Pervasive and Mobile Computing 28 (2016) 51–68

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

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

Modeling patterns of activities using activity curves

Prafulla N. Dawadi[a,*], Diane J. Cook[a], Maureen Schmitter-Edgecombe[b]

[a] School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA, United States

[b] Department of Psychology, Washington State University, Pullman, WA, United States

ARTICLE INFO

Article history:
Available online 1 October 2015

Keywords:
Activity curve
Smart environments
Functional assessment
Permutation

ABSTRACT

Pervasive computing offers an unprecedented opportunity to unobtrusively monitor behavior and use the large amount of collected data to perform analysis of activity-based behavioral patterns. In this paper, we introduce the notion of an *activity curve*, which represents an abstraction of an individual's normal daily routine based on automatically-recognized activities. We propose methods to detect changes in behavioral routines by comparing activity curves and use these changes to analyze the possibility of changes in cognitive or physical health. We demonstrate our model and evaluate our change detection approach using a longitudinal smart home sensor dataset collected from 18 smart homes with older adult residents. Finally, we demonstrate how big data-based pervasive analytics such as activity curve-based change detection can be used to perform functional health assessment. Our evaluation indicates that correlations do exist between behavior and health changes and that these changes can be automatically detected using smart homes, machine learning, and big data-based pervasive analytics.

1. Introduction

Many pervasive computing applications such as home automation, activity aware interventions, and health assessment require analyzing and understanding activity-based behavioral patterns. The performance of such applications depends on the ability to correctly learn a model of general daily activity behavior from a large amount of data and be able to predict when such daily behavior is likely to continue or change. These big data-based approaches to activity modeling can then in turn be used to provide effective activity-aware services such as improved healthcare.

Activity recognition lies at the heart of any pervasive computing approach to modeling behavioral routines. An activity recognition algorithm maps a sensor reading or sequence of readings to a corresponding activity label. In order to answer general questions related to daily activity patterns, such information needs to be transformed to a higher-level representation. For example, questions such as how average daily activity patterns have changed over a year, or generally what hours did a particular individual sleep last month are difficult to answer using raw output from activity recognition algorithms. However, many pervasive computing applications such as home automation and health assessment require answering such questions.

Obtaining higher-level representations or models of activities has several additional advantages. Higher-level representations can abstract variations in day-to-day activity routines. For example, wake-up times in the morning may be slightly different each day even if the overall routine is fairly stable. Additionally, such representations simplify the task of modeling an

* Corresponding author.
E-mail addresses: pdawadi@eecs.wsu.edu (P.N. Dawadi), cook@eecs.wsu.edu (D.J. Cook), schmitter-e@wsu.edu (M. Schmitter-Edgecombe).

http://dx.doi.org/10.1016/j.pmcj.2015.09.007
1574-1192/

individual's daily routine and at the same time make visualization and interpretation of daily activity routines easy. Collecting big datasets over long periods of time allows us to abstract activity models over such daily variations. As we will demonstrate in this paper, such representations aid with the process of identifying long-term changes in a behavioral routine.

For example, consider the following description highlighting aspects of an individual's routine at two different points in time:

- *Month of March* 2012: Sleep at 10:00 PM, get up at 6:00 AM, eat breakfast at 7:00 AM, eat lunch at 12:00 PM, go out for a walk at 4:00 PM, and dine at 8:00 PM.
- *Month of September* 2013: Sleep at 8:00 PM, wake up frequently during the night, get up at 10:00 AM, no breakfast, eat lunch at 11 AM, no going out for a walk, and dine at 7:00 PM.

Note that each of these sample activity-based descriptions is aggregated over a one-month period and therefore describes a general routine that is maintained over a prolonged period of time. Based on these descriptions we also note changes in the routine from the first observation to the second. From this example, we can infer that by September 2013 the observed individual was experiencing disturbances in sleep, was skipping meals, and stopped exercising. Determining if the overall daily activity patterns has changed may be difficult based only on the raw sensor data or even based on event-by-event labels from an activity recognition algorithm. Such questions can be more easily answered by comparing two higher-level representations of these activity patterns.

In our current work, we propose a novel ***activity curve*** to model an individual's generalized *daily activity routines*. The activity curve modeling algorithm uses activity-labeled sensor events to learn a higher-level representation of the individual's regular routine. These activity labels are *automatically-recognized* using an activity recognition algorithm. We also introduce a Permutation-based Change detection in Activity Routine (PCAR) algorithm to compare activity curves between different time points in order to detect changes in an activity routine. To validate our algorithm, we make use of longitudinal smart home sensor data collected by monitoring everyday behavior of residents over two years. Finally, we demonstrate how the activity curve and the PCAR algorithm can be used to perform important pervasive computing tasks such as automated assessment of an individual's functional health.

2. Related work

The work that we describe in this paper is unique in its ability to automatically characterize behavioral routines based on recognized activities and to detect changes in generalized routines over time. However, other work has focused on alternative approaches to recognize and discover daily activity routines and two-sample tests to detect changes between two sample populations similar to the core component of the proposed PCAR algorithm. Additionally, other studies have investigated the relationship between changes in activity patterns and changes in health using the statistical and visualization techniques.

Discovering activity routines: Researchers have studied the problem of automated discovery and recognition of daily activity routines using the data collected from wearable sensors [1,2], GPS signals [3] and mobile phones [4] using algorithms such as topic modeling [1] and collaborative filtering [5]. In these approaches, raw sensor data are converted to a bag-of-words representation which contains the histogram of activity label occurrences or histogram of location/proximity information. Data from wearable sensors can be used to discover daily routines such as *having lunch* and *brushing teeth* [1]. Similarly, data from mobile phones can be used to recognize routines such as *commuting to office* and *working*.

In contrast to these earlier works that focus on discovering and recognizing daily activity routines from sensor data, our focus is to model the discovered and recognized daily activity routines using an activity curve model. The proposed activity curve model is a generic model that can be calculated both from the output of an activity recognition algorithm as well as using the algorithms mentioned in the aforementioned studies. Furthermore, our proposed model facilitates answering more complex questions related to activity routines such as whether changes in an activity routine have occurred or not. In our current work, the proposed activity curve model uses the output from an activity recognition algorithm.

Visualization of activity routines: Researchers have proposed visualization techniques to visualize daily activity patterns. For example, Galambos et al. [6,7] developed methods to visualize activity level, time spent away from home, deviations in activities of daily living, and behavioral patterns. Similarly, other researchers have developed techniques to visualize deviations in activity routines and behavioral patterns using smart home sensors [8,9]. These methods provide a tool to understand sensor-monitoring data and to study daily activity routines. However, these approaches rely on manual inspection of the data in order to make any higher-level conclusions regarding daily routines.

Two-sample test: The two-sample test is a widely used statistical analysis tool to compare between two sample populations. Classical two-sample tests such as the t-test are used to compare the means of two populations having the same or different variances. However, the t-test is a parametric test that is limited to comparing between two Gaussian distributions. Other examples of non-parametric classical versions of two-sample tests are the Wald–Wolfowitz runs test, the Anderson–Darling test and the Kolmogorov–Smirnov test [10].

Recently, Maximum Mean Discrepancy (MMD) was proposed as another non-parametric two-sample test technique [11]. MMD compares the means of two distributions in a universal reproducing kernel Hilbert space and has superior performance to several of the classic two-sample tests. However, the superior performance of MMD relies on a valid choice of a kernel and