Detecting Daily Routines of Older Adults Using Sensor Time Series Clustering

Zahra Hajihashemi, *Member, IEEE*, Maria Yefimova, and Mihail Popescu, *Senior Member, IEEE*

***Abstract*— The aim of this paper is to develop an algorithm to identify deviations in patterns of day-to-day activities of older adults to generate alerts to the healthcare providers for timely interventions. Daily routines, such as bathroom visits, can be monitored by automated in-home sensor systems. We present a novel approach that finds periodicity in sensor time series data using clustering approach. For this study, we used data set from TigerPlace, a retirement community in Columbia, MO, where apartments are equipped with a network of motion, pressure and depth sensors. A retrospective multiple case study (N=3) design was used to quantify bathroom visits as parts of the older adult's daily routine, over a 10-day period. The distribution of duration, number, and average time between sensor hits was used to define the confidence level for routine visit extraction. Then, a hierarchical clustering was applied to extract periodic patterns. The performance of the proposed method was evaluated through experimental results.**

I. INTRODUCTION

Using technology to monitor functional abilities of older adults will improve management of healthcare, housing and social service [1]. Exacerbations of chronic conditions associated with aging lead to adverse events, hospitalization, and higher use of healthcare resources [2], preventing older adults from living independently. Identifying illnesses early will improve timing and adjustment of services to manage or reverse functional decline.

To achieve this goal, the living environment can be equipped with passive sensor networks to monitor the daily life of the older adult. Functional ability can be reflected in the performance of routine tasks that include activities of daily living (such as bathing, dressing, hygiene, and bowel movement), more complex instrumental activities of daily living (s. a. housework, finance management, and shopping), and general life activities (s. a. hobbies, leisure past time and socialization) [3]. Routines develop within temporal patterns of day and night, weekday and weekend, and seasonal change [3]. This periodicity of routine behavior can be modeled with mathematical algorithms. Extracting and analyzing attributes of daily routines may be useful in understanding trajectories of functional decline of older adults and can suggest timely interventions.

To identify daily routines within large datasets, one approach is to use association rule mining techniques [4]. In sequences of nominal variables that do not have natural ordering, such as our sensor data, the goal is to find analogous subsets, or "primitive shapes", rather than similar subsequences [5]. In [6, 7], authors addressed the challenge of multidimensionality and huge volume of sensor data by detecting frequent patterns (motifs) to summarize and visualize massive time-series datasets. They identified both globally occurring motifs and locally representative trends in time series with high probability, even in the presence of noise or "don't care" symbols.

To detect temporal motifs in multi-dimensional temporal streams, labeled data has been used [8]. First, the more important dimensions are selected using SAX [9]. Then, the frequent motifs of each dimension are discovered using a sliding window and metrics, such as probability of detection (POD) and the false alarm ratio (FAR). POD is the number of times an event was correctly predicted divided by the total number of observed events. To address the problem of extending the structure of motif extraction, [10] proposed a suffix-tree-based Flexible and Accurate Motif Detector (FLAME) that allows mining frequent combinations of motifs under relaxed constraints.

In this paper, we propose a novel computational approach for mining periodic patterns of bathroom visits, as part of the daily routine, based on the similarity of sensor firing sequences (SFS) and frequent SFS patterns. This paper is organized as follows: section II presents our system architecture and dataset; section III describes our method; and section IV concludes with our experimental results.

II. SYSTEM ARCHITECTURE

With the IRB approval of the University of Missouri, we deployed our in-home monitoring system in 47 apartments at TigerPlace, a unique aging-in-place retirement community. Data has being collected since fall 2005. Based on focus group studies with TigerPlace residents, we decided to use only non-wearable sensors for monitoring since they are unobtrusive and more acceptable by older adults [11]. Figure 1 shows the architecture of our data acquisition system.

Various sensors have been deployed in each apartment: motion, radar, Microsoft Kinect, and pressure bed sensors. When activated, each motion sensor sends an X10 signal (firing) indicating its ID that is logged with a time stamp in our database (see more details: http://eldertech.missouri.edu). Our previous work [12, 13, 14] involved 23 motion sensors to propose a new sensor sequence similarity measure named Temporal Smith Waterman (TSW). In this paper, we focus on the detection of periodic bathroom visits, considering three particular motion sensors: bathroom, shower, and closet/laundry.

Research supported by the NSF SHWB Grant Award #IIS-1115956 and the John A. Hartford Foundation's NHCGNE Award Program.

Z. Hajihashemi is with Computer Science Department, University of Missouri, Columbia, MO 65203 USA (zhmr5@ mail.missouri.edu).

M. Yefimova is with the School of Nursing, University of California, Los Angeles, CA 90025 USA (m.yefimova@ucla.edu).

M. Popescu is with the HMI Department, University of Missouri, Columbia, MO 65203 USA (popescuM@missouri.edu).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

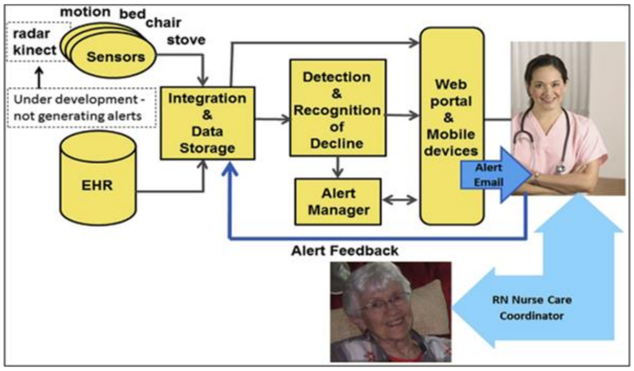

Figure 1. Integrated sensor network under development at TigerPlace.

Our sample consists of three residents with different bathroom habits. Their characteristics are described in Table I. Resident #1 does not have any reported urinary problems, but takes diuretic medications for high blood pressure. It may alter his bathroom habits, causing more frequent and varied visits. Resident #2 is diagnosed with urinary retention from an enlarged prostate and has difficulty emptying the bladder. Resident #3 is occasionally incontinent. All three take laxatives to improve bowel motility.

TABLE I. TIGERPLACE RESIDENT CHARACTERISTICS

Resident ID	Age	Gender	Urinary Problems	Ambulates with Walker	Diuretic Medication
#1	88	Male	None	No	Yes
#2	99	Male	Retention	Yes	No
#3	90	Female	Incontinence	Yes	No

Table II shows an example of a bathroom visit as recorded by our in-home monitoring system. As the resident moves through the apartment, the motion sensor firing data is added to the log file database. In this example, on October 5, 2005, around 12:30AM Resident #3 was in the bathroom for about 3 minutes. The visit begins with the first bathroom sensor hit and ends when the person enters another room (walk-in closet). The sequence has the length of 5 sensor hits.

TABLE II. EXAMPLE OF SENSOR FIRINGS FOR A BATHROOM VISIT

Resident ID	Sensor	Year	Month	Day	Hour	Min	Sec
#3	Bathroom	2005	10	5	12	34	38
#3	Bathroom	2005	10	5	12	36	52
#3	Bathroom	2005	10	5	12	37	04
#3	Bathroom	2005	10	5	12	37	11
#3	Closet	2005	10	5	12	37	26

Table III shows the total number of known bathroom visits from the apartments of the three residents for the period of 10 consecutive days used in this study.

TABLE III. TIGERPLACE PILOT DATASET

Resident ID	Number of bathroom visits	Number of shower visits
#1	74	3
#2	82	4
#3	69	1

Because routine behaviors are person-specific, mathematical models cannot be transferrable from one resident to another. Hence, we view our data as three separate experiments with more than 200 samples of bathroom visits, rather than one experiment with only three samples.

III. METHOD

We focus on bathroom visits, as an important daily routine, and extract three features: duration, number of sensor hits, and the average time between hits. We use distributions of these parameters to define the confidence level for validating the automated bathroom visit extraction process. Then, a hierarchical clustering approach is applied to cluster visits and extract frequent bathroom routine patterns.

A. Automated Routine Extraction Process

We apply a rule-based approach to automatically extract bathroom visits from sensor log files. The rules are based on the apartment floor maps and geospatial locations of the motion sensors in each apartment. Then, we apply a hierarchical approach to validate the extracted visits.

Given n days of data, assume the automated process extracts $\{m_i\}_{i=1,n}$ bathroom visits for day i. Let's denote the k^{th} visit, $1 < k < m_i$, of day i by a sensor sequence $T_k=\{(C_{k1}\ t_{k1}),\ (C_{k2}\ t_{k2}),\dots,(C_{kp}\ t_{kp})\}$ where p is the length of the visit, $p \in N$, C_{ij} belongs to alphabet Σ, and t_{ij} are the time of the day of the firing C_{ij}. For each visit we define three parameters as duration of the visit in seconds (D), length of the visits that is the number of sensor hits in the visit (S), and the average time difference between consecutive hits in the visit (H_{AVG}) as formulated in equations 1-3.

$$D = |\ t_{kp} - t_{k1}\ | \tag{1}$$

$$S = p \tag{2}$$

$$H_{AVG}=\left(\left(\sum_{i=1}^{p-1} |t_{k(i+1)} - t_{ki}|\right)/p\right) \tag{3}$$

Assume we found out that parameters S, D, and H_{AVG} follow a *"LogLogistic"* distribution with parameters σ, μ, i.e $\{F_a(\sigma_a,\mu_a)\}_{a=S,D,H_{AVG}}$. The confidence that x with parameters $(x_S, x_D, x_{H_{AVG}})$ is a valid bathroom visit, $C(x)$ is 0 if $\{x_a < \min(F_a(\sigma_a,\mu_a))\}_{a=S,D,H_{AVG}}$ and equal to $\{Average(F_a(\sigma_a,\mu_a))\}_{a=S,D,H_{AVG}}$ where F_a is the likelihood that x_a comes from $F_a(\sigma_a,\mu_a)$ and is calculated using:

$$F_a(x_a,\sigma_a,\mu_a) = \frac{e^z}{\sigma_a x(1+e^z)^2}\ , \tag{4}$$

where $z = \frac{\log(x)-\mu_a}{\sigma_a}$.

B. Hierarchical Clustering on Routines

After retrieving valid bathroom visits, we use a hierarchical clustering approach to find repeated patterns [15]. In the clustering process we use TSW to measure the similarity of two sensor sequences [13]. Figure 2 shows the TSW algorithm. Essentially, TSW considers the time of the day between the sensor firings as a gap and computes the gap penalty using time stamps (as shown in equation 7 below). We use the "time of the day" metric, since we would like to find similar routine behaviors across different days that happen at roughly the same time.

Applying agglomerative hierarchical clustering, we seek to build a hierarchy of bathroom visit clusters. In this bottom-up approach, each observation is considered as a cluster, and pairs of clusters are merged as one move up the hierarchy. Figure 3 illustrates the hierarchical clustering with TSW.

$$H_{i0}=H_{0j} \qquad i\epsilon[1,n] \text{ and } j\epsilon[1,m]$$

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

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

$$Final\ Score=\frac{\max\{H_{ij}\}}{\min\{n,m\}}$$

Figure 2. Temporal Smith Waterman algorithm.

Algorithm: Hierarchical Clustering using TSW
Input: *N* sensor sequences, and an *N*N* similarity matrix;
Output: a dendrogram representing the hierarchical structure;
Steps:

1. Consider each sensor sequence as its own cluster.
2. Find the most similar pair of sensor sequences using TSW and merge them into a single cluster.
3. Compute distances between clusters as follows $D(C_i,C_j)=max\{1-TSW(S_{ik},S_{jk'})\}_{k=1,n,k'=1,m}$ where cluster i has n and cluster j has m sensor sequences.
4. Repeat steps 2 and 3 until all sensor sequences are clustered into a single cluster of size N.

Figure 3. Hierarchical clusterig using TSW algorithm.

IV. EXPERIMENTAL RESULTS

A. Bathroom Visit Data Set

We run a set of experiments to investigate the performance of the proposed routine detection based on frequent sensor sequences. Figure 4 shows three scatter plots of the duration and length of bathroom visits in terms of number of sensor hits in each visit (shown by their ids in x-axis) presented in Table III.

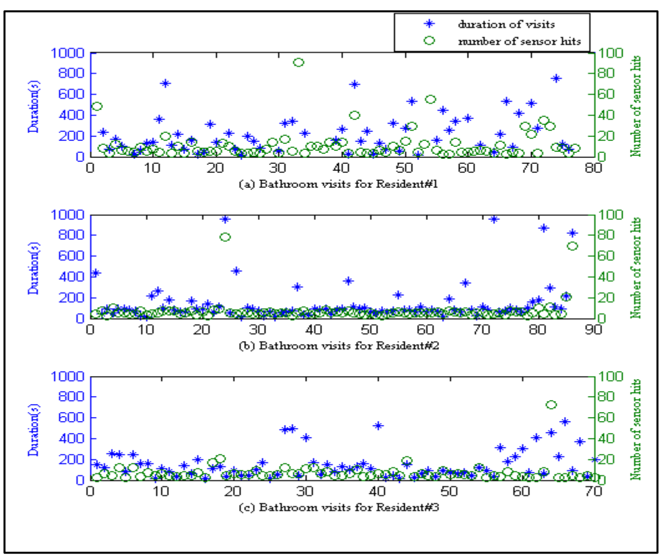

Figure 4. Bathroom visits scatter plots.

Bathroom visits range in duration from few seconds to about 16 minutes. The average duration for bathroom visit across people was 2.5 minutes (150 s) and the average number of sensor hits is 7.8. The average duration of for Resident #1 was 3.5 min (214.9 s), 1.87 min (112.5 s) for Resident #2, and 2.8 min (149 s) for Resident #3. While the majority of bathroom visits are shorter, possibly indicative of regular bowel movements, there are a few outliers that may represent more complex hygiene and showering. It is evident that each resident has their own pattern and distribution of visits.

B. Automated Routine Extraction Results

To verify the performance of the extraction process on the dataset presented in Table III, we applied the 10 cross-fold validation approach. We used a library of manually extracted sensor sequences of bathroom visits to calculate the parameters of the ROC curve, as shown in Figure 5. Table IV presents the area under the curve (AUC) separately for each of the residents. The average recognition performance is 0.84. Residents #2 and #3 use a walker that affects their speed reflected in the sensor hits. Therefore, the extraction process has a lower specificity, compared to Resident #1. Modeling walk patterns with walker for elderly resident is a complicated task that increases the false alarm rate of our model.

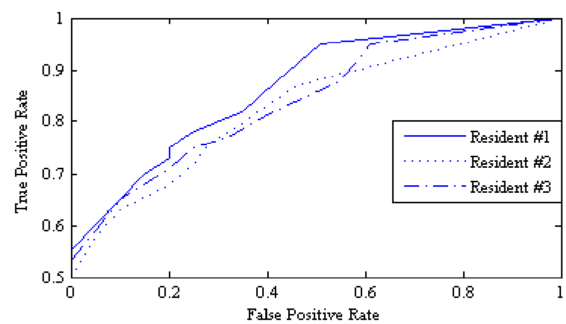

Figure 5. ROC curve of automated bathroom visit extraction.

TABLE IV. AREA UNDER ROC CURVES

AUC #1	AUC #2	AUC #3
0.86	0.83	0.84

C. Hierarchical Clustering Results on Bathroom Routines

We initialized the parameters of TSW as g=0 and c=0.028, setting the relevant time interval at about 35 seconds [13]. Figure 6 visualizes the results of the hierarchical clustering using (1-TSW) as a distance measure. The color map visually represents the strength of the relationships between bathroom visits for each resident separately. We can see that Resident #1 has a large number of small clusters, while Resident #2 has at least two large and distinct clusters (red and blue rectangles). Meanwhile Resident #3 has a mix of large and small clusters.

The number of routine clusters may be related to the functional status of the older adult. Resident #1 is the youngest and most active, which allows him to have a less rigid routine, with bathroom visits that are more varied in time. Also he is the only one taking a diuretic medication that promotes urine production, and in turn may affect the quality of bathroom visits. Both Residents #2 and #3 have urinary problems and may be more mindful of their needs. They use walkers that may force them to have more regimented and planned bathroom trips.

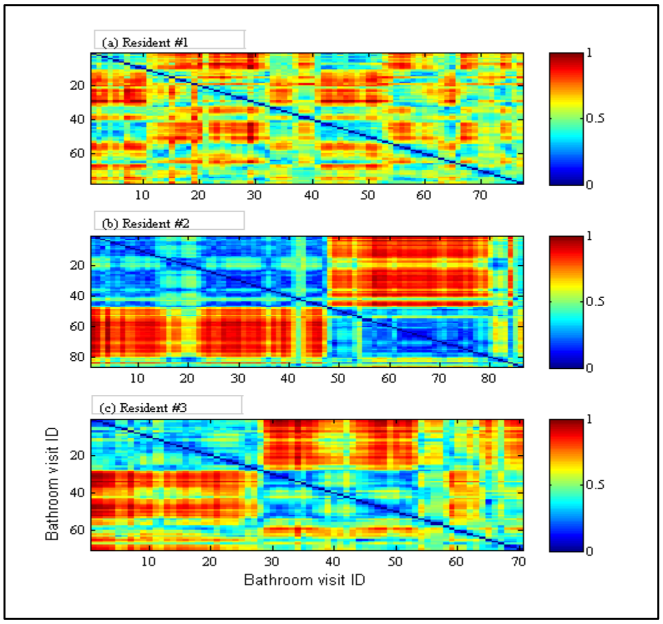

Figure 6. Hierarchical clustering results.

In Figure 7, the dendrogram shows clusters of bathroom visits based on parameters D, S and H_{avg} from eq. 3 (the height of each U represents the distance between the two objects being connected). Based on these characteristics we can identify types of bathroom visits and predict future changes. For example, for Resident #3 the first cluster (dashed red box) includes approximately 1.5 minute visits that happen in the afternoon when the resident is more active. Now that we validated the extraction and identification algorithm, we can explore this relationship between routines and functional abilities with a larger number of residents in future studies.

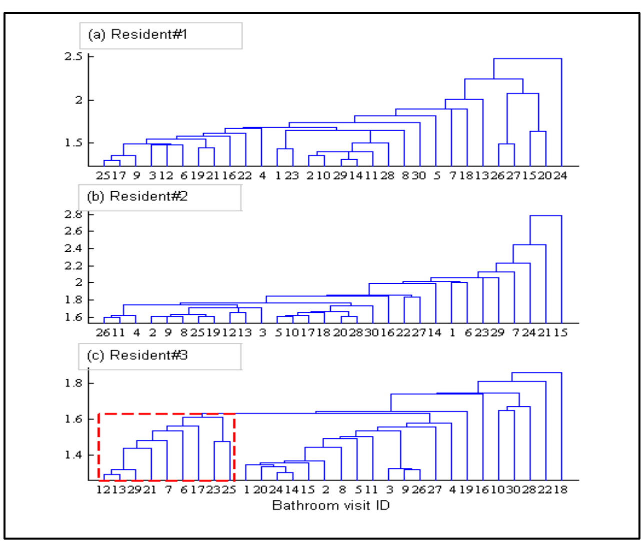

Figure 7. Hierarchical dendrograms.

V. CONCLUSION

Our computational algorithm successfully extracted routine bathroom visits of three TigerPlace residents. Using a new hierarchical clustering approach we detected frequent sensor sequences with an average recognition rate of 0.84, supporting the robustness of the proposed method. Identified clusters of routine behaviors may be related to the functional status of the older resident. Our approach tackles the problem with the big data collected by in-home monitoring systems by conceptualizing target periodic patterns as daily routines. In the future, we propose to use this algorithm to detect other routines, such as mealtimes and social activities. Our computational algorithm can be incorporated into the existing alert system for the healthcare providers to recognize early functional changes reflected in deviations from daily routines.

ACKNOWLEDGMENT

This project has been funded by a NSF SHWB grant, award #IIS-1115956 and the John A. Hartford Foundation's NHCGNE Award Program.

REFERENCES

[1] Rantz, M.J., et al., "Improving nurse care coordination with technology". CIN:Computers, Informatics, Nursing, 2010. 28, p.325-332.

[2] Vogeli, C., et al., "Multiple Chronic Conditions: Prevalence, Health Consequences, and Implications for Quality, Care Management, and Costs". Journal of General Internal Medicine, 2007. 22, pp. 391-395.

[3] Zisberg, A., H.M. Young, and K. Schepp, "Development and psychometric testing of the Scale of Older Adults' Routine". Journal of Advanced Nursing, 2009. 65, pp. 672-683.

[4] Tompa, M., "Finding motifs using random projections", The 5th Int'l Conference on Computational Molecular Biology, Montreal, Canada, 2001, pp.67-74.

[5] Zhenhui Li, J. H., "Mining Periodicity from Dynamic and Incomplete Spatiotemporal Data", Data Mining and Knowledge Discovery for Big Data, 2014, pp.41-81.

[6] Wang. W., Y. J., "Meta-patterns: revealing hidden periodical patterns", Proceedings of the 1st IEEE International Conference on Data Mining, 2001, pp.550-557.

[7] Indyk, P. K., "Identifying representative trends in massive time series data sets using sketches", The 26th Int'l Conference on Very Large Data Bases, Cairo, Egypt, 2000, pp. 363-372.

[8] Bill Chiu, E. K., "Probabilistic Discovery of Time Series Motifs", KDD '03 Proceedings of the ninth ACM SIGKDD Int'l conference on Knowledge discovery and data mining, 2003, pp. 493-498.

[9] Amy McGovern, D. H., "Identifying predictive multi-dimensional time series motifs: an application to severe weather prediction", Data Min. Knowl. Disc, 2011, pp. 232–258.

[10] Lin J, K. E., "Experiencing SAX: a novel symbolic representation of time series", Data Min. Knowl. Discov., 2007, pp. 107–144.

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

[12] Hajihashemi Z, Popescu. M., "Predicting Health Patterns Using Sensor Sequence Similarity and NLP", IEEE International Conference on BIBM, Philadelphia, USA, 2012, pp. 948-950.

[13] Hajihashemi Z., Popescu M., "An Early Illness Recognition Framework Using a Temporal Smith Waterman algorithm and NLP", AMIA Annual Symposium, 2013, pp. 548–557.

[14] Hajihashemi Z., Popescu M., "Detection of Abnormal Sensor Patters in Eldercare", The 4th International Conference on E-Health and Bioengineering- EHB '13, Iasi, Romania, 2013, pp. 1-4.

[15] Johnson S. C., "Hierarchical Clustering Schemes", Psychometrika, vol. 32, Issue 3, 1967, pp. 241-254.