Human behavior drift detection in a smart home environment

Andrea MASCIADRI[1], Anna A. TROFIMOVA, Matteo MATTEUCCI and Fabio SALICE
Department of Electronics, Information and Bioengineering
Politecnico di Milano - via Anzani 42, 22100, Como, ITALY

Abstract. The proposed system aims at elderly people independent living by providing an early indicator of habits changes which might be relevant for a diagnosis of diseases. It relies on Hidden Markov Model to describe the behavior observing sensors data, while Likelihood Ratio Test gives the variation within different time periods.

Keywords. Behavioral drift detection, smart home, HMM, Likelihood ratio test.

1. Introduction

The rapid aging of the world population is a phenomenon that threats worldwide health care systems. According to the technical report of the United Nations [1], between 2015 and 2030, the number of elderly is projected to grow by 56 percent. The challenge posed to worldwide Social Health Systems is to extend the period from the moment when a person, getting older, passes from independent living to the need of care-giving services. According to Becker et al. [2], this goal can be achieved enhancing Emergency Treatment Service in Smart Environments; indeed, the recognition of behavioral changes could be an effective solution to detect critical conditions, predict them on early stages, and alert caregivers [3].

It is challenging to define what a "human behavior drift" is, indeed, the behavior of a person constantly deviates due to personal factors such as aging, physical and mental well-being, and external environmental factors. According to Elbert et al. [4], these changes affect the way Activities of Daily Living are performed and thus human behavior. In this work, the behavior of a user is characterized by the sequence of activities that he/she performs, which are detected from Smart Home data through probability-based algorithms [5] that can be found in the literature with the name Activity Recognition (AR) [6, 7]. We defined behavioral drift as a long-term (gradual) deviation of the schedule and performance of Activities of Daily Living (ADLs). This definition aims to identify a term representing the observable effect of a possible condition that affects the human behavior which might threaten healthy (independent) living. Moreover, the contribution of the term "gradual" allows us to focus on drifts characterized by a long duration, e.g., a long decline that leads a depressed person to isolate itself from the world in possibly one year.

[1] Corresponding Author.

2. Background

Several papers discuss about the importance of monitoring the behavior of an individual in relation to his/her well-being [8, 9]. Much less attention has been given instead to identifying long term changes in the behavior from sensors data [10, 4].

Hung et al. [11] have presented an abnormality detection model based on Support Vector Machines and Hidden Markov Models to define a decision boundary to distinguish between normal and abnormal human behaviors. Elbert et al. [4] have presented a method able to compare information about the performed ADLs in one day, against those performed the previous one to detect long-term deviations in behavior. Their proposal combines three approaches: a cosine analysis based on the theory of circadian rhythms, a history-based approach based on movement data and a probabilistic model of behavior based on ADLs. Unfortunately, no quantitative result about the performances of this method has been reported in the literature.

Most of the publicly available datasets [12, 13, 14] that can be used for a proper quantitative evaluation of new methods have been collected with no explicit drift and for a relatively short period; thus, they are not suitable for Behavioral Drift Detection since it is required to have a dataset with an annotated Behavioral Drift that occurs on a widespread period. Considering this lack of data, an alternative approach for the dataset collection phase could be the generation of synthetic data with the consequent reduction of time and costs. Several works have been done in the field of Agent Based Modeling and Smart Environment Simulations [15].

In this work, we decided to use a Simulator of Human Activities, ROutines and Needs (SHARON) [16] that it can be easily configured to fits our requirements. SHARON is a tool developed in the frame of the BRIDGe project [17] to face the lack of data for AR. The simulator has been designed to be able to reproduce the typical ADLs scheduling of a person inside his/her dwelling; the cross-validation process applied on a real-world dataset (ARAS [12]) showed good results as reported by Veronese et al. [16].

3. Methods

According to the definition of behavioral drift that has been previously introduced, this phenomenon consists of a long-term gradual change in human behavior. This article is mainly focused on computing a drift detector as a measure of the intensity of behavioral changes between two periods, omitting every further qualitative analysis. This differential indicator is a first step to assess the emergence of a behavioral drift.

Let a *sensorset* be the set of sensors statuses in the house in each instant of time, θ is a Hidden Markov Model [18], where each node corresponds to a sensorset, and arches represent the transition probabilities among them. This model can be easily computed over a portion of the dataset in a given period and can be used as a surrogate of the model of the behavior. Thus, an indicator of Behavioral Drift is obtained comparing the sensorset transition model θ evaluated in two different time periods: if a drift occurs in the time that elapses between these two observed periods, an alarm should be rise.

The Likelihood Ratio (LR) is a statistical test that can be used to measure the "distance" between two observations. This situation perfectly fits our requirement because it would be possible to train the model over the data observed in a given period and to compute the LR with respect to data observed in another period.

The Likelihood Ratio Test is a statistical test providing a mechanism for making quantitative decision on whether the hypothesis should be rejected or not. Given two different time periods T_1 and T_2, the model θ representing the behavior in T_1 and the data observations x_1 and x_2 respectively at T_1 and T_2, we would like to get whether the model θ fits the data x_2 or not. If this hypothesis is rejected, a behavioral drift has occurred.

Let θ be the sensorset transition model and P the related transition probability matrix computed in the period T_1, the likelihood ratio between the observations x1 and x2 can be easily computed through the log-likelihood $L_\theta(x_i)$ of a model θ. If the LR between observations x_1 and x_2 is small, then the hypothesis holds true, otherwise a behavioral drift occurred in the period between T_1 and T_2. Since there is no indication about the threshold to be imposed to reject the hypothesis, in the following section we will expand the test by introducing a confidence interval.

4. Results

Using the SHARON simulator a synthetic dataset has been generated for an overall duration of three years, reproducing a decline consistent with the characteristics of Alzheimer's Disease (AD). According to the literature, the relationship between sleep disorders and dementia has been widely observed by monitoring elderly people [19], providing very useful data to configure SHARON: after three years of disease progression, the time taken to perform complex tasks such as "Take a shower" increases by 20%, the rate at which the activity is performed fells by 15%, the time of night-time sleep passes from an average of continuative 8h per night to 4.5h fragmented up to 5 times. Three portions of 90 days data have been extracted from the original dataset: *Ref90* is constituted by the initial part of the simulation, *ND90* refers to a period before the appearance of the AD' symptoms, while *D90* is obtained after one year from the appearance of the first symptoms. Considering the amount of drift in the dataset Ref90 as a reference, the H0 hypothesis states that: "the observations of a given dataset are produced from a model equal to the model that produced the reference observations". If the LR is relatively low the hypothesis is accepted, otherwise a drift is detected. To set a threshold between high and low values of LR, four datasets using the same parameters used for Ref90 have been generated. Calculating for each of them the value of LR compared to the reference dataset, it is possible to define the Confidence Interval (CI) at 99% for the mean of LR. Since the number of observed samples is low (<15), a Student's distribution is used in the computation of the CI. Therefore, the hypothesis H0 is accepted for a dataset if its LR value falls within the found Confidence Interval.

As result, the log-likelihood ratio computed on two periods of 90 days and referring to a stable behavior, has been successfully classified as "No Drift detected" with a 99% confidence interval. Moreover, the 90 days of observation referring to a drifted behavior has been classified as "Drift detected" since the LR falls outside of the CI.

5. Conclusions

A very simple model representing the behavior of a person based on Home Automation Environment Response data is presented. The Likelihood Ratio has been introduced to assess the emergency level associated to a behavioral drift: this differential indicator is computed comparing sensors observation in different time periods. The proposed model can be enhanced considering the activities performed by the inhabitant to retrieve more information about the underlying cause of the drift. For this reason, a module of unsupervised activity recognition can be integrated in the current BDD system. Finally, as soon as real life dataset containing long-term behavioral drift will be published, a new evaluation of the system can be performed.

References

[1] United Nations. (2015). World population ageing 2015, technical report, Department of Economic and Social Affairs, Population Division, New York.

[2] Becker, M., Werkman, E., Anastasopoulos, M., & Kleinberger, T. (2006, November). Approaching ambient intelligent home care systems. In Pervasive Health Conference and Workshops, 2006 (pp. 1-10). IEEE.

[3] Veronese, F., Comai, S., Saidinejad, H., & Salice, F. (2015). Elderly monitoring and aal for independent living at home: Human needs, technological. Optimizing Assistive Technologies for Aging Populations.

[4] Elbert, D., Storf, H., Eisenbarth, M., Ünalan, Ö., & Schmitt, M. (2011, May). An approach for detecting deviations in daily routine for long-term behavior analysis. In Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on (pp. 426-433). IEEE

[5] Kim, E., Helal, S., & Cook, D. (2010). Human activity recognition and pattern discovery. IEEE Pervasive Computing, 9(1).

[6] Mehr, H. D., Polat, H., & Cetin, A. (2016, April). Resident activity recognition in smart homes by using artificial neural networks. In Smart Grid Congress and Fair (ICSG), 2016 4th International Istanbul (pp. 1-5). IEEE.

[7] Krishnan, N. C., & Cook, D. J. (2014). Activity recognition on streaming sensor data. Pervasive and mobile computing, 10, 138-154.

[8] Riboni, D., Bettini, C., Civitarese, G., Janjua, Z. H., & Bulgari, V. (2015, March). From lab to life: Fine-grained behavior monitoring in the elderly's home. In Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on (pp. 342-347). IEEE.

[9] Peetoom, K. K., Lexis, M. A., Joore, M., Dirksen, C. D., & De Witte, L. P. (2015). Literature review on monitoring technologies and their outcomes in independently living elderly people. Disability and Rehabilitation: Assistive Technology, 10(4), 271-294.

[10] Sprint, G., Cook, D., Fritz, R., & Schmitter-Edgecombe, M. (2016, May). Detecting Health and Behavior Change by Analyzing Smart Home Sensor Data. In Smart Computing (SMARTCOMP), 2016 IEEE International Conference on (pp. 1-3). IEEE.

[11] Hung, Y. X., Chiang, C. Y., Hsu, S. J., & Chan, C. T. (2010, June). Abnormality detection for improving elder's daily life independent. In International Conference on Smart Homes and Health Telematics (pp. 186-194). Springer Berlin Heidelberg.

[12] Alemdar, H., Ertan, H., Incel, O. D., & Ersoy, C. (2013, May). ARAS human activity datasets in multiple homes with multiple residents. In Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare (pp. 232-235). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[13] Cook, D., Schmitter-Edgecombe, M., Crandall, A., Sanders, C., & Thomas, B. (2009, April). Collecting and disseminating smart home sensor data in the CASAS project. In Proceedings of the CHI Workshop on Developing Shared Home Behavior Datasets to Advance HCI and Ubiquitous Computing Research (pp. 1-7).

[14] Van Kasteren, T., Noulas, A., Englebienne, G., & Kröse, B. (2008, September). Accurate activity recognition in a home setting. In Proceedings of the 10th international conference on Ubiquitous computing (pp. 1-9). ACM.

[15] Armac, I., & Retkowitz, D. (2007, July). Simulation of smart environments. In Pervasive Services, IEEE International Conference on (pp. 322-331). IEEE.

[16] Veronese, F., Masciadri, A., Trofimova, A. A., Matteucci, M., & Salice, F. (2016). Realistic human behaviour simulation for quantitative ambient intelligence studies. Technology and Disability, 28(4), 159-177.

[17] "Assistive technology group of Politecnico di Milano" http://www.atg.deib.polimi.it, 2015

[18] Häggström, O. (2002). Finite Markov chains and algorithmic applications (Vol. 52). Cambridge University Press.

[19] Cricco, M., Simonsick, E. M., & Foley, D. J. (2001). The impact of insomnia on cognitive functioning in older adults. Journal of the American Geriatrics Society, 49(9), 1185-1189.