Engineering Applications of Artificial Intelligence 45 (2015) 90–102

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Unsupervised discovery of activities of daily living characterized by their periodicity and variability

Julie Soulas*, Philippe Lenca, André Thépaut

Institut Mines-Telecom; Telecom Bretagne, UMR 6285 Lab-STICC, Technopôle Brest Iroise, CS 83818, 29238 Brest Cedex 3, Université Européenne de Bretagne, France

ARTICLE INFO

Article history:
Received 18 February 2015
Received in revised form
7 May 2015
Accepted 2 June 2015
Available online 15 July 2015

Keywords:
Activities of daily living
Regular patterns
Episodes
Periodicity
Variability
Aging at home

ABSTRACT

Habits characterize the activities of elderly people. Monitoring their habits and their ability to carry out the activities of daily living is a great challenge in order to improve aging at home. In particular, the detection of changes in regular behavior may help to detect emerging disorders. The emergence of smart homes and sensor networks allows the non-intrusive collection of data describing the activities in the home. The collected data is indeed an objective source to mine periodic patterns representing the habits of a particular individual. Extended Episode Discovery (xED) algorithm is described and discussed. This algorithm searches for regular patterns, highlighting the periodicity and variability of each discovered pattern. This approach allows a high adaptability to different users and lifestyles. Experiments on six real-life datasets illustrate the interest of xED.

1. Introduction

The population in many countries is getting older. For example, it is expected that in 2025, a third of the European population will be over 60 years old (Sölvesdotter et al., 2007). Meeting the needs of this aging population is thus one of the big challenges of this century. Elderly people are frailer and more prone to chronic diseases. Thus, it becomes more dangerous for them to live alone at home. Falls and malnutrition problems are frequently observed. These threats to their health can be a source of anxiety for them and their families. That is notably why families tend to push them to move to a nursing-home. Health is in fact the first reason why people over 75 move. However, most elderly people would rather continue to live in their own home (Kotchera, 2005). In addition, solutions enabling aging at home are also usually cheaper for the society than the funding of nursing homes. Helping people stay in their home longer and in better conditions is thus an active effort.

In order to reach this goal, different approaches are used, including the traditional home visits of medical staff and the use of technical aids, like support bars and emergency pendants. But thanks to the development of sensor technologies, Smart Home and Ambient Assisted Living (SHAAL) systems have gained much attention during the last decade, see for example the surveys (De Silva et al., 2012; Rashidi and Mihailidis, 2013), as well as the assistive smart home projects CASAS (Cook et al., 2013) and Domus (Pigot, 2010). SHAAL systems use sensors and other devices that are either wearable or integrated in the home infrastructure. They capture data describing health-related events as well as the Activities of Daily Living (ADLs). The ADLs are things we do as part of our daily life: feeding, taking care of one's personal hygiene, dressing, moving about, etc. Being able to perform them is a necessary condition for living independently and so for aging at home.

We focus here on the unsupervised discovery of living habits. Indeed, elderly people tend to follow daily routines (Bergua and Bouisson, 2008): habits help them remain in control of their environment. Changes in the routines may indicate that carrying out everyday activities is becoming harder, or even reveal the onset of a disorder, like Alzheimer's disease. Sensor data is thus used as a source for the discovery of interesting patterns describing the habits, i.e. the discovery of periodic patterns.

For that purpose, we proposed xED to discover periodic patterns (Soulas et al., 2013). We here deeply extend the analysis of the qualities and drawbacks of xED. We also propose and discuss experiments with five additional datasets. Next section describes some of the prominent current research on the use of sensor dataset mining for improving aging at home, with a focus set on the unsupervised approaches. Section 3 presents the formalism and definitions used for frequent episode mining. Episode Discovery algorithm is detailed in Section 4. Extended

* Corresponding author.
E-mail addresses: julie.soulas@telecom-bretagne.eu (J. Soulas), philippe.lenca@telecom-bretagne.eu (P. Lenca), andre.thepaut@telecom-bretagne.eu (A. Thépaut).

http://dx.doi.org/10.1016/j.engappai.2015.06.006
0952-1976/

Episode Discovery algorithm is extensively described and analyzed in Section 5. Section 6 presents experiments on six real-life datasets, as well as a comparison of the performances of xED with another relevant algorithm: Episode Discovery (ED, Heierman et al., 2004). Conclusions and perspectives are proposed in Section 7.

2. Related works

SHAAL systems aim to support elderly people with their daily activities, and to help them continue to live independently, healthily, and safely. Thus, they cover a large number of topics directly related to daily living, like fall detection (Botía et al., 2012; Mirchevska et al., 2013), recognition of the visits from family members or professional caregivers (Aicha et al., 2013), sleep pattern recognition (Ni et al., 2012), identification and prediction of abnormal behaviors (Jakkula et al., 2009; Roy et al., 2011; Lotfi et al., 2012; Meffre et al., 2014). SHAAL systems may also aim to fight against isolation (Bothorel et al., 2011), consider user acceptance, privacy, accuracy and cost issues (Chernbumroong et al., 2013), and make efforts to propose efficient designs of sensor networks for the reduction of energy consumption (Francisco et al., 2014).

Popular data mining algorithms have been widely investigated, in particular state-space based models, such as Dynamic Bayesian Network (Philipose et al., 2004) and Hidden Markov Models (Patterson et al., 2005). Gu et al. (2009) use emerging patterns as powerful discriminators to differentiate sequential, interleaved and concurrent activities. Fleury et al. (2010) successfully apply support vector machines to classify activities into a predefined set of seven activities. Focusing on dry cooking, Hirano et al. (2013) used decision tree classifiers to predict doneness for waffles and popcorn. The use of ontologies for the dynamic inference of the context and the choice of an adapted reaction was also investigated (Allégre et al., 2012; Aloulou et al., 2014).

Despite their success, none of the above work is sufficient for all needs and encountered situations. As a consequence there is an increasing trend to propose integrated frameworks which include various machine learning tools (Chen et al., 2010; Cao et al., 2012; Ugolotti et al., 2013). Moreover, the combination of machine learning tools with domain knowledge is also an interesting perspective to increase the systems performance (Luštrek et al., 2012; Simonin et al., 2015).

Unsupervised analysis: A very large proportion of the aforementioned algorithms rely on supervised learning. They require annotated data, which is hard to get in the context of SHAAL systems (Cleland et al., 2013): two main strategies are currently used. In the first one, the inhabitant is asked to write down the name of every activity and the time when it was undertaken. In the second one, cameras are installed in the home, and someone later watches the video footprints and annotates the activities. Either method is very demanding and time-consuming. They are also intrusive into the user's daily life, and prone to annotation errors. This leads to an increasing interest for unsupervised algorithms for the discovery of frequent and periodic activities. The next paragraphs present some unsupervised approaches found in recent literature.

Rodner and Litz (2013) use association rules: the recorded events are preprocessed to generate transactions of items. This preprocessing uses data extraction, aggregation and transformation in order to enrich the raw events with geographic data; the numeric data and dates are discretized, etc. For instance, a transaction corresponding to an event "`2015-04-24 7:12:50 Kitchen light on`" could be {`between 6 and 8`, `light in kitchen`, `activity in kitchen`, `followed by kitchen motion detector`, `Friday`}. The transactions are then mined within the association rules framework to extract frequent and confident rules. This approach depends a lot on the expressiveness of the transactions, and hence on the richness of the preprocessing.

Table 1
Summary of the characteristics of unsupervised approaches for activity monitoring in SHAAL systems.

Bibliographic reference	Objective for the end user	Technical keywords	Episode Discovery	Periodicity	Inclusion of temporal variability	Criteria for candidate construction	Pattern quality assessment measure
Rodner and Litz (2013)	Describe data with association rules	Association rules, data enrichment in preprocessing	No	No	preprocessing: timestamp discretization	Support, confidence	Support, confidence
Nazerfard et al. (2010)	Describe activity occurrences and their succession	Clustering, GMM[a], Temporal association rules	No	No	**yes**: with GMMs	Inter-occurrence temporal distance	Confidence
Rashidi et al. (2011)	Discovery of activity patterns (for unsupervised classification)	Episode Discovery, Clustering of sequences, MDL[b]	**Yes**: sequential episodes	No	No (timestamps are not taken into account)	Frequency, edit distance, description length (MDL)	Description length
Salah et al. (2013)	Activity segmentation and classification	Segmentation, Clustering of sequences, HSMM[c]	**Yes**: sequential episodes	No	No (timestamps are not taken into account)	Geographic / contextual segmentation (thresholds)	Feature vector similarity
Heierman et al. (2004)	Periodic episode discovery and description	Episode Discovery, Interval cycles, MDL	**Yes**: parallel episodes	**Yes**	Preprocessing: timestamp discretization	Episode length and capacity	Description length
xED	Periodic episode discovery, periodicity description	Episode Discovery, DBSCAN, EM[d], GMM, periodicity analysis	**Yes**: parallel episodes	**Yes**	**Yes**: clustering and description with GMM	Episode duration, frequency	Accuracy, compression power (MDL)

[a] Gaussian mixture model.
[b] Minimum description length.
[c] Hidden semi-Markov model.
[d] Expectation-maximization.