Automated Health Alerts Using In-Home Sensor Data for Embedded Health Assessment

MARJORIE SKUBIC1, (Senior Member, IEEE), RAINER DANE GUEVARA1,2, AND MARILYN RANTZ3, (Member, IEEE)

1Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO 65211 USA
2Cerner Corporation, Kansas City, MO 64117 USA
3Sinclair School of Nursing, University of Missouri, Columbia, MO 65211 USA

CORRESPONDING AUTHOR: M. SKUBIC (skubicm@missouri.edu)

The work was supported in part by the U.S. National Science Foundation under Grant IIS-0428420 and in part by the National Institutes of Health under Grant 1R01NR014255 and Grant 1R21NR011197-01.

ABSTRACT We present an example of unobtrusive, continuous monitoring in the home for the purpose of assessing early health changes. Sensors embedded in the environment capture behavior and activity patterns. Changes in patterns are detected as potential signs of changing health. We first present results of a preliminary study investigating 22 features extracted from in-home sensor data. A 1-D alert algorithm was then implemented to generate health alerts to clinicians in a senior housing facility. Clinicians analyze each alert and provide a rating on the clinical relevance. These ratings are then used as ground truth for training and testing classifiers. Here, we present the methodology for four classification approaches that fuse multisensor data. Results are shown using embedded sensor data and health alert ratings collected on 21 seniors over nine months. The best results show similar performance for two techniques, where one approach uses only domain knowledge and the second uses supervised learning for training. Finally, we propose a health change detection model based on these results and clinical expertise. The system of in-home sensors and algorithms for automated health alerts provides a method for detecting health problems very early so that early treatment is possible. This method of passive in-home sensing alleviates compliance issues.

INDEX TERMS Behavioral bio-markers, eldercare monitoring, health alerts, in-home sensing.

I. INTRODUCTION

Our view of embedded health assessment is the on-going assessment of health changes based on an individual’s behavior and activity patterns and baseline health conditions. Sensors embedded in the environment are used to collect behavior and activity patterns for the purpose of detecting health changes. Early detection is the key to promoting health, independence, and function as people age [1], [2]. Identifying and assessing problems early, while they are still small, provides a window of opportunity for interventions to alleviate problems before they become catastrophic. Older adults will benefit from early detection and recognition of small changes in health conditions and get help early when treatment is the most effective. Most importantly, function can be restored so they can continue living independently.

Recently, there has been an increased focus on technology for enabling independent living and healthy aging. A major challenge for studies in this area is the capture of ground truth data sufficient for training and testing purposes. For example, students have been enlisted to act out activities of daily living (ADLs) to create labeled data sets, e.g., for studying statistical activity recognition methods [3], [4]. Other work has used smaller datasets from a few volunteers, such as the statistical predictive algorithm to model circadian activity rhythms [5], mixture model analysis to infer activities of one user, validated with a manual log [6], and fuzzy rules used to classify activities in the home [7]. Although progress continues, the difficulties associated with collecting longitudinal sensor data along with real health data of subjects have hindered studies on embedded health assessment.

In this paper, we present an example of unobtrusive, continuous monitoring in the home for the purpose of embedded health assessment, to address the management of chronic conditions as people age. An embedded sensor
network collects data on behavior and activity patterns. A one-dimensional (1-D) alert algorithm is used to generate health alerts to clinicians in a senior housing facility. Clinicians analyze each alert using an electronic health record (EHR) and an interactive web interface for visualizing the sensor data. Based on their clinical expertise, they rate the clinical relevance of the alert. Here, we use the ratings as ground truth labels for health changes and test multi-dimensional approaches for classifying alerts as good or poor.

The approach is tested in a senior housing site called TigerPlace, which comprises independent living apartments that support seniors in the same apartment through the end of life. Nearly all residents have at least one chronic health condition. Over half have multiple chronic conditions, which include ailments such as arthritis, heart disease, diabetes, and hypertension. None of the study participants had significant cognitive impairments.

In Section II, we include a review of related work. Section III describes the sensor network used in this study, an initial investigation of feature space using the in-home sensors for detecting health decline, and the automated health alert process used for testing the embedded health assessment concept. Section IV describes four health alert classification methods tested; Section V presents the results. The discussion in Section VI includes advantages and limitations of the health alert system. Based on the study results, as well as our clinical experience in using the sensors, we propose a model for detecting health decline using continuous in-home monitoring. We conclude in Section VII.

II. RELATED WORK

The variety of related work shows the interest and potential of embedded health assessment. Both daytime and night time activity have been investigated using in-home sensors. For example, passive infrared (PIR) motion sensors have been used to capture activity in a particular location in the home [8], [9]. The pattern of room to room activity has been studied as a means of investigating health changes [5], [6], [10], [11]. Motion density from PIR motion sensors (i.e., number of events per unit time) can capture overall activity level that may be linked to health condition [12]–[14]. ADLs have been captured by a variety of sensor types including motion sensors, discrete switches, powerline monitoring, and computer vision [3], [4], [15]–[18]. In addition, sleep patterns have been studied using motion sensors [19], [20], bed mats [21], [22], or load cells [23], [24]. Other work has focused on the detection of cognitive changes, using a combination of motion, bed and door sensing, medication tracking, use of a home computer, cognitive computer games for monitoring and remediation, and a phone sensor for detecting incoming and outgoing calls [8], [25]–[27].

Walking gait in the home is of interest because of the link to both physical and cognitive health [28]–[32]. Walking speed has been captured using motion sensors [33], [34], video [35], [36], radar [37], and depth images [38], [39]. In addition, significant effort using in-home sensing has focused on safety, e.g., fall detection [40]–[46] and wandering [47], [48].

One challenge for researchers is collecting enough in-home sensor data to establish correlations with clinical assessments, although there is some recent work in this area [49]–[51]. Much of the recent research is specifically focused on cognitive decline; there is less work on detecting early health changes for generalized health management. Also, it is important to identify the best parameters to track for health change; some parameters may be too late for very early health change detection. For example, tracking ADLs is important for determining whether independent living is possible. However, explicit ADL recognition may be too late in some cases for early interventions; e.g., changes in night time activity due to urinary tract infections have been detected before ADL changes appear [52].

Detecting early health change is further complicated by the variety of home configurations and health conditions. Many seniors have multiple chronic health conditions to manage, and the interaction may present changes in a variety of ways. In the work presented here, we address this variability by exploring behavioral bio-markers that generalize across different chronic health conditions and different home sizes and room layouts.

There is also work on wearable sensor networks and sensing incorporated into clothing. Some wearable systems have been studied extensively or advanced to the stage of commercialization, e.g., accelerometers for fall detection [42], [44], [53], counting footsteps [54], worn on the wrist as actigraphy sensors [55], [56], or assessment in clinics [57], [58]. These remain a challenge for continuous monitoring and explicit health change detection of seniors in naturalistic home settings. Many seniors refuse to wear them, forget to charge batteries, or are unable to operate them [59], [60]. Thus, they are not practical for continuous, long term monitoring. In addition, many wearable sensors are still at the prototype stage and not yet ready for long-term use in an unstructured home environment. Challenges still need to be addressed, such as unobtrusiveness, miniaturization, and robustness before they can be adopted widely [61].

III. SENSOR NETWORK

Our team has developed a sensor monitoring system for embedded health assessment in the home, as illustrated in Fig. 1. The focus is on detecting very early changes in health status for older adults who are managing chronic health conditions. Data from sensors installed in apartments are logged and stored on a secure server; data are logged by identifier number only as part of the IRB-approved research study. A typical installation for a one bedroom apartment as used in this study is shown in Fig. 2 with 11 motion sensors, a bed sensor, and a temperature sensor for capturing stove and oven activity.
Passive infrared (PIR) motion sensors are used to capture motion in a room area and also for localized activity, e.g., in the refrigerator, in kitchen cabinets, on the ceiling over the shower, and on the ceiling over the front door to detect apartment exits. The PIR motion sensors, which use the wireless X-10 protocol for data transmission, generate an event every seven seconds if there is continuous motion. This is used as an artifact to capture the general activity level in the home by computing a motion density as motion events per unit time. For example, a resident with a sedentary lifestyle may generate only 50 motion events per hour, whereas a resident with a very active life style may generate 400 or more motion events per hour [62].

A pneumatic bed sensor [21] is installed on the bed mattress and used to capture sleep patterns. The bed sensor generates events for restlessness in bed (four levels) as well as low, normal, and high events for pulse rate and respiration rate. For those residents who often sleep in a recliner chair, the bed sensor is installed in the chair. Sensor networks with motion, bed, chair and stove sensing have been deployed in TigerPlace apartments since 2005. Automated monitoring is used to detect the absence of sensor data, e.g., in the case of battery failures. However, there is still some data loss due to the brittleness of the X10 transmission.

### A. INVESTIGATING BEHAVIORAL FEATURE SPACE FOR CAPTURING HEALTH DECLINE

The in-home sensor data were investigated for use in detecting health decline through machine learning techniques coupled with input from the clinical research partners. Initially, clinicians provided a list of features that might be important for capturing health changes, related to behavior activity and sleep patterns. This set was further investigated using a systematic feature selection process retrospectively on data collected in senior apartments.

We did a broad investigation of 22 features extracted from motion and bed sensor data using a forward feature selection process. Case studies on two elderly residents were used for this initial investigation, where ground truth labels were generated manually by looking at the health records and nurses’ notes. Each day was labeled either as a normal day (a non-alert day) or an abnormal day (a day that would warrant a health alert). For this feature investigation, days with extended visitor activity, more than 12 hours out of the home, or significant sensor failures were removed from the datasets. After filtering these days, case study #1 included 267 normal days and 38 alert days; case study #2 included 567 normal days and 32 alert days.

Features extracted from motion sensor data included activity in the bathroom, bedroom, and living room, time out of the home, time with visitors, and overall activity level estimated from motion density [13], [63]. Kitchen activity was not included due to sensor failures. Also, TigerPlace residents seldom use their kitchens for preparing food due to the common dining facility; thus, kitchen features are not as important for this population. Features extracted from bed sensor data included time in bed, restlessness in bed, pulse events, and respiration events. All features were normalized for this investigation.

The feature selection process used a one class classification (OCC) method due to the class imbalance. Several OCC methods were tested, including the Support Vector Machine (SVM), nearest neighbor (NN-d), mixture of Gaussians, and Parzen density [63], [64]. The NN-d method showed more stable results overall; thus, we present the feature selection results using NN-d. OCC performance is based on the area under the Receiver Operating Characteristics (ROC) curve. For the forward search, each feature is first tested individually and the best feature is selected. Then, additional feature combinations are tested, and features are added one at a time until the
performance stabilizes or declines. The best combination is considered to be the one that first achieves the best classification performance.

Results of the forward search showed a different best combination for each case study. Table 1 shows the features tested and the best feature combination selected for each case [63]. Features common to both best sets include bathroom time, visitor time, and bed sensor events. In summary, bathroom visits, sleep patterns, and socialization activity were shown to be key in recognizing early health changes for these two residents. These results were used to inform the approach used for generating health alerts in the study described below.

<table>
<thead>
<tr>
<th>TABLE 1. Results of the forward feature selection process.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Features tested</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Day time motion density</td>
</tr>
<tr>
<td>Number of bathroom visits</td>
</tr>
<tr>
<td>Total time in bathroom</td>
</tr>
<tr>
<td>Bedroom events</td>
</tr>
<tr>
<td>Living room events</td>
</tr>
<tr>
<td>Number of times away from home</td>
</tr>
<tr>
<td>Total time away from home</td>
</tr>
<tr>
<td>Number of visitor events</td>
</tr>
<tr>
<td>Total time of visitor activity</td>
</tr>
<tr>
<td>Night time motion density</td>
</tr>
<tr>
<td>Number of times in bed</td>
</tr>
<tr>
<td>Total time in bed</td>
</tr>
<tr>
<td>Bed restlessness 1 events &lt; 3 sec</td>
</tr>
<tr>
<td>Bed restlessness 2 events 3-6 sec</td>
</tr>
<tr>
<td>Bed restlessness 3 events 6-9 sec</td>
</tr>
<tr>
<td>Bed restlessness 4 events &gt; 9 sec</td>
</tr>
<tr>
<td>Low pulse events &lt; 30/min</td>
</tr>
<tr>
<td>Normal pulse events</td>
</tr>
<tr>
<td>High pulse events &gt; 100/min</td>
</tr>
<tr>
<td>Low breathing events &lt; 6/min</td>
</tr>
<tr>
<td>Normal breathing events</td>
</tr>
<tr>
<td>High breathing events &gt; 30/min</td>
</tr>
<tr>
<td>Area under the ROC for best set</td>
</tr>
</tbody>
</table>

Retrospective analysis also revealed that the data form clusters in feature space. Our observation is that normal days tend to cluster, and abnormal days appear as outliers [63]. Fig. 3 shows the data for case study #2. There are still challenges in determining valid health alert days. Outliers may result from real health changes or may be due to "noise," such as a sensor failure, unusually long visitor activity, an extended absence, or an isolated abnormal day because, e.g., someone exercised excessively.

### B. GENERATING HEALTH ALERTS

To test the use of in-home sensor data for capturing health decline, an automated health alert system has been developed for prospective use. The logged sensor data are automatically analyzed on a daily basis, looking for changes in an individual’s data patterns. If a change is detected for the current day, an alert email is sent to clinicians [65].

The health alert email includes two web links. One is a link into the web portal which facilitates fast access to the sensor data visualization, showing a two week window of data before the alert and supporting an interactive interface for zooming out, drilling down, or displaying other parameters. This provides context to the clinician and helps determine whether the alert is relevant for this resident from a clinical perspective.

The second link provides access to a feedback web page that allows the clinician to rate the clinical relevance of the alert on a five point scale, from 1 (not clinically relevant) to 5 (very clinically relevant). This rating is then used as ground truth labels to aid in evaluation and further development of the alert algorithms.

On average, the clinician takes about two minutes to display the sensor data, analyze the alert, determine whether action is warranted, and provide feedback [66]. For the study reported here, 4 clinicians provided feedback (one physician and three nurses), based on their clinical expertise with older adults [66].

The alert algorithm was developed through retrospective analysis and collaboration with clinicians. Health events of senior residents with sensors were examined retrospectively, looking at emergency room visits, hospitalizations, and falls. A researcher manually reviewed sensor data leading up to the health events for signs of changes before the event occurred. Potential algorithms were then tested iteratively with a group of clinical researchers until consensus was reached. The approach looks at the sensor values per day, compared to a moving baseline of two weeks immediately before the day examined, i.e., relative sensor values are used rather than actual counts. Thresholds were set to increase the likelihood of detecting critical health changes even if it resulted in a high percentage of false alarms.

The two week moving baseline was chosen on the recommendation of clinicians after retrospective analysis of sensor
data patterns, as a compromise to capture both sudden and gradual health changes [66]. Each resident has a personalized normal that is reflected uniquely in the sensor data patterns, depending on health condition, usual lifestyle pattern, the size of the apartment, and the number of sensors. This strategy of change detection has facilitated the testing of health alerts even in a diverse group of seniors with varying levels of health and different chronic ailments.

**TABLE 2. Alert parameters and sensor data monitored for the alerts.**

<table>
<thead>
<tr>
<th>Alert parameter</th>
<th>Sensors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bathroom Activity</td>
<td>Sum of motion sensor events in the bathroom (bathroom, shower)</td>
</tr>
<tr>
<td>Bed Restlessness</td>
<td>Number of all bed restlessness events</td>
</tr>
<tr>
<td>Bed Breathing Low/Normal/High</td>
<td>Number of bed breathing low/normal/high events</td>
</tr>
<tr>
<td>Bed Pulse Low/Normal/High</td>
<td>Number of bed pulse low/normal/high events</td>
</tr>
<tr>
<td>Kitchen Activity</td>
<td>Sum of kitchen motion sensor (kitchen, fridge, etc.) events and stove/oven temperature high</td>
</tr>
<tr>
<td>Living Room Activity</td>
<td>Number of living room motion sensor events</td>
</tr>
</tbody>
</table>

Table 2 shows the alert parameters and sensor data monitored for the health alerts, chosen as a result of the collaborative analysis. For each parameter, the system computes a mean and standard deviation for the two week baseline window. If the current day’s values vary from the mean beyond a pre-determined number of standard deviations, an alert is generated. The standard deviation multiplier varies somewhat for different behaviors (from 2.5 to 4), according to the clinical research team’s view of the relative importance of the parameters. Relative changes are computed for three time periods: (1) a 24-hour day, midnight to midnight, (2) daytime, 8am to 8pm, and (3) nighttime, midnight to 6am. The email alerts generated include the parameter that caused the alert, the time period of the change, the direction of change (increase or decrease), and the number of standard deviations from mean (how big is the change).

Although both increasing and decreasing changes of all parameters are used to flag alerts, the clinical research team recognized that not all such changes would denote a health decline. For example, increasing living room activity during the day is not typically a cause for concern, while increasing living room activity during the night can indicate a change in health status. In contrast, any change in bathroom activity might indicate a health problem. For completeness in testing, all of these changing conditions are used for generating automated health alerts.

With this 1-D strategy, about half of the alerts generated are false alarms. Through manual investigation, it appears we are capturing nearly all of the obvious health changes.

We tracked changes in health status (emergency room visits, hospitalizations, and falls) and compared with alerts to investigate potential false negatives. Only ten false negative events were detected during a one year period, across 21 senior participants. However, other potential alerts might be missed. Nonetheless, capturing a clinical rating on the health alerts has allowed us to create a unique dataset for investigating more advanced algorithms for health alerts beyond this simple 1-D approach.

**IV. METHODS AND PROCEDURES**

In this section, we describe the classifiers investigated for determining whether a particular day’s sensor data should be classified as an alert day or not. The health alert ratings provided by the clinicians are used as ground truth labels in training and testing. Four classifiers were studied. The first is a fuzzy pattern tree (FPT) that does not require training but rather takes advantage of domain knowledge from our clinical partners. The remaining classifier methods rely on labeled training data and include the fuzzy K-nearest neighbor (FKNN), the neural network (NN), and the support vector machine (SVM). These classifiers were chosen for the study to provide a comparison between the use of domain knowledge only vs. trained classifiers that support a nonlinear decision boundary. The nonlinear decision boundary was deemed essential to match our observations about the normal vs. abnormal days as illustrated in Fig. 3.

**A. FEATURE SPACE**

In analyzing the health alerts generated for the parameters listed in Table 2, it was observed that some of the parameters do not typically cause alerts and others generate a few alerts but not enough to be used for supervised learning. Thus, for the study described here, we looked at the following four alert parameters: bathroom activity, bed restlessness, kitchen activity, and living room activity. If both increasing and decreasing changes (the current day’s count compared to the baseline period) are considered for all three time periods (daytime, night time, and full day), the dimensionality of the feature space is 24. We also investigated a 12-dimensional (12-D) feature space, considering only increases in each parameter. Fig. 4 shows a principal components analysis (PCA) reduction of the 24-dimensional (24-D) and 12-D feature spaces, where each blue × indicates a good alert day and each red o indicates a bad alert day (normal day); as shown, there is not good separation between the good alert and bad alert classes.

After further analysis of the alert ratings and discussion with our clinical partners, the feature space was reduced to consider the following six features: increasing nighttime activity in the living room, kitchen, and bathroom, increasing full day activity in the bathroom, and increasing bed restlessness at both nighttime and during the full day. A PCA reduction of this 6-dimensional (6-D) feature space
B. FUZZY PATTERN TREE

A fuzzy pattern tree (FPT) [67] was investigated as a method that uses domain knowledge and does not require training. The six features described above were combined in a FPT using an OR operator, providing a “rule” that is easy for clinicians to interpret [65]. Intuitively, the output is as follows:

- IF Bathroom activity for the full day is an Increase
- OR Bathroom activity at night time is an Increase
- OR Bed restlessness for the full day is an Increase
- OR Bed restlessness at night time is an Increase
- OR Kitchen activity at night time is an Increase
- OR Living room activity at night time is an Increase

THEN Alert is Clinically Relevant

Gaussian-based membership functions were used for the input parameters. The Yager t-conorm [68] was chosen as the OR operator to explore the additive combination of parameters as opposed to the standard maximum. That is, if small changes were observed in several parameters, these resulted in a cumulative effect in determining whether an alert was warranted. The Yager parameter $w$ sets the degree of optimism (how much greater the output is over the standard maximum operator) when two inputs are OR-ed together. For the work presented here, $w = 3$ generated the best classification empirically. For comparison, a 12-D FPT was also tested.

C. FUZZY K-NEAREST NEIGHBOR

A FKNN classifier was trained and tested with both 6-D and 12-D feature vectors (using Matlab functions). The standard Euclidean distance is used as a dissimilarity measure. For the crisp K-Nearest Neighbor (KNN) method, the classifier finds the K neighbors in the training set that are closest to the test vector. K is chosen as an odd-numbered value; the majority class of the K neighbors determines the class of the test vector. In contrast to the crisp KNN method, the FKNN produces a membership value in [0,1] that assigns partial membership in each class to the test vector. The class with the highest membership value is used here to determine the class of the test vector.

The method was tested using 10-fold cross validation with 90% of the data used for training and 10% used for testing in each fold; the cumulative results are reported. Each set of FKNN experiments varied K using 1, 3, 5, 7, 9, and 11. The best results are reported here, with K = 11.

D. NEURAL NETWORK

Two sets of 10-fold cross validation experiments were performed to test 6-D and 12-D NN classifiers (using Matlab functions). Both sets of NNs used 5 hidden nodes and one output node and the sigmoid activation function at each node. After each training sample, the NNs were updated using scaled gradient back propagation. Each fold randomly split the dataset into 70% training data, 15% validation data, and 15% testing data. The training data was used to update the NNs and after each epoch,
a sum of squared error (SSE) function was computed using the validation data. Training was terminated when the SSE increased for several epochs. Cumulative results on the cross-validation experiments are reported.

E. SUPPORT VECTOR MACHINE
The support vector machine (SVM) was also tested to investigate a supervised learning approach [65]. We investigated both a linear and radial basis function (RBF) kernel (using Matlab functions) and tested both the 6-D and 12-D features as described above. The RBF kernel performed better than the linear kernel; the RBF results are reported here. Again, 10-fold cross validation is used for the experiments, and the cumulative results are reported.

V. RESULTS
The four classifiers were tested with a set of health alert ratings spanning nine months on 21 senior residents. Table 3 shows the number of alerts used for each alert parameter that were rated as either good alerts or poor alerts. Good alerts included the alerts that were rated as a 4 or 5 (clinically relevant or very clinically relevant). The poor alerts included those rated as a 1 or 2 (not clinically relevant or less clinically relevant). The alerts rated as a 3 were interpreted as being neutral and were not included in this test.

TABLE 3. Health alerts used for testing (1-D).

<table>
<thead>
<tr>
<th>Alert parameter</th>
<th>Good alerts</th>
<th>Poor alerts</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bathroom Activity</td>
<td>42</td>
<td>43</td>
<td>85</td>
</tr>
<tr>
<td>Bed Restlessness</td>
<td>57</td>
<td>21</td>
<td>78</td>
</tr>
<tr>
<td>Kitchen Activity</td>
<td>7</td>
<td>63</td>
<td>70</td>
</tr>
<tr>
<td>Living Room Activity</td>
<td>17</td>
<td>85</td>
<td>102</td>
</tr>
<tr>
<td>Total</td>
<td>116</td>
<td>183</td>
<td>299</td>
</tr>
</tbody>
</table>

Table 4 shows the results of the four classification methods for both the 6-D and 12-D feature sets. The best classifications are achieved with the 6-D FPT and SVM and the 12-D FKNN. For comparison, we have included the ROC curves for the 6-D FPT and SVM classifiers, shown in Fig. 6. Although the accuracy rate is slightly higher for the FPT, the ROC curves show their classification performance is very similar. As shown in Table 3, the percentage of

TABLE 4. Classification results (in percentages).

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>FPT</th>
<th>FKNN K=11</th>
<th>NN</th>
<th>SVM RBF</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-D Accuracy</td>
<td>86.4</td>
<td>84.5</td>
<td>82.7</td>
<td>85.9</td>
</tr>
<tr>
<td>6-D False Positives</td>
<td>8.2</td>
<td>12.2</td>
<td>17.2</td>
<td>10.0</td>
</tr>
<tr>
<td>6-D False Negatives</td>
<td>24.1</td>
<td>20.9</td>
<td>17.5</td>
<td>20.9</td>
</tr>
<tr>
<td>12-D Accuracy</td>
<td>61.2</td>
<td>86.5</td>
<td>80.7</td>
<td>85.9</td>
</tr>
<tr>
<td>12-D False Positives</td>
<td>0.0</td>
<td>13.3</td>
<td>14.5</td>
<td>12.2</td>
</tr>
<tr>
<td>12-D False Negatives</td>
<td>1.0</td>
<td>13.6</td>
<td>28.0</td>
<td>17.3</td>
</tr>
</tbody>
</table>

good alerts using the 1-D alert algorithm was 39%. Thus, the multi-dimensional classifiers performed significantly better.

VI. DISCUSSION
A preliminary feature selection investigation examined 22 features from in-home motion and bed sensors. This initial retrospective study showed the importance of bathroom activity, sleep patterns, and socialization behavior in detecting health change for the two subjects. A 1-D health alert algorithm was then developed for testing in a prospective study with 21 seniors; feedback on clinical relevance was provided by clinicians. Using data from this larger prospective study, we showed that multi-dimensional classifiers performed significantly better than the 1-D algorithm, with the best 6-D performance at 86% compared to 39% at 1-D. The results also show that a 6-D classifier based only on domain knowledge performed slightly better than the best 6-D classifier using supervised machine learning.

One motivation for this study was to test the embedded health assessment concept in a home setting. The results show that continuous monitoring with in-home sensors can be used effectively to detect health changes. Note, however, that no diagnoses are made; there is only detection of a potential health problem. The system functions as a clinical decision support system with the clinician determining whether an intervention is warranted.

A second motivation was to identify appropriate sensor-derived features for detecting health decline. Both the feature selection study and the prospective health alert study illustrate the importance of tracking bathroom activity and bed parameters. The labeled data in Table 3 shows a much higher percentage of clinically relevant alerts for these activities compared to kitchen and living room activity.

The best features found here are reflective of the study participants and their health conditions. The best discriminating features might be different for a group of cognitively impaired individuals. It is also likely that other features will be important in different housing contexts. For example, we would expect kitchen activity to play a stronger role where seniors are responsible for preparing their own meals.
The TigerPlace clinical staff liked the health alerts and wanted to continue them in spite of the high false alarm rate. At the end of the prospective study, hyperlinks were added to the alerts to allow clinical staff to open new progress and visit notes in the EHR directly from the alert emails. Thus, their clinical workflow was improved at the same time.

### A. NEW SENSORS NEEDED

The results of the studies showed the importance of bed sensor data for capturing early signs of health decline. However, the clinical team felt that quantitative data on pulse and respiration were necessary to fully use these parameters. This led to the development of an improved bed sensor that estimates quantitative pulse and respiration rates in addition to reporting restlessness in bed. The improved sensor uses hydraulic transducers (flexible tubes of water) placed under the mattress [22] which has proven to be less invasive and more comfortable than the original pneumatic transducers placed on top of the mattress. Quantitative estimates of pulse rate, respiration rate, and restlessness are stored every 15 seconds. The new sensor also provides a clean reading of on/off bed transitions without the noise of the previous sensor. Thus, this captures a rich set of data on physiological state as well as sleep patterns.

The clinical research team also noted the importance of logging walking speed and gait patterns as a means of tracking health and fall risk. Changes in gait patterns have been linked to both physical and cognitive decline [28]–[32] and are particularly important for detecting very early signs of health decline. As a result, our team investigated the capture of in-home gait parameters with Doppler radar [37], dual webcams [36], and depth imagery from the Microsoft Kinect [38]. While walking speed and stride information can be captured with all three sensing modalities, the Kinect has proven to be the most robust and easiest to use. A gait model of walking speed, stride time, stride length, and height is learned for each resident [39]. The segmented resident walks appear as a cluster with visitors as outliers. Thus, visitor walks can be excluded from the model. Algorithms have also been developed for fall detection using Kinect depth images [46].

### B. MODELING HEALTH DECLINE WITH BEHAVIORAL BIO-MARKERS

The results of the studies show that some changes indicate health decline and some do not. After reviewing the results and considering their clinical experience in using sensor-based health alerts for longitudinal studies, our clinical research team revisited the question of which parameters are important for detecting early signs of health decline. Six clinicians contributed to the discussion: three nurses, two social workers, and one physician. Table 5 shows the changes considered and ultimately chosen for health alerts. This represents our proposed model for detecting health decline with in-home sensors.

One principle guiding this model is that daytime increases in normal daytime activity should not generate an alert. Likewise, night time increases in normal night time activity should not generate an alert. However, night time increases in normal daytime activity is a concern. Some changes are left out completely, e.g., bedroom activity and visitor activity. The clinical staff report that these changes have too many inconsistencies across different seniors to be reliable behavioral bio-markers. We have included changes in time away from home as part of the model. However, the clinical researchers were not all in agreement, noting that it is important to capture the reason seniors leave home. For example, excursions due to social engagements are viewed as a sign of good health, whereas increased time away from home due to medical appointments is likely to indicate a health problem. The social workers also noted that changes in bed sensor data are likely to appear sooner than changes in time away from home. Other sensor changes in the model are included because of our experience in both retrospective and prospective research studies such as the studies described in this paper (e.g., bed restlessness) or because of research documenting signs of health decline, e.g., gait changes [28]–[32].

The investigation of in-home sensors for early detection of health changes has been aided by an interdisciplinary team and the use of TigerPlace as a research site. From the beginning, TigerPlace was designed to be a living laboratory affiliated with the University of Missouri, Sinclair School of Nursing (SSON) [69]. The housing facility was built and is operated by Americare; however, the clinical operation is run by the SSON. About 65 senior residents reside in 54 independent apartments. An electronic health record was developed by our research team for the purpose of tracking changes in residents’ health status [70]. Thus, it provides a unique opportunity to conduct longitudinal studies on embedded health assessment.

We began installing sensors in TigerPlace apartments in October, 2005 and initially studied a small number of

<table>
<thead>
<tr>
<th>Sensor-Based Parameter</th>
<th>Increases</th>
<th>Decreases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Day</td>
<td>Night</td>
</tr>
<tr>
<td>Motion density</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Bathroom activity</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Bedroom activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living room activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kitchen activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time away from home</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Visitor activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total time in bed</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Bed restlessness</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Pulse rate</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Breathing rate</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>In-home walking speed</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>In-home stride length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-home stride time</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>
residents as case studies. We now operate 20-25 active sensor networks continuously for IRB-approved longitudinal research studies. The sensor data for the study described here was collected as part of a year-long intervention study in which 21 seniors had in-home sensors (PIR motion sensors and pneumatic bed sensors) with 1-D health alerts and 20 seniors were control subjects receiving normal care [66], [71]. With the health alert system, clinicians were able to detect early signs of urinary tract infections and other acute infections such as pneumonia, upper respiratory infections, heart failure, post-hospitalization pain, delirium, and hypoglycemia [66]. At the end of the study the intervention group showed significantly better health outcomes than the control group [66], [70], [71].

A randomized control study is currently underway in other senior housing with 70 seniors as controls and 70 seniors with sensors and health alerts. This current study uses the new hydraulic bed sensor and the Kinect-based in-home gait analysis system. A commercialization effort is also underway with a goal of keeping the system cost effective. The equipment of PIR motion sensors, bed sensor, and depth camera gait system is relatively inexpensive (about $2000). The bigger cost will be managing the deployment and providing trained clinicians for receiving and interpreting the health alerts. However, even if the annual costs are $4000-$5000, the system can be very cost effective due to the early treatment of health problems when costs are less. Avoiding hospitalizations or re-admissions, as has occurred with TigerPlace residents, can offer a substantial overall cost savings. We anticipate that the health alert system will also allow seniors more flexibility in housing choices, including the ability to stay in an existing private home. This too would offer cost savings and provide an improved quality of life for seniors as they age.

VII. CONCLUSION

In this paper, we present studies designed first to investigate embedded health assessment. A forward search was first used to retrospectively investigate the feature space of embedded in-home sensors. We also described a prospective study using 1-D health alerts. Clinical ratings on the health alerts were provided by clinicians and used to train and test multi-D classifiers. The best 6-D performance was achieved by a FPT based on domain knowledge only, although the SVM (trained on labeled training data) had a similar performance. To improve the current performance, we will investigate on-line learning using the alert ratings as feedback. The work presented here shows that domain knowledge could be used for initial classification to build up enough data to support on-line learning methods.

Finally, based on the study results and our experience using health alerts prospectively, we proposed a model for detecting health decline with in-home sensors. A randomized control study using this model with the hydraulic bed sensor, motion sensors, and in-home gait is underway to further test the potential of embedded health assessment. A system that recognizes very early signs of health decline passively, without requiring the user to wear anything, charge batteries, or do anything special, has enormous implications for seniors’ health trajectories. Identifying health decline early provides a window of opportunity for early treatment and intervention that can address health problems before they become catastrophic. This offers the potential for improved health outcomes, reduced healthcare costs, continued independence, and better quality of life.

ACKNOWLEDGMENT

The authors would like to thank the Eldertech team at the University of Missouri for their help and support with this work.

REFERENCES


MARJORIE SKUBIC (S’90–M’91–SM’13) received the Ph.D. degree in computer science from Texas A&M University, College Station, TX, USA, in 1997. She specialized in distributed telerobotics and robot programming by demonstration with Texas A&M University. In addition to her academic experience, she has spent 14 years working in industry in real-time applications, such as data acquisition and automation. She established the Center for Eldercare and Rehabilitation Technology, University of Missouri, in 2006, where she serves as the Center Director of an interdisciplinary team. The center supports proactive models of health care, such as monitoring systems that noninvasively track the physical and cognitive health of elderly residents in their homes and generate alerts that flag health changes. The center’s recent work also includes the investigation of automated screening of athletes and pianists to flag injury risks, with the support of preventative exercises to reduce the risk. She is currently a Professor with the Electrical and Computer Engineering Department, University of Missouri, with a joint appointment in Computer Science. Her current research interests include human–robot interaction, spatial referencing interfaces, computational intelligence, sensor networks for eldercare, and preventative screening tools.

RAINER DANE GUEVARA received a M.S. degree in computer engineering from the University of Missouri, Columbia, MO, USA in 2012. His thesis focused on using in-home sensor data for early illness detection of seniors. He is currently a Software Engineer with Cerner Corporation, Kansas City, MO. His interests include computational intelligence, pattern recognition, eldercare technology, and wireless sensor networks.

MARILYN RANTZ (M’12) received the Ph.D. degree in nursing from the University of Wisconsin, Milwaukee, USA, in 1992. She specialized in long-term care health policy and public program evaluation and gerontological nursing with the University of Wisconsin. She was elected into the Institute of Medicine and also became the only individual to be twice named as an American Academy of Nursing Edge Runner for two different innovations in 2012. She is currently a Curators’ Professor with the Sinclair School of Nursing, University of Missouri, Columbia, where she holds the position of University Hospital Professor of Nursing, and has an appointment as a Professor with the Department of Family and Community Medicine, University of Missouri School of Medicine.

Dr. Rantz has developed and sustained a research program to improve the quality of care of the elderly. Her innovative work on nursing home quality spans nearly 30 years, in practice and as a leading Researcher, and establishes her as a Premier International Expert in quality measurement in nursing homes. In 2012, she secured a U.S.$14.8 million grant from the Centers for Medicare and Medicaid Services for their initiative to reduce avoidable hospitalizations among nursing facility residents. In total, she and her multidisciplinary research teams have garnered over U.S.$60 million in funds to support their work measuring effectiveness of nurse care coordination, conduct cutting edge research in long-term care, new delivery models of care for older adults, and most recently, for technology development to enhance aging in place of community-dwelling elders.