The 4th IEEE International Conference on E-Health and Bioengineering - EHB 2013
Grigore T. Popa University of Medicine and Pharmacy, Iaşi, Romania, November 21-23, 2013

Detection of Abnormal Sensor Patterns in Eldercare

Zahra Hajihashemi
Computer Science Department
University of Missouri,
Columbia
MO, USA
Zhmr5@mail.missouri.edu

Mihail Popescu
Health Management and Informatics Department
University of Missouri,
Columbia
MO, USA
PopescuM@missouri.edu

Abstract— ***In TigerPlace, an aging in place facility from Columbia, MO, we deployed 47 sensor networks together with an electronic health record (EHR) system to provide early illness recognition. In this paper, we describe a framework for predicting abnormal health patterns using non-wearable sensor sequence similarity. To compute the similarity between two sensor patterns we employ the Temporal Smith Waterman. A sensor pattern is classified as "abnormal" if it is much smaller than the mean of the distribution of "normal" patterns similarities. "Abnormal" days are defined by unusual sensor activity patterns that require a nurse's assessment of the resident. On a pilot data set of 1685 sensor days and 626 nursing records, we obtained a classification performance with an average precision of 0.70 and a recall of 0.30.***

Keywords: sensor networks, early illness detection, Smith Waterman algorithm.

I. Introduction

Elderly population (aged 65 and older) is rapidly increasing worldwide from 13% in 2010 to 19% in 2030, while working-age population (age between 15 and 64 years) is projected to decline [1]. The preference of older adults to live independently regardless of many health conditions associated with old age, such as frailty, dementia, and risk of falling requires more attention and sophisticated health monitoring systems. However, independent living may lead to infrequent health assessments due to the lack of continuous monitoring or fear of being institutionalized. Late health assessments may miss unreported complications, which in turn lead to poor long-term prognosis and quality of life [2]. A possible solution to preventing unreported health problems in independently living older adults is through automatic health monitoring systems.

One efficient approach to health monitoring is to use sensor networks to collect information about the older adult's activity. In the last decade, many of such health monitoring systems have been piloted. For example, MIT's PlaceLab, Georgia Tech's aware House, Honeywell's Independent Lifestyle Assistant, and University of Missouri' s TigerPlace have demonstrated possible approaches to activity monitoring [3,4,5]. A variety of methodologies for detecting activity and assess medication compliance have been reported in the literature [6,7,8,9,10]. Even though these and many other systems are successful examples of applying sensor networks to monitor activity patterns, the major unsolved challenge is to consider the health context of the monitored older adults.

In TigerPlace [12,13] the goal is to provide a stable home environment using independent living facility for senior citizens. As a result of our collaboration between Sinclair School of Nursing, Columbia, Missouri, and Americare Systems Inc. of Sikeston, Missouri, elder residents live as healthy and independent as possible [14]. We utilized our facility with sensor technology to provide early illness detection. We have installed sensor networks in the apartments of 47 residents, a system that has been active since fall of 2005. In the previous work [2], we described a version of our framework system for early illness recognition. In [15], we introduced a new sensor similarity algorithm based on a modified Smith-Waterman algorithm, called Temporal Smith Waterman (TSW). Moreover, we improved medical term extraction method based on Unified Medical Language System (UMLS), which have been tested on multiple resident datasets.

In this paper, we present a new early illness recognition approach based on TSW pattern similarity. In the next section, we briefly describe the system architecture and sensor data. In the section "Method", we use the similarity measurement (TSW) and introduce a new confidence level for the purpose of classification. Section IV shows our experimental results and discussion. Finally, in the last section we give conclusions and future directions.

I. System Architecture

Technology has a tremendous impact on elderly by offering them full productive and independence lives. In TigerPlace, we utilize sensor technology to help elderly residents not only manage their illness but also stay as healthy and independent as possible. We deployed our integrated monitoring system in 47 TigerPlace apartments with the University of Missouri IRB approval. After focus groups with TigerPlace residents early on in our research [9,10], we decided to use only non-wearable sensors for monitoring, since they are unobtrusive and more acceptable by older adults. The monitoring has started in fall 2005. On average, we have two years of data for each resident. Figure 1 shows the architecture of our data acquisition system.

The main components of the monitoring system are: an integrated motion/video/Kinect sensor network, a data logger, a reasoning system that analyses sensor patterns, an electronic health record (EHR) system, an alert manager to notify clinicians of potential problems, and a secure Web-based interface to display the data for the clinicians and researchers. The logger unit collects data from the sensor network of the monitoring system, date-time stamps the data, and logs it into a database. The computer from each apartment is connected to a main storage server through a secure wired network connection. The main server receives, stores, and monitors sensor data of all apartments in TigerPlace for research purposes.

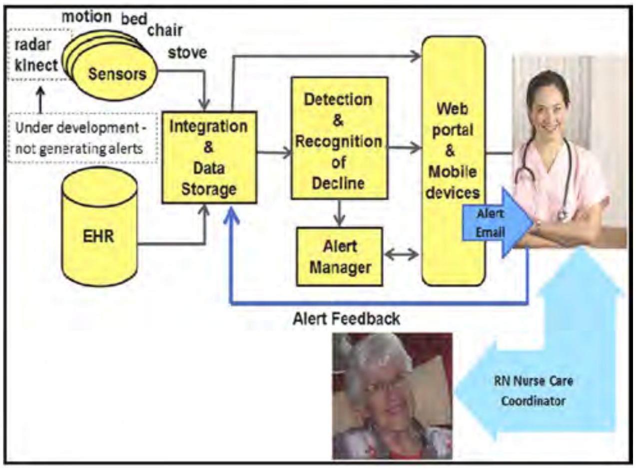

Fig. 1. TigerPlace monitoring system architecture.

The sensor network employs a variety of sensors to capture various aspects of the residents' behavior. Figure 2 displays a floor map of a typical TigerPlace apartment that shows the placement of some of the sensors.

The sensor network includes sensors placed in the living room, bedroom, bath room, and kitchen. Pulse-Doppler radar and Microsoft Kinect sensors (Microsoft, Redmond WA) are also included to the system (not shown). The passive infrared motion sensors detect activities in each room. For example, kitchen sensors such as stove temperature sensors, cabinet motions, and refrigerator recognize presents in the kitchen and imply food preparation activities. Bathroom activities are monitored by a motion sensor installed above the shower.

Sleep patterns, as an important night time patterns, are captured by bed sensors. The bed sensor is a pneumatic strip which lies on top of the mattress under the sheets and detects presence in bed. The bed sensor measures qualitative pulse and respiration in three levels (low, normal, high), and restlessness in four levels of (low, normal, high, very high) while the resident sleeps.

We developed a web interface to present the sensor data in a way that health care providers can easily understand and interpret. Users can filter their searches by resident information and a date interval. The sensors are grouped by type: motion, bed restlessness, bed pulse, and bed respiration. Histograms are used to display motion and bed sensor events, which are aggregated to a daily level. Figure 3 shows the user interface of the sensor display for health care providers and research staff. This interface shows aggregated motion sensors data from 2013-04-12 to 2013-07-26 for a specific resident.

II. METHOD

The aim of this study is to predict the change in resident's health status based on sensor data produced by in-home monitoring system. The power of sensor networks for predicting health patterns using logistic regressions, one class classifiers, multiple instance learning and temporal clustering has been previously investigated [14, 16, 17, 18]. In this paper, we explore the detection of abnormal patterns by employing the distribution of similar sensor sequences. We assume that an abnormal sensor pattern is due to a medical condition, fact that is true in general for the elderly population from TigerPlace. The difficulty of finding similar sensor patterns is given by the multidimensional nature of sensor data and their huge volume. We argue that, in general, sensor data represents a case of big data.

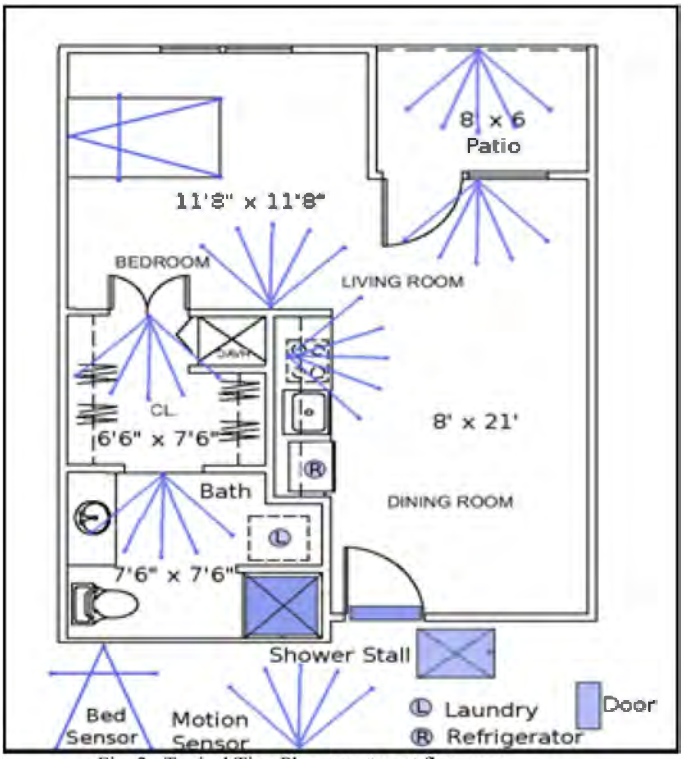

Fig. 2. Typical TigerPlace apartment floor map.

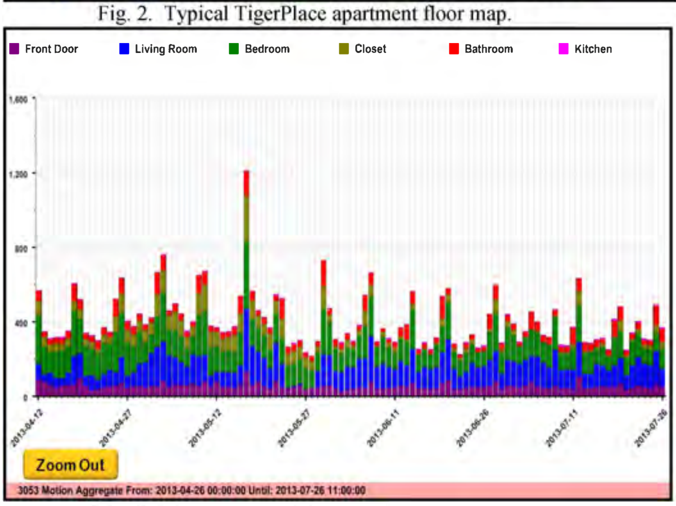

Fig. 3. Motion sensors histogram.

Temporal Smith Waterman (TSW)

In [15] we introduced a new multi-dimensional sensor sequence similarity, Temporal Smith Waterman (TSW) that considered sensor patterns as 1-dimensional sequences of characters together with their associated time stamps. In TSW, the time stamp shows the time when a character (firing) is emitted by the related sensor network. Consider two time-stamped sequences of characters as $T_1=\{(C_{11}\ t_{11}), (C_{12}\ t_{12}), \ldots, (C_{1m}\ t_{1m})\}$ and $T_2=\{(C_{21}\ t_{21}), (C_{22}\ t_{22}),\ldots,(C_{2n}\ t_{2n})\}$ where $m, n \in N$ and C_{ij} belongs to alphabet Σ and t_{ij} are the time of the day of the firing C_{ij}. Figure 4 shows the TSW algorithm.

$$H_{i0}=H_{0j},\ i \in [1,n] \text{ and } j \in [1,m] \quad (1)$$

$$H_{ij}=\max\{0,\ H_{i-1,j-1}+S(C_{1i}, C_{2j}),\ \max_{k\geq 1}\{H_{i-k,j} - W_{\Delta t}\},\ \max_{k\geq 1}\{H_{i,j-k} - W_{\Delta t}\}\} \quad (2)$$

$$W_{\Delta t} = g + c|t_{1i}-t_{2j}| \quad (3)$$

$$Final\ Score = \frac{\max\{H_{ij}\}}{\min\{n,m\}} \quad (4)$$

Fig. 4. Temporal Smith Waterman algorithm.

Essentially, TSW considers the time of the day between the sensor firings as a gap and computes the gap penalty $W_{\Delta t}$ by using time stamps (as shown in equation 3 above). We use the "time of the day" metric, since we would like to find similar behaviors across different days that happen at roughly the same time. Note that the type of time used in (3) is critically important and depends on the application. In our case, the time in (3) is input in seconds. In this paper, we use TSW to classify sensor sequences as "normal" or "abnormal" days. "Abnormal" days are defined by unusual sensor activity patterns that require a nurse's assessment of the resident and were documented in EHR. The system automatically sends an email alert to the nurse in case that some unusual pattern is detected. After assessing the resident, the nurse records pertinent comments in the EHR. However, on a "normal" day the sensor activity pattern is not flagged to prompt a nurse's visit to the resident. We note that a simplified alert system based on individual sensor values is already in place at TigerPlace [19]. The disadvantage of this system is that a given medical condition generates multiple alerts (one for each sensor) that may confuse the clinical staff. Instead, our system generates only one alert per incident.

Abnormal pattern detection algorithm

Given *n* (training) normal sensor sequences $\{S_i\}_{i=1,n}$ we compute the pair-wise similarities between them, $\{s_{ij}\}_{i,j=1,n}$ using TSW. We, then, calculate the distribution of the $\{s_{ij}\}$ similarities of these "normal" days assuming they follow a Gamma distribution. Assume that we found out that $\{s_{ij}\}$ follows a distribution Gamma with parameters *a* and *b*, i.e. $\Gamma(a,b)$. To classify an unknown sequence S_x, we start by computing its similarities with all normal sequences $S_1,\ldots S_n$, obtaining similarities $\{s_{ix}\}_{i=1,n}$. Then, we find the maximum of $\{s_{ix}\}$, $s_{x,max}$. The confidence that S_x is abnormal, $C(S_x)$ is 0 if $s_{x,max} > mean(\Gamma(a,b))$ and equal to $1-P$ where P is the likelihood that $s_{x,max}$ comes from $\Gamma(a,b)$ and is calculated using:

$$P(s_{x,max}, a, b) = \int_0^{s_{x,max}} \left(\frac{1}{\Gamma(a)b}\left(\frac{t}{b}\right)^{a-1} e^{-\frac{t}{b}}\right) dt \quad (5)$$

While in our experiments we use all normal days available, in a real system implementation we would use the data from the last two weeks.

C. Classification experiments

The classification experiments are performed using a leave-one-out approach. We compare our results to the ones obtained using a *k*-nearest neighbor (*k-NN*) approach with *k=1* and using the same TSW distance as above. In the next section we present our experimental results.

III. Experimental Results

A. Data Set

To investigate the performance of the proposed illness prediction methodology based on TSW, we run a set of experiments to train a classifier for each resident separately. Table I shows the set of sensors in the resident's apartment. We processed a log file of the sensor events (displayed in the table II) to extract the sensor sequences of "normal" and "abnormal" days in a predefined time interval. Table III shows the pilot sensor data from the apartments of two TigerPlace residents in this study. For each resident, we also retrieved visit notes about physical, emotional and other health complaints, recorded by the nurses in the TigerPlace nursing EHR.

B. Classification Results

We initialized the parameters of TSW as *g*=0 and *c*=0.002 which means that the relevant time interval is about 500s~8 minutes [15]. Figure 5 shows similarity distribution in the morning time interval (8am to 10am). "Normal" day's similarity has a Gamma distribution. We used this

TABLE I
List of Sensors in the Resident's Apartment

ID | Sensors firings	6.Breathing2	12. Living Motion	18. Drawer
1. BedMovement1	7.Breathing3	13. BathRoom	19. Cabinet
2. BedMovement2	8. Pulse1	14. Off Chair	20. Cup Cabinet
3. BedMovement3	9. Pulse2	15. On Chair	21. Refrigerator
4. BedMovement4	10. Pulse3	16. TempHigh	22. Plate Cabinet
5. Breathing1	11. Bedroom Motion	17. TempLow	23. Silverware Drawer

TABLE II
Sensor Sequence Snippet for TigerPlace Resident #2

User ID	Sensor ID	Year	Mo.	Day	Hour	Min.	S.
3	3	2005	10	5	12	34	38
3	2	2005	10	5	12	36	52
3	2	2005	10	5	12	37	04
3	2	2005	10	5	12	37	11
3	1	2005	10	5	12	37	26

TABLE III
TigerPlace Pilot Dataset

Resident #	Number of sensor days	Number of comments	Number of abnormal days
#1	440	83	81
#2	745	44	35
#3	500	499	335

distribution to set the threshold in the training classifier phase. From this point of view our method represents a one class classifier: if the similarity between a new day and the previous normal ones has a low likelihood, then the day is either abnormal or represents a new pattern.

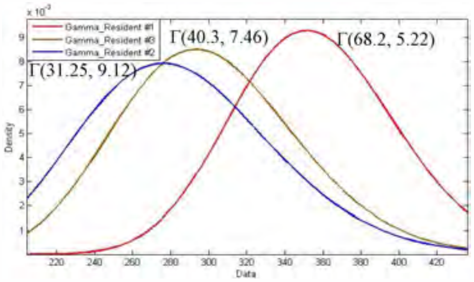

Fig. 5. Distribution of normal day pattern similarity

For similarity scores less than the threshold, we used formula (5) to calculate the confidence level of the similarity between two sensor sequences. This formula inspires the Gamma probability distribution to consider not only similarity scores but also the density of the scores of similar days. We used this similarity score and confidence level to train a classifier. For comparison purposes, we run a set of experiments with one class *k*-nearest neighbor (*k*=1) classifier where we use the similarity score without confidence level. We show the results of this comparison in table VI as area under ROC curve. Our proposed method outperforms *k*-nearest neighbor (*k*=1) in terms of accuracy. The main reason for this outcome is in *k*-nearest neighbor (*k*=1) method we do not consider the distribution of the similarity scores. In elderly health monitoring systems, usually emergency situations are so different from the elderly health history. Therefore, classifying based on the similar abnormal patterns does not produce satisfactory results. However, TSW algorithm with distribution similarities considers the history of normal patterns and using a confidence level classifies all abnormal patterns in a more accurate way. Consequently, this method is more accurate compare to *k*-nearest neighbor (*k*=1). We present the ROC curves comparisons in figure 6.

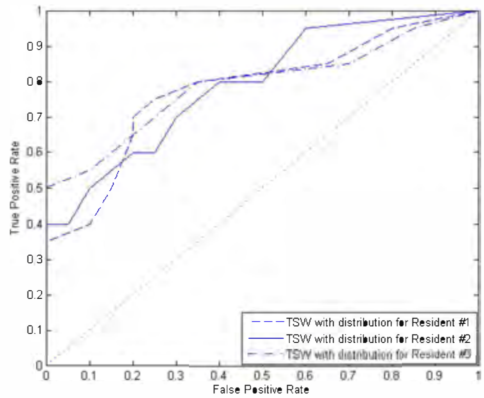

Fig. 6. Comparison between the results obtained using the similarity distribution and *k-nn*

TABLE VI
AREA UNDER ROC CURVES

Similarity Method	AUC #1	AUC #2	AUC #3
TSW with Distribution	0.79	0.78	0.79
TSW with *k-NN*	0.54	0.56	0.60

IV. CONCLUSION

In this paper we enhanced our framework for illness prediction using sensor networks. We used TSW as similarity measure to train a k-nearest neighbor classifier. We compared our results with binary *k*-nearest neighbor. In this comparison, the proposed TSW framework outperformed binary *k*-nearest neighbor. In future research we plan to add more sensors and residents data to our study.

ACKNOWLEDGMENT

This project has been funded by a NSF SHWB grant, award #: IIS-1115956.

REFERENCES

[1] KV Grayson, AV Victoria, “The next four decades: the older population in the United States: 2010 to 2050”, U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau, 2010, pp.1-14.

[2] Z Hajihashemi , M Popescu, “ Predicting Health Patterns Using Sensor Sequence Similarity and NLP”, 2012 IEEE International Conference on BIBM, 2012, pp.948-950.

[3] S Intille, K Larson, E Munguia-Tapia, J Beaudin, P Kaushik, J Nawyn, et al, “Using a live-in laboratory for ubiquitous computing research”, Proc. of Pervasive Health, 2006.

[4] CD Kidd, RJ Orr, GD Abowd, IA Atkeson, B Essa, B MacIntyre, et al., “The Aware Home: A living laboratory for ubiquitous computing research”, Proc of the 2nd International Workshop on Cooperative Buildings-CoBuild'9, 1999.

[5] KZ Haigh, LM Kiff, G Ho, “Independent Lifestyle Assistant: Lessons learned”, Assistive Technology, 2006, pp.87-106.

[6] D Heise, M Skubic, “Monitoring pulse and respiration with a non-invasive hydraulic bed sensor”, Proc 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Buenos Aires, Argentina. 2010.

[7] R Beckwith, “Designing for ubiquity: The perception of privacy”, Pervasive Computing, 2003, pp.40-6.

[8] D Mack, M Alwan, B Turner, R Suratt, R Felder, “A passive and portable system for monitoring heart rate and detecting sleep apnea and arousals: Preliminary validation”, Proceedings Transdisciplinary Conference on Distributed Diagnosis and Home Healthcare (D2H2). 2006.

[9] MJ Rantz, K Dorman-Marek, M Aud, HW Tyrer, M Skubic, G Demiris, et al., “A technology and nursing collaboration to help older adults age in place”, Nursing Outlook, 2005, pp.40-45.

[10] M Skubic, G Alexander, M Popescu, M Rantz, J Keller, “A Smart Home Application to Eldercare: Current Status and Lessons Learned”, Journal of Technology and Health Care, 2009, pp.183-201.

[11] TF Smith, MS Waterman, “Identification of common molecular subsequences”, Journal of Molecular Biology, 1981,147, pp.195–197.

[12] JS Goodwin, B Howrey, BD Zhang, Y-F Kuo, “Risk of Continued Institutionalization after Hospitalization in Older Adults”, J Gerontol A Biol Sci Med Sci., 2011, pp.1-7.

[13] MJ Rantz, KD Marek, MA Aud, RA Jounson, D Otto, D Porter, “TigerPlace: A New Future for Older Adults”, J. of Nursing Care Quality, 2005, Vol. 20(1), pp.1−4.

[14] M Popescu, E Florea, “Linking Clinical Events in Elderly to In-home Monitoring Sensor Data: A Brief Review and a Pilot Study on Predicting Pulse Pressure”, Journal of Computing Science and Engineering, March 2008, vol. 2(1), pp.180-199.

[15] Z Hajihashemi, M Popescu, “An Early illness recognition framework using a temporal smith waterman algorithm and NLP”, AMIA 2013, Chicago, November, 2013.

[16] M Popescu, A Mahnot, “Early Illness Recognition in Older Adults Using In-Home Monitoring Sensors and Multiple Instance Learning”, Methods of Informatics in Medicine, 2012, Vol. 51(4), pp. 359-67.