Testing Classifiers for Embedded Health Assessment

Marjorie Skubic[1], Rainer Dane Guevara[1], and Marilyn Rantz[2]

[1] Electrical and Computer Engineering
[2] Sinclair School of Nursing
University of Missouri, Columbia, MO 65211
{SkubicM,RantzM}@missouri.edu, rdg5xc@mail.missouri.edu

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 activity patterns. Changes in the activity patterns are detected as potential signs of changing health. A simple alert algorithm has been 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 in developing classifiers. Here, we present the methodology and results for two classification approaches using embedded sensor data and health alert ratings collected on 21 seniors over nine months. The results show similar performance for the two techniques, where one approach uses only domain knowledge and the second uses supervised learning for training.

Keywords: in-home sensing, eldercare monitoring, health alerts.

1 Introduction

Our view of embedded health assessment is the on-going assessment of health changes based on an individual's activity patterns and baseline health conditions. Sensors embedded in the environment are used to collect activity patterns for the purpose of early detection of 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 the application of technology for enabling independent living and healthy aging. A recent review of health related smart home projects [3] included 114 relevant publications, with 71% of the projects including technologies for functional monitoring. The outcomes that have been assessed include in-home activity and restlessness captured using passive infrared (PIR) sensors [4][5], or video [6]; activities of daily living (ADLs) captured by multiple sensor types [7] [8]; sleep patterns captured using PIR motion sensors [9],

M. Donnelly et al. (Eds.): ICOST 2012, LNCS 7251, pp. 198–205, 2012.

bed mats [10][11], or load cells [12]; and walking speed using PIR sensors [13], video [14], radar [15], or depth images [16]. The variety of work in this area shows the interest and potential of the embedded health assessment approach.

A major challenge for classifier 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 ADLs to create labeled data sets, often used for studying statistical activity recognition methods, e.g., [8][17]. Other work has used much smaller datasets from a few volunteers, such as the statistical predictive algorithm to model circadian activity rhythms [18], mixture model analysis to infer activities of one user, validated with a manual log [19], and fuzzy rules used to classify activities in the home [20]. 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 activity patterns. A simple, one-dimensional 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. They then rate the clinical relevance of the alert. Here, we use the ratings as ground truth for health changes and test multi-dimensional approaches for classifying alerts as good or poor. Results are shown for 2 classifiers using embedded sensor data and health alert ratings collected on 21 seniors over nine months.

2 Sensor Network

Fig. 1 shows the sensor monitoring system for embedded health assessment. Data from sensors installed in seniors' apartments are logged and stored on a secure server. A typical installation for a one bedroom apartment consists of about 12 motion sensors, a bed sensor, and a temperature sensor for capturing stove and oven activity. 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 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 [21]. A pneumatic bed sensor [11] 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 senior 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.

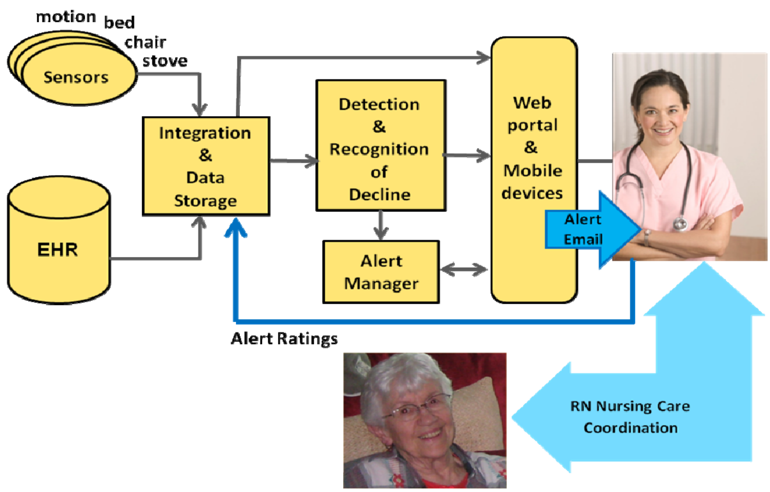

Fig. 1. Integrated sensor network with health alerts and ratings on clinical relevance captured from clinicians.

2.1 Health Alerts

The logged sensor data are automatically analyzed, looking for changes in an individual's data patterns. If a change is detected, an alert is sent to clinicians in the form of an email. The alert email includes two web links. One is a link into the web portal which facilitates fast access to the sensor data for the resident, 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 on health changes to aid in the 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. For the study reported here, 4 clinicians provided feedback (one physician and three nurses), based on their clinical expertise with older adults [22].

2.2 Alert Parameters

The alert algorithm was developed through collaboration with clinicians and was intentionally designed as a simple algorithm to cast a wide net so that critical health changes were captured even if it resulted in a high percentage of false alarms. 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. The two week moving baseline was chosen on the recommendation of clinicians after retrospective analysis, as a compromise to capture both sudden and gradual health changes [22]. Each resident has a personalized *normal* that is reflected uniquely in the sensor data patterns, depending on the chronic health condition(s), the 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 chronic ailments.

Table 1 shows the alert parameters and sensor data monitored for the health alerts. 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, according to the 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).

With this one-dimensional 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; however, it is difficult to tell how many potential alerts we are missing. 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 one-dimensional approach.

Table 1. Alert parameters and sensor data monitored for the alerts.

Alert parameter	**Sensors**
Bathroom Activity	Sum of motion sensor events in the bathroom (bathroom, shower, laundry)
Bed Restlessness	No. of all bed restlessness events
Bed Breathing Low/Normal/High	No. of bed breathing low/normal/high events
Bed Pulse Low/Normal/High	No. of bed pulse low/normal/high events
Kitchen Activity	Sum of kitchen motion sensor (kitchen, fridge, etc.) events and stove/oven temperature high
Living Room Activity	No. of living room motion sensor events

3 Classifier Methodology

We have investigated different classifiers 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 in training and testing. Here, we discuss the application of two classifiers to this problem. The first is a fuzzy pattern tree that does not require training but rather takes advantage of domain knowledge from our clinical partners. The second is the support vector machine which uses training data and supervised learning to train the classifier. These classifiers were chosen for the study to provide a comparison between the use of domain knowledge vs. a trained classifier that supports a nonlinear decision boundary.

3.1 Feature Space

In analyzing the health alerts generated for the parameters listed in Table 1, 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. At this point in the study, there were also very few decrease alerts. Thus, in the end, we looked at the increases in the following four alert parameters: bathroom activity, bed restlessness, kitchen activity, and living room activity. If increased 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 12. Fig. 2 shows a PCA reduction of the 12-dimensional feature space, where blue indicates the good alert days and red indicates the bad alert days; 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 as relative sensor values: increased nighttime activity in the living room, kitchen, and bathroom, increased full day activity in the bathroom, and increased bed restlessness at both nighttime and during the full day. A PCA reduction of this 6-dimensional feature space is shown in Fig. 3. Normal days (i.e., poor alert days) tend to cluster, and the abnormal days (good alert days) tend to be outliers around the cluster center. We report the methods and results using these six features.

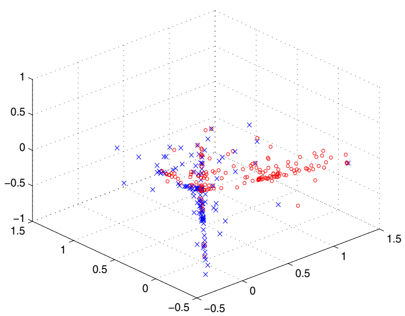

Fig. 2. PCA reduction of 12-D feature vectors from the health alert study. Red are poor alert days; blue are good alert days as rated by clinicians.

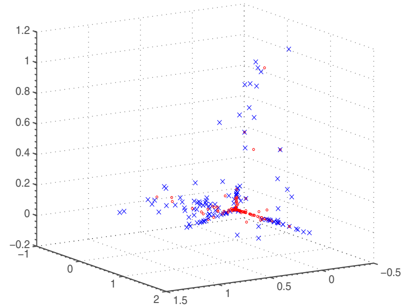

Fig. 3. PCA reduction of 6-D feature vectors from the health alert study. Red are poor alert days; blue are good alert days as rated by clinicians.

3.2 Fuzzy Pattern Tree

A fuzzy pattern tree (FPT) [23] was investigated as a method that uses domain knowledge only 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. 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 [24] 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 experimentally.

3.3 Support Vector Machine

The support vector machine (SVM) was also tested to investigate a supervised learning approach. We investigated both a linear and radial basis function (RBF) kernel (using Matlab functions). We also tested both the 12 features and the six features described above. The RBF kernel performed slightly better than the linear kernel. The performance of the 12 features and six features was almost identical. The results reported here and compared to the FPT are the RBF kernel with six features.

4 Results

The FPT and SVM classifiers were tested with a set of health alert ratings spanning nine months on 21 senior residents. Table 2 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 but will be included in future work.

Fig. 4 shows the ROC curves of the two classification methods using the six features described in Sec. 3. The FPT was constructed with domain knowledge only as described above. The SVM results reported here use a RBF kernel and were evaluated with 10-fold cross validation. The performance of the two classifiers, as shown in the ROC curve, is very similar. Both achieved about 85% correct classification at the highest rate. As shown in Table 2, the percentage of good alerts using the one-dimensional alert algorithm was less than 40%. Thus, the multi-dimensional classifiers performed significantly better.

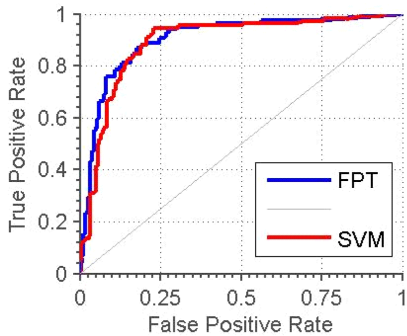

Fig. 4. ROC curves for the SVM and fuzzy pattern tree (FPT) using 6D feature vectors

Table 2. Health Alerts Used for Testing

Alert parameter	Good alerts	Poor alerts	Total
Bathroom Activity	42	43	85
Bed Restlessness	57	21	78
Kitchen Activity	7	63	70
Living Room Activity	17	85	102
Total	116	183	299

5 Conclusions

In this paper, we present work on testing two classification approaches for detecting early health changes. The ground truth was taken from health alert ratings provided by clinicians at a senior housing facility in which embedded sensors monitor seniors in their apartments. The two multi-dimensional classifiers tested include a fuzzy pattern tree that uses only domain knowledge of the clinicians and a support vector machine trained by supervised learning. Both classifiers achieved about 85% correct classification compared to less than 40% for a single-dimensional algorithm. In future work, we will explore other classification methods and test different baseline time periods. To improve over the current performance, we will investigate on-line learning using the alert ratings as feedback. The work presented in this paper shows that domain knowledge could be used as an initial classification scheme to build up enough data to support on-line learning methods.

References

1. Boockvar, K.S., Lachs, M.S.: Predictive value of nonspecific symptoms for acute illness in nursing home residents. J. of the American Geriatrics Society 51(8), 1111–1115 (2003)
2. Ridley, S.: The recognition and early management of critical illness. Annals of the Royal College of Surgeons of England 87(5), 315–322 (2005)
3. Demiris, G., Hensel, B.K.: Technologies for an aging society: a systematic review of "smart home" applications. Year. Med. Inform., 33–40 (2008)

4. Ohta, S., Nakamoto, H., Shinagawa, Y., Tanikawa, T.: A health monitoring system for elderly people living alone. Journal of Telemedicine & Telecare 8(3), 151–156 (2008)
5. Chan, M., Campo, E., Esteve, D.: Assessment of activity of elderly people using a home monitoring system. Intl. Journal of Rehabilitation Research 28(1), 69–76 (2005)
6. Harvey, N., Zhou, Z., Keller, J.M., Rantz, M., He, Z.: Automated estimation of elder activity levels from anonymized video data. In: Proc., Intl. Conf. IEEE Eng. Med. Biol. Soc., pp. 7236–7239 (2009)
7. Fleury, A., Vacher, M., Noury, N.: SVM-based multimodal classification of activities of daily living in Health Smart Homes: Sensors, algorithms, and first experimental results. IEEE Transaction on Information Technology in Biomedicine 14(2), 274–283 (2010)
8. Singla, G., Cook, D.J., Schmitter-Edgecombe, M.: Recognizing independent and joint activities among multiple residents in smart environments. J. Ambient Intell. Humaniz. Comput. 1(1), 57–63 (2010)
9. Hayes, T.L., Riley, T., Pavel, M., Kaye, J.A.: A method for estimating rest-activity patterns using simple pyroelectric motion sensors. In: Proc., Intl. Conf. IEEE Eng. Med. Biol. Soc. (2010)
10. Mack, D.C., Patrie, J.T., Suratt, P.M., Felder, R.A., Alwan, M.: Development and preliminary validation of heart rate and breathing rate detection using a passive, ballistocardiography-based sleep monitoring system. IEEE Transaction on Information Technology in Biomedicine 13, 111–120 (2009)
11. Heise, D., Rosales, L., Skubic, M., Devaney, M.J.: Refinement and Evaluation of a Hydraulic Bed Sensor. In: Proc., Intl. Conf. IEEE Eng. Med. Biol. Soc., pp. 4356–4360 (2011)
12. Beattie, Z.T., Hagen, C.C., Pavel, M., Hayes, T.L.: Classification of breathing events using load cells under the bed. In: Proc., Intl. Conf. IEEE Eng. Med. Biol. Soc., pp. 3921–3924 (2009)
13. Hagler, S., Austin, D., Hayes, T.L., Kaye, J., Pavel, M.: Unobtrusive and ubiquitous in-home monitoring: a methodology for continuous assessment of gait velocity in elders. IEEE Trans. Biomed. Eng. 57(4), 813–820 (2010)
14. Stone, E., Anderson, D., Skubic, M., Keller, J.M.: Extracting footfalls from voxel data. In: Proc. Intl. Conf. IEEE Eng. Med. Biol. Soc. (2010)
15. Yardibi, T., Cuddihy, P., Genc, S., Bufi, C., Skubic, M., Rantz, M., Liu, L., Phillips, C.: Gait characterization via pulse-Doppler radar. In: Proc., IEEE Intl. Conf. Pervasive Computing and Communications: SmartE Workshop, pp. 662–667 (2011)
16. Stone, E., Skubic, M.: Evaluation of an Inexpensive Depth Camera for In-Home Gait Assessment. Journal of Ambient Intelligence and Smart Environments 3(4), 349–361 (2011)
17. Cook, D., Schmitter-Edgecombe, M.: Assessing the quality of activities in a smart environment. Methods of Information in Medicine 48(5), 459–467 (2009)
18. Virone, G., Alwan, M., Dala, S., Kell, S., Stankovic, J.A., Felder, R.: Behavioral patterns of older adults in assisted living. IEEE Transaction on Information Technology in Biomedicine 12(3), 387–398 (2008)
19. Barger, T.S., Brown, D.E., Alwan, M.: Health-status monitoring through analysis of behavioral patterns. IEEE Trans. SMC-A 35(1), 22–27 (2005)
20. Brown, S., Majeed, B., Clarke, N., Lee, B.-S.: Developing a well-being monitoring system-Modeling and data analysis techniques. In: Mann, W., Helel, A. (eds.) Promoting Independence for Older Persons with Disabilities, Washington, DC (2006)
21. Wang, S., Skubic, M., Zhu, Y.: Activity density map dis-similarity comparison for eldercare monitoring. In: Proc., Intl. Conf. IEEE Eng. Med. Biol. Soc., pp. 7232–7235 (2009)
22. Rantz, M.J., Skubic, M., Koopman, R.J., Alexander, G., Phillips, L., Musterman, K.I., Back, J.R., Aud, M.A., Galambos, C., Guevara, R.D., Miller, S.J.: Automated technology to speed recognition of signs of illness in older adults. J. of Gerontological Nursing (in press)
23. Theodoridis, S., Koutroumbas, K.: Pattern Recognition. Academic Press (2006)
24. Yager, R.R.: On a general class of fuzzy connectives. Fuzzy Sets and Systems 4(3), 235–242 (1980)