting the home, and ADLs. If a particular motion pattern has already been recorded in the literature regarding the identified diagnosis this is recorded as a 'known' measure. Known measures are applied when the expert reviews the participant's raw sensor data. Currently there are few known motion sensor measures recorded related to particular diagnoses or conditions.

In addition to measures identified in the literature, newly identified measures discovered as part of this research are documented. Once the measure is deemed stable it is used in all future analysis and actively informs interpretations. Stable measures are those encountered across a minimum of 3 transcripts, 5 or more times total, consistently exhibiting similar attributes for each occurrence (i.e., the same cluster of sensors activated about the same time of day related to the same health state), and consistently informing understandings of motion patterns (e.g., bed time routines).

Step 3. Identify a change in health state. This is done by reviewing the spreadsheet of organized health data (i.e., from Step 1). The spreadsheet includes reports of changes in health states (e.g., a fall). We pay attention to symptom complaints, medications changes, and reported health events. Once a change has been identified and its associated sensor data located by date and time, the data is set aside for pending review.

Step 4. Verify reliability of raw data. Concurrent variables can impact reliability of the sensor data at the time the change in health state occurred. Reliability issues occur when there is poor internet connectivity, the sensor batteries are low, or the Wi-Fi relay or server box becomes disconnected. Noisy data can also be unreliable; for example, when visitors are in the home. Unreliable data should be disregarded.

Step 5. Conduct a systematic and scoped review of data. The expert initiates a review of raw sensor data on the day of the identified change in health state. Experiential clinical knowledge is applied as well as the measures identified in the literature to look for motion patterns that regard the participant's

current condition. Careful review of the entire 24 hours around the change in health state is important in determining exactly 'where' in the data the change occurred. A floorplan that is annotated with sensor types and locations is used to create a mental image of the older adult's movement around their home on the identified day. Concurrently and in circular fashion, the expert considers any previously acquired knowledge regarding the older adult's routines (e.g. wake time, bedtime, bathroom use). Once the change is identified, comparisons are then made between normal motion patterns and abnormal motion patterns that occurred on the day the event occurred.

Next, the expert identifies the individual's baseline patterns. To discover what these patterns are, a scoping review is conducted of all sensor data spanning 48 hours of data on each side of the identified date. Then, the review is extended to 1 week, or longer if further review of nearby dates is needed to find baseline. From there, 5 random dates are chosen for comparison at one month, and again at 2-5 months and at 6-12 months from the noted change in health state. If needed, the review of data is further extended on each of the randomly chosen dates beyond 24 hours of the date to gain a better understanding of routines at that time. Dates are chosen randomly while concurrently considering: (a) equal distribution of time of day (morning, mid-day, evening, night) and day of week (weekday, weekend, or holiday), (b) appropriateness of time distance from event (e.g., onset and duration of symptoms), (c) interfering factors creating noise in the data (e.g., visitors known to be in the home at that time, other known health events), and (d) reliability of data. Effort is given to finding contrasting patterns across time; both similarities to normal patterns and pattern anomalies. If, upon choosing a random date, a pattern abnormal to the emerging understanding of baseline is noted, that comparison is documented and the date associated with the abnormal data is labeled as abnormal for future reference. That date is subsequently disregarded and another date is chosen for review.

Throughout the scoping review the clinical expert documents what is observed as baseline routines and any abnormal motion patterns. These descriptions include information regarding patterned norms by time of day (e.g., morning versus evening) and by type and length of activity, as well as activity repetition. Abnormal patterns described in comparison to baseline are analyzed for their potential clinical relevance. For example, data indicating that an older adult with congestive heart

failure (CHF) has changed from sleeping in his bed every night to sleeping in the recliner chair has clinical relevance and would be an important finding. Heavily informing qualitative interpretation of data is an understanding of individual as well as clinically anticipated changes.

During this stage of analysis, it is important to consider the different types of sensors that are activated. Discrimination between data derived from area sensors (motion sensors that monitor a large room or area) versus direct sensors (motion sensors that monitor a focused area with a 2-meter diameter) informs interpretation of motion patterns. Area sensors are sensitive to motion across broad areas of a room while direct sensors are only sensitive to motion *directly beneath* them. Activation of different types of sensors, either solely or in combination, inform understandings of movement.

B. Interdisciplinary Reporting of Findings

It is important to report abnormal data patterns in a way that is meaningful to scientists in both the fields of engineering and nursing. After exploring several communication formats with regard to sensor data and health events, we concluded that a spreadsheet was most effective. On this spreadsheet, we organized combined health events and their associated sensor activations. Specifically, we labeled the change in health state (e.g., shortness of breath) and the corresponding participant's coded identifier, diagnoses (associated with the change in health state), date and time of change, time of day associated with change (day or night), the measure used to identify the event in the data (e.g. sleep), routine motion patterns, the change in motion patterns, associated lines of sensor data, and relevant clinical and contextual comments. Changes in health states were sorted chronologically. This communication process proved convenient for engineers. It also allowed clinical organization of thought using normed methods of synthesizing health-related information within the field of nursing while providing situational context, a traditionally valued and informative component of nursing communication.

C. Exemplar (Restless Legs Syndrome)

An 89-year-old female living alone and independent complained of not sleeping because her legs were restless. She had a medical diagnosis of Restless Legs Syndrome. On the night of January 14, 2017 she reported that she only slept for two sessions of less than one hour. Upon examining the raw sensor data, the clinician noted the bedroom bed sensor, a direct sensor placed directly above the bed and labeled BedroomABed, rapidly fired all night long. This sensor activation was periodically interspersed with activation of the bedroom area sensor. There were only a few times where all sensors in the bedroom were silent. Beginning about 2200 and throughout most of the night, the sensor data appeared to indicate her body was nearly continuously moving while in bed. The direct sensor over the bed and the bedroom area sensor were intermixed and were both activated multiple times a second. From this pattern of sensor activation, the clinical expert inferred that the participant was in bed (as opposed to anywhere else in the house, no other sensors in the home were active). Also inferred was that the participant was moving around in bed with enough movement to cause activation of both the direct sensor located over the bed and the bedroom area sensor.

Table 1. Sensor Activations, No Sleep

Rest Without Sleep
2017-01-14 22:44:50.652739 BedroomABed OFF
2017-01-14 22:44:51.202735 BedroomABed ON
2017-01-14 22:44:52.323667 BedroomABed OFF
2017-01-14 22:44:57.946372 BedroomABed ON
2017-01-14 22:44:59.072191 BedroomABed OFF
2017-01-14 22:44:59.631225 BedroomABed ON
2017-01-14 22:45:00.764955 BedroomABed OFF
2017-01-14 22:45:01.323373 BedroomABed ON
2017-01-14 22:45:03.006263 BedroomABed OFF

This information can be contrasted with the following data. Note the data has a distinct break in time of sensor activations. (See bolded text.)

Table 2. Sensor Activations, Sleep

Rest With Sleep
2017-01-15 01:52:18.314491 BedroomABed ON
2017-01-15 01:52:19.438018 BedroomABed OFF
2017-01-15 01:52:20.559191 BedroomABed ON
2017-01-15 01:52:22.246756 BedroomABed OFF
2017-01-15 03:21:47.748431 BedroomABed ON
2017-01-15 03:21:47.888572 BedroomAArea ON
2017-01-15 03:21:49.018683 BedroomAArea OFF

Relevant data not shown here was the pattern of sensor activations just before and after the break, which includes an intermixed rapid firing of direct and area bedroom sensors indicating rather continuous physical body movement. The logical assumption is that this break in time, resulting from a

lack of sensor activation, indicates a period of time in which the legs relaxed and the participant was able to sleep. This can be inferred because neither the direct sensors, nor the area sensors, were activated; meaning there was no movement of the physical body at that time. The clinical expert retrospectively verified in the nursing assessment notes that the absent bedroom sensor activations aligned with the approximate time of night sleep occurred, length of time she slept, and her location (i.e., bedroom bed) and the event did match. The interpretation was further validated while interviewing the participant during the next weekly nursing visit.

III. Discussion

In this section, we discuss the characteristics of movement, highlight the benefits of the clinician-in-the-loop model, and identify areas that need further exploration.

A. Motion Patterns

Patterns of motion are detected in the Smart Home by observing common sequences of sensor activations. The detected patterns are noteworthy if they occur at equivalent times of day (e.g., morning, evening, night) or in relation to clustered activations occurring in relation to the event (i.e., just before or after the data-cluster of interest). For example, in our study we have one participant who routinely goes to the kitchen sink for a glass of water immediately upon arising every morning while another participant uses a specific bathroom toilet upon arising. In each scenario, for the respective older adult, the same cluster of sensors are activated about the same time of day and for about the same length of time. Furthermore, on either side of the cluster of sensor activations a similar data-cluster pattern existed. These patterns inform our understandings of each participant's overall patterns of motion.

Motion patterns have unique sensor characteristics that appear to correlate to changes in health states. Motion pattern characteristics deemed valuable to our research are: (a) timed movement from one location to another (bed to toilet), (b) timed-activity comparisons by percentage, (c) sensor activation combinations, and (d) bookends. Here, we further describe these characteristics.

Timed-activity comparisons are determined by looking at the length of time a cluster of sensors remain continuously activated. Comparing the amount of time a particular cluster of sensors remains activated against the quietness of those same sensors can inform understandings regarding health. The summative time spent in bed, or combined time spent in bed and recliner, changes as health declines because there is an increased need for the body to rest. This change can be detected with ambient sensors and is meaningful clinical information.

Sensor activation combinations are represented in data where 2 or 3 sensors are continuously activating in back-to-back fashion. A meaningful combination is the consecutive activation of area sensors WITHOUT interspersed activations of direct sensors, which occurs over a period of time and can be variable (e.g. BedroomAArea – BedroomAArea – BedroomAArea). Another meaningful combination is seen in the Restless Leg Syndrome scenario where intermixed sensors were activated (i.e. BedroomABed and BedroomAArea). In that scenario, no other sensors were activated for a period of time. In a scenario where a fall occurs under a direct sensor then a combination of direct and area sensors will be noted and the length of time in that location with that particular sensor combination will be abnormal.

The term 'bookends' regards a section of abnormal data found between larger sections of normal data; the abnormal data has normal data bookends. These bookends are clusters of data commonly occurring on either side of a health event. During qualitative analytic review a common theme emerged. This theme was the bookending of event data by clusters of intermixed data from multiple other sensors. The data that bookended the 'data in the middle' often included sensor activations from 4 or more various sensors in 2 different rooms of the home occurring within 15-30 minutes before and after the event. This, in contrast to the 'data in the middle' which exhibited back-to-back activation of a single sensor (e.g., BedroomAArea), or the activation of a narrow cluster of sensors such as direct and area sensors within the same room (e.g., BedroomABed and BedroomAArea).

B. Future Considerations

The benefit of having a clinician-in-the-loop is multifaceted. Most importantly, the clinician facilitates the Smart Home's ability to recognize clinical changes in health states *as a nurse would*. This will likely improve health outcomes of patients using the Smart Home.

A major limitation of this work is the novel nature of this qualitative approach to analyzing massive amounts of quantitative data. There is a lack of information about how to identify health states in low fidelity ambient sensor data. Our work is

exploratory in nature and as such, knowledge regarding its efficacy is low. Clinicians and engineers should continue to collaborate in this arena.

Interdisciplinary communication remains important and future research should address additional ways for clinicians to engage with engineers in development of health-assistive technologies. There is a need to identify additional measures of motion. Adoption by older adults also needs further exploration.

IV. Conclusion

Health-assistive smart homes that can detect and predict changes in health states are under development. In this article, we discussed the importance of including clinical knowledge when training machine learning algorithms and we shared a 5-step process used by a nurse to infuse clinical knowledge as the 'clinician-in-the-loop' for an expert guided approach to machine learning. We also discussed the limitations of our neoteric qualitative analytic approach.

REFERENCES

[1] National Institute on Aging. "Global Health and Aging," 2016. Retrieved May 10, 2017 from https://www.nia.nih.gov/research/publication/global-health-and-aging/preface

[2] National Council on Aging. "Healthy Aging Facts: Chronic Disease," 2016. Retrieved May 10, 2017 from https://www.ncoa.org/news/resources-for-reporters/get-the-facts/healthy-aging-facts/

[3] Office of Disease Prevention and Health Promotion. "Healthy People 2020," *Healthy People.*gov, 2017.

[4] D.J. Cook, A.S. Crandall, B.L. Thomas, N.C. Krishnan, "CASAS: A Smart Home in a Box," *Computer*, vol. 46, no. 7, pp.62-69, Jul 2015

[5] J.M. Wilkinson, L.S. Treas. *Fundamentals of Nursing, 3rd Ed.,* vol. 3. Philadelphia: F.A. Davis.

[6] G. Sprint, D.J. Cook, R. Fritz, and M. Schmitter-Edgecombe "Using Smart Homes to Detect and Analyze Health Events." *Computer*, *49*(11), pp.29-37, Nov 2016

[7] M. Schmitter-Edgecombe, D.J. Cook, A. Weakley, P. Dawadi [in R.L. Kane and T.D. Parsons, eds., 2017.] *The Role of Technology in Clinical Neuropsychology*. Oxford University Press.

[8] M. Rantz, M. Skubic, M. Popescu, C. Galambos, R.J. Koopman, G.L. Alexander, L.J.Phillips, K. Musterman, J. Back, S.J. Miller. "A new paradigm of technology enabled 'vital signs' for early detection of health change for older adults" Gerontology, vol 61, no. 3, pp. 281-290, 2015.

[9] L. Liu, E. Stroulia, I. Nikolaidis, A. Miguel-Cruz, A.R. Rincon. "Systematic Homes and Home Health Monitoring Technologies for Older Adults: A Systematic Review," *International Journal of Medical Informatics,* vol 91, pp. 44-59, Apr 2016.