Discovery of Personal Processes from Labeled Sensor Data – An Application of Process Mining to Personalized Health Care

Timo Sztyler[1], Johanna Völker[1], Josep Carmona[2],
Oliver Meier[1], Heiner Stuckenschmidt[1]

[1]University of Mannheim, Germany
{timo,johanna,heiner}@informatik.uni-mannheim.de
[2]Universitat Politècnica de Catalunya, Spain
jcarmona@cs.upc.edu

Abstract. Currently, there is a trend to promote personalized health care in order to prevent diseases or to have a healthier life. Using current devices such as smart-phones and smart-watches, an individual can easily record detailed data from her daily life. Yet, this data has been mainly used for *self-tracking* in order to enable personalized health care. In this paper, we provide ideas on how process mining can be used as a fine-grained evolution of traditional self-tracking. We have applied the ideas of the paper on recorded data from a set of individuals, and present interesting conclusions and challenges.

1 Introduction

Physical inactivity is a major risk factor for certain types of diseases. Indeed, physical activity does not only prevent or relieve diseases, but also improves public health and well being [2]. In this context, personalized health solutions and lifestyle monitoring can help to ensure that people doing the right activity at the right time. However, the regular use of such methods is critical to achieve the desired result. Hence, barriers for the adoption must be low, and using both software and devices should be as comfortable as possible.

Thanks to the technological progress in the development of wearable devices, sensor technology, and communication, we are nowadays able to setup a body sensor network based on smart-phones, smart-watches, and wristbands which does not affect people during their daily routine. In contrast, most of the available software requires substantial user input to specify, e.g., the current activity or even vital parameters like the heart rate or blood pressure.

We want to develop an application which monitors the personal lifestyle of the users and provides appropriate visualizations. However, this still needs a sufficient acceptance because the user has to view and interpret the visualizations. Therefore, we also want to provide automatically generated recommendations resulting from the monitoring data and, e.g., references (practical guidance). In the long term, we also have to automatically recognize a person's daily activities such as different types of sports and desk work. This is necessary to ensure that the required user input is a minimum which also is a requirement to make the application practical.

Due to the fact that the activities of a user can be seen as process instances, process mining can help us to elicit and analyze these processes. It allows discovering a process model from an event log focused on personal activity, and combined with, e.g., conformance checking, to explore deviations with respect to reference models. The results could be useful in the context of monitoring to provide a meaningful feedback but also to create recommendations.

In this paper, we present the data set we created for our first experiments (see Section 2), and we outline initial ideas about how process mining could help us to address our main use cases (see Sections 3, 4, and 5):

Monitoring We want to help users to monitor their personal behavior by presenting them a daily or weekly visual summary of their personal processes. This summary could highlight behavior which is unknown or unconscious to the user. As a result, the user could correct the behavior.

Deviations We want compare their personal processes with reference processes to detect deviations. This allows making suggestions regarding the procedure of certain activities, and point out missing activities. As a result, the user learns to optimize the daily routine in respect of a healthy lifestyle.

Operational Support Historical data that combines both activity and environmental data (e.g., geographic position) can then be used for the operational support based on individual's process models, enabling predictions and recommendations in order to accomplish certain goals.

We do not deal with activity recognition but address succeeding problems. The created data set is a training data set with manually labeled data. Commonly, machine learning techniques are used for activity recognition [9]. Therefore, the data set can be used to build or evaluate activity recognition systems, but in the following we want to use the result of such a system in combination with process mining to create personal processes by using the manually created activity labels. The resulting personal process models should allow to benefit the users health by making visualizations, recommendations, and predictions.

2 Data Gathering

This section provides the details of the data set used in this paper. The data set can be obtained by contacting us.

General Settings. Seven individuals (age 23.1±1.81, height 179.0±9.09, weight 80.6±9.41, seven males) collected Accelerometer, Orientation, and GPS sensor data and labeled this data simultaneously (see Table 1). The data was collected using a smart-phone and smart-watch combined with a self-developed sensor data collector and labeling framework (see Figure 1). The subjects were not supervised but got an introduction and guidelines. The subject group covers five students, a worker, and a researcher.

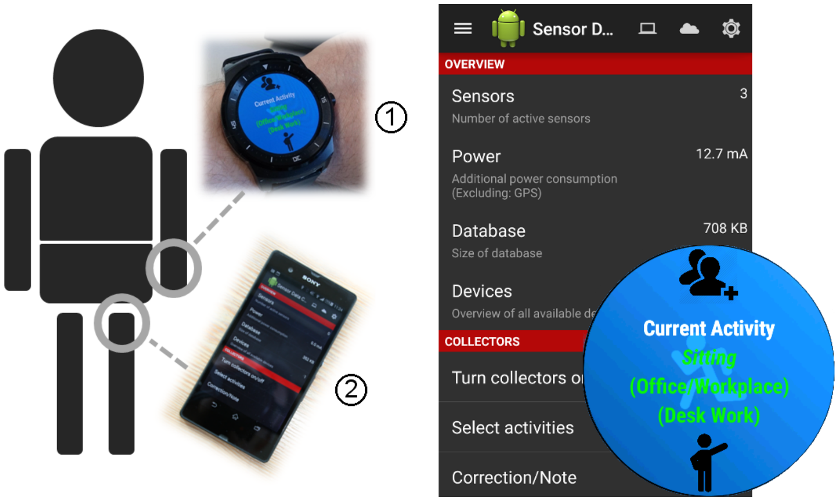


Fig. 1. Collector and labeling framework: Wear App (smart-watch, 1) and Hand App (smart-phone, 2). The positions of the devices may vary.

Setup	
Subjects	7 males
Devices	Smart-phone (2) and Smart-watch (1)
Sensors	Acceleration (50Hz), Orientation (50Hz), GPS (every 10 min.)
Labels	Activity, Device Position, Environment, Posture
Storage	Local Database, SD-Card
Duration	10 hours a day, 12 days

Table 1. Equipment and Settings of the data gathering.

Devices and Labeling. The framework consists of a *Wear* (1) and *Hand* (2) application which interact with each other via Bluetooth. The *Wear* application allows updating the parameters (see Table 1) immediately where the *Hand* application manages the settings of the sensors and the storing of the data. The sampling rate (50Hz) was chosen with consideration of battery life as well as with reference to previous studies [12, 19]. Table 1 summaries the equipment and settings.

The individuals should collect data during their daily routine and it was up to them to decide where the device should be positioned on the body. We focused on the activity, device position, environment, and the posture which occur during the daily routine. The values for these parameter were predefined (see Tables 2 and 3) and could not be changed or extended.

Activities. The activity labels allow recording the daily routine. We focused on food intake, sport, different type of movements, but also (house) work so that we can compare the daily routine of several individuals to detect common activity

Parameter	Values
Device Position	Chest, Hand, Head, Hip, Forearm, Shin, Thigh, Upper Arm, Waist
Environment	Building, Home, Office, Street, Transportation
Posture	Climbing, Jumping, Lay, Running, Sitting, Standing, Walking

Table 2. Labeling parameters that were updated immediately when the *device position*, *environment*, or *posture* had changed.

Activity	Sub-Activity
Desk Work^1	*n/a*
Eating/Drinking	Breakfast, Brunch, Coffee Break, Dinner, Lunch, Snack
Housework	Cleaning, Tidying Up
Meal Preparation^1	*n/a*
Movement	Go for a Walk, Go Home, Go to Work
Personal Grooming^1	*n/a*
Relaxing	Playing, Listen to Music, Watching TV
Shopping^1	*n/a*
Socializing	Bar/Disco, Cinema at Home
Sleeping^1	*n/a*
Sport	Basketball, Bicycling, Dancing, Gym, Gymnastics, Ice Hockey, Jogging, Soccer
Transportation	Bicycle, Bus, Car, Motorcycle, Scooter, Skateboard, Train, Tram

Table 3. Activity and sub-activity labels. The subjects had to select at least one of these activity labels to specify their current action. The selection of a sub-activity is optional but allows to be more precise. ^{1}Please note, that there are activities without sub-activities.

patterns but also to analyze the different behaviors. The set of activity labels was minimized and structured to decrease the time which the individual needs to decide and choose a suitable label. Thus, there are 12 activities and 32 sub-activities where an activity could be “Eating/Drinking” and a corresponding sub-activity “Breakfast”[1]. It is possible to select several activity labels at the same time to record the current situation with a high accuracy (e.g., “Movement - go to Work”, “Transportation - Train”, and “Sleeping”). Thus, the individual can describe the current situation from several points of view.

To keep the set of activity labels as small as possible, we provided some generic labels such as “Desk Work”. This label should be used if the individual works in an office (worker), attends a lecture or class room (student), or visits a school (pupil). During the introduction phase, we explained this to the individuals to avoid that they choose different labels in the same situation.

[1] Please note: So far, we do not consider the sub-activities in the presented use-cases.

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14	D15	D16	D17	D18	D19	D20	D21	D22	D22	D23
S1											▨	▨	▨	▨		▨	▨	▨	▨	▨				
S2								▨	▨	▨	▨	▨			▨	▨	▨	▨	▨					
S3		▨	▨			▨	▨	▨	▨	▨		▨	▨		▨	▨		▨						
S4	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨									
S5		▨	▨	▨	▨		▨		▨	▨	▨		▨	▨	▨	▨	▨	▨						
S6		▨		▨																				
S7									▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨

Table 4. Timetable. Overview of how many and which days where recorded for each individual. The X-axis represents the [D]ays whereas the Y-axis illustrates the [S]ubjects. The grey colored day labels (D[0-9]+) are weekend days. The grey squares indicate that data was recorded.

Profiling. We recorded 74 cases which cover 1, 386 events. A case is represented by one individual in one particular day and has an average duration of 12.1 hours. The events comprise the activities and sub-activities which were performed by the individuals. Table 4 describes when and how long each individual recorded data. Tables 5 and 6 illustrate the related recorded data. The number of records of acceleration and orientation differs, because one subject selected a lower frequency for the orientation sensor. The high standard deviation of the numbers of postures results from the different behavior of the individuals. Hence, some individuals move a lot (e.g., walking, standing, walking) while others label the posture less accurate (e.g., standing just for a second).

Labels	**Records** (avg±sd)
Activities	20 ± 7
Postures	80 ± 62
Environment	16 ± 4
Dev. Position	8 ± 6

Table 5. Annotated labels per day and individual.

Raw Data	**Records** (absolute)
Acceleration	$2.7 * 10^6$
Orientation	$2.3 * 10^6$
geo. Location	70

Table 6. Number of recorded values per day and individual.

3 Use Case 1: Monitor Personal Behavior

Since a picture is worth a thousand words, the deployment of graphical representations of event data may open the door to a precise awareness of the activities carried out by an individual. We believe graphs are a strong visualization aid to understand aggregated behavior, and thus consider this direction as the first use case for understanding personal activity data.

Interesting information a user can get periodically (every day or week) is the personal process model that describes the main activities and their dependencies.

In this process model, one can find frequent sequences of activities, alternatives, concurrency (moving while eating) and so on. This deviates from the typical information that is provided by current tools for self-tracking individuals. In general, such tools focus only on showing correlations between the tracked variables (e.g., eating vs. sport) or the evolution of single variables (weight over the week). In this section, we take the training data that were described in the previous section and illustrate how traditional process discovery techniques can be used to elicit the personal process model of an individual. The preliminary conclusions reported in this section should not be considered as a general rule but instead are meant to illustrate the capabilities of process discovery techniques in providing a fresh look for self-tracking.

3.1 Focusing on the Frequent Paths

Due to the variability in personal activity data, there is not a simple process model that represents all possible paths for an individual, even for the reduced number of individuals monitored in this paper. In this section, we focus on the most frequent paths taken by each individual. To this end, the discovery of *fuzzy models* [8] using the Disco tool [4] is considered. The reason for using a frequency-based discovery technique is to handle the variability and noise of a self-tracking log. Alternative techniques like the *heuristic miner* [18] or the *inductive miner* [10] which can be applied in this scenario may be considered as well.

To illustrate the potential of a personal process model with respect to analyzing tons of raw data, we focus on two simple aspects: the difference in activity between work and weekend days on the one hand, and the differences across the individuals on the other hand.

During the week vs. weekend. Figures A.1 and A.2 (see Appendix), show the main activity models during the working week and the weekend, respectively. The process models depicted in the figures have a very different structure. This clearly denotes a variation in the personal activity during the week and weekend, when considering the main activity by individuals. For instance, while in the week days the main behavior is centered towards “Desk Work” which is also the most frequent activity, the frequency of paths and activities is more balanced in the weekends. This tendency is also satisfied in the average duration of activities (not shown in the process models).

Personal activity across users. Figure A.3 (see Appendix) shows each individual’s main activity models. As it was explained in the previous section, three types of individuals were monitored: student (5 instances), researcher (1 instance) and worker (1 instance). Although the details of the models are not visible in this figure, one can see significant differences across individuals. Commonalities between students are also elicited in the models, for example, the global tendency to structure the model around “Desk Work” and the well-structured relation between the activities for most of the students.

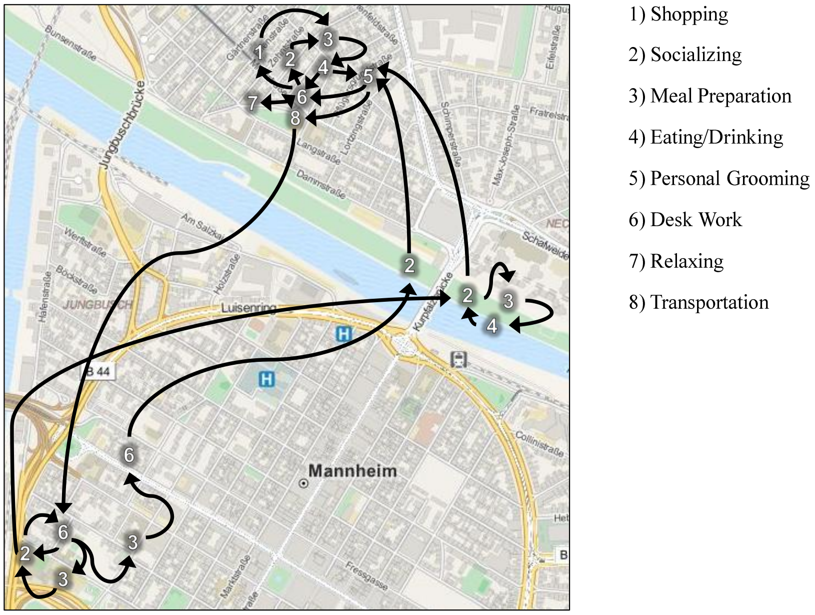


Fig. 2. Main personal activity for an individual including geographical position data: numbers correspond to different activities, and arcs denote control-flow relations extracted from the activity data.

3.2 Model Enhancement Using Personal Data

As shown in Section 2, not only activity data is stored from individuals but also important data like the geographical position, acceleration and, orientation of the device. In the following, as an example, we explain how to combine the control-flow process models (e.g., see Figure A.1) with the geographical position data to derive *personal activity-position maps*. This kind of map illustrates geographically the control-flow with respect to the real geographical position of activities. Figure 2 depicts an example of such a map for the data gathered from one of the individuals. The computation of personal activity-position maps can be done by simply aligning the timing information (*start*, *end*) recorded for each activity event with the one obtained from the geographical position of individuals. This way, for every activity, its geographical position in a case will be extracted. Events corresponding to the activity name will be then analyzed to compute a set of locations that represents the different locations where the activity has been carried out. For instance, in Figure 2, activity 2 ("Socializing") has four different nodes in the graph. Ideally, to have a simpler graph, only one location per activity is desired. The locations for an activity can be computed by clustering the set of locations with a fixed radius of k meters and selecting the centroids, or by using the frequency of locations, or a combination of both. Finally, the nodes corresponding to each activity in a certain location are displayed on top of a real map, the area of which corresponds to the minimal enclosing box that includes all locations depicted. Arcs from the control-flow are then routed from the corresponding locations in the map.

The personal activity-position maps are strongly related to trajectory pattern mining [20]. A trajectory pattern consists of chronologically ordered geographical locations combined with the duration. The provided algorithms allow to detect frequent behaviors in space and time (daily, weekly), and in this context to aggregate movement behavior of a person [7] or a group [14] [11] to keep track on specific movements. This facilitates to discover highly frequented places as well as underlying patterns in movements which might be related to other persons, and can help to identify semantic relations between persons [11]. Related to this, a previous work [13] explored the principle limitations of predicting human dynamics based on mobility patterns of smart-phone users.
Concerning our scenario, we focus on the daily routine of a person and the related activities which means that we have to connect the spatiotemporal information explicitly with the activity information. If we can combine this information and apply the mentioned techniques then it could help to influence the daily routine of a person in terms of achieving a healthier life by optimizing specific kind of patterns. Considering health care, the kind of transportation between locations might be also important in the context of energy consumption but this is not covered by the mentioned techniques (see [7]).

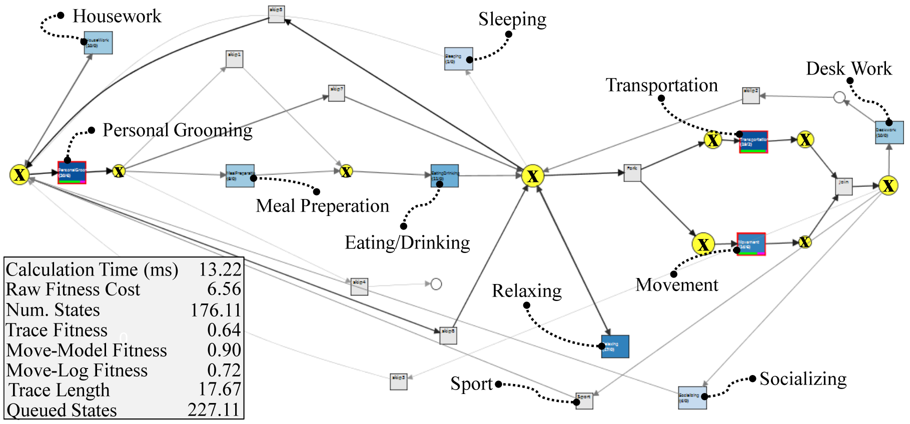
4 Use Case 2: Deviations from Reference Models

Self-tracking may be a meaningful way to verify if certain requirements with respect to reference quantities are accomplished. For instance, many associations advise to do at least 30 minutes of moderate physical activity per day or eat fish at least twice a week. Those guidelines for a good lifestyle offer a rough description for individuals, mainly concerning about quantities and frequencies. However, some ways of satisfying these guidelines are probably less healthy than others, e.g., it may not be the best decision to eat fish while doing physical activity.

Hence, there may be reference models that describe precisely how activities should be carried out in order to satisfy a guideline. Thus, the reference model has to provide the opportunity to describe certain actions in a specific *order* (e.g., “Sport” should be followed by “Personal Grooming”), should allow explicit *choices* (e.g., after “Desk Work” only “Eating/Drinking”, “Socializing”, or “Transportation” are expected actions) and should also consider *concurrency* actions. (e.g., “Transportation” and “Movement” may be overlapping activities).

Reference models can be obtained in several ways. One possibility would be to ask a domain expert to create manually the desired reference model for a given goal. A second option would be to collect event logs from successful individuals. These logs can be combined with the introduced techniques of the previous section to discover a reference model. Finally, a third option would be to translate the

[3] http://www.promtools.org



Calculation Time (ms)	13.22
Raw Fitness Cost	6.56
Num. States	176.11
Trace Fitness	0.64
Move-Model Fitness	0.90
Move-Log Fitness	0.72
Trace Length	17.67
Queued States	227.11

Fig. 3. Example of fitness analysis in ProM[3]of an individual with respect to a reference model: places with yellow background (X) represent situations where the individual deviates from the process model. Transitions without a label denote silent events not appearing in the event log.

textual guidelines into process models, using recent techniques that apply *natural language processing* to elicit process models [6].

When a reference model is available, *conformance checking* techniques can be applied to assess the adequacy of the reference process model in representing the traces of individuals [15]. Since the reference model describes the ideal behavior, it is meaningful to focus the analysis on the *fitness* of the reference model with respect to the traces of individuals. A process model fits a given trace if it can reproduce it. An example of such analysis can be seen in Figure 3 where an individual is analyzed with respect to an invented process model meant to represent a healthy behavior.

Fitness checking can also be extended to consider other perspectives, i.e., costs or quantities for additional event data [5]. For instance, one typical advice on dietary guidelines is *to eat as many calories as one burns* [1]. These kind of checks can be incorporated into the reference model by using the data conformance approach from [5]. Therefore, deviations on quantities can also be verified with respect to the reference model.

If reference models are not available, simple rules can be used which should be satisfy by individuals on their daily routine. These rules may describe patterns that should satisfy an individual, e.g., "taking medicines" should be followed by "eating". This can be formally specified with Linear Temporal Logic (LTL) formulas to be satisfied by the event log of activities [17].

5 Use Case 3: Operational Support

Historical data of an individual is a rich source of information which may be crucial to influence the daily routine in order to reach a particular goal. In this context, process models can be enhanced and used at each decision point to assess the influence of the next step in satisfying the targeted goal. For instance, following the guideline of the previous section that advice to eat as many calories as one burns, activities can be annotated with respect to calorie levels (e.g., "Eating/Drinking" produces an amount of calories while "Movement" takes an amount of calories). Then, historical activity data can be aggregated with this information to learn for all decision points the impact of the decision regarding the likelihood of satisfying the targeted goal, e.g., the balanced consumption of calories. Figure 4 shows an example for the case of the balance of calories in a diet, i.e., states (nodes) are labeled with the probability of reaching a balanced diet at the end of the day.

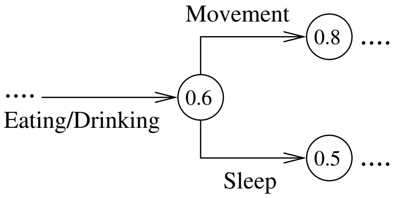


Fig. 4. Excerpt of a state-based prediction model for balance of calories. The nodes illustrate the probability for reaching the balance.

Thus, when an individual is about to start a new activity, recommendations can be provided on the basis on the model's aggregated data corresponding to the current state. This deviates from current prediction and recommendation practices that do not consider the current state of the model explicitly.

The precision of the prediction may vary due to the fact that the available information can have a different granularity. Hence, events can carry information such as the amount of calories but also only cover complete cases with the resulting label (e.g., good, satisfactory, medium, bad). In such a case, standard techniques [15] for the operational support of process models can be applied to predict and recommend the next steps.

6 Future Work

In the following, we outline a few general directions of future work and possible next steps.

When process mining is applied, e.g., to identify and visualize the most frequent paths, it should take into account a given hierarchy of activities and subactivities. Such a hierarchy could facilitate, for instance, the aggregation of collected data on different levels of abstraction.

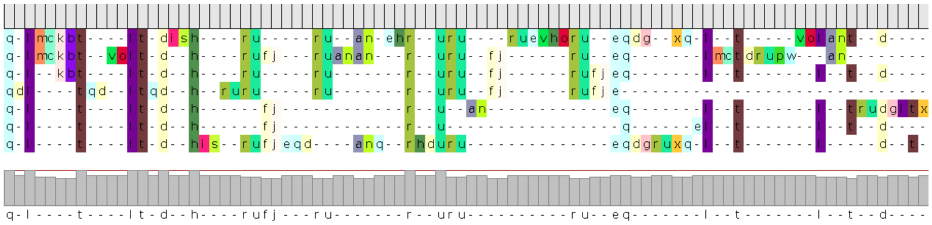


Fig. 5. Example of discovered trace cluster: letters in the bottom denote activities with high consensus. The Y-axis represents seven different traces where the X-axis illustrates the different events per traces.

Future applications of process mining might also require dealing with uncertain data. In particular, the data generated by classification-based methods for activity recognition will most probably be uncertain, since these methods are never a hundred percent accurate. However, provenance information such as explicit uncertain values will be available in most cases, and might serve as an additional input to process mining methods.
Further directions include the investigation of more expressive process models. For example, reference models, which describe an ideal sequence of daily routines, should include information about frequencies, time and locations.
Finally, we would like to bootstrap activity recognition by creating and leveraging synergies between activity recognition and process mining techniques. A possible bootstrapping approach would generate process models from automatically recognized activities, and use the resulting process models to improve the accuracy of the activity recognition.

More concrete ideas that we would like to investigate are the following:

Exploring the log via trace alignment. Section 3.1 focused on the main activity paths followed by individuals, thus ignoring less frequent behavior that may mislead the conclusions. An alternative will be to preprocess the log with the goal of extracting patterns, and then transform the log accordingly, either by introducing hierarchy, or by ignoring outlier activities not following the learned patterns. For this purpose, *Trace alignment* techniques from [3] can be applied. For instance, in Figure 5 seven traces have been aligned together from the log of workdays.

Process Cubes. Recently, *process cubes* have been proposed also as a means to apply process mining in a exploratory manner, similar to *online analytical processing* (OLAP) techniques [16]. The intuitive idea is to mine event logs by restricting events under a particular perspective. For example, extracting a process model for activity in a bank, focusing only on clients from a given region that got married within the last three years. With the data available in a personal activity context, process cubes can be a promising way to slice the data and mine particular contexts. For instance, one can be interested in process models where “Desk Work” is mainly situated in a given location.

7 Conclusions

This paper discusses challenges and opportunities for process mining in the area of personalized health care. We described the acquisition of a real-world data set consisting of manually labeled sensor data from smart-phones, and outlined interesting use cases. We then took a look at existing methods for eliciting, analyzing and monitoring individuals' daily routines, and described the results of our preliminary experiments. We presented our ideas on future directions and challenges in this application context which may require significant advances with respect to algorithmic support for process mining.

Acknowledgments. This work as been partially supported by funds from the Spanish Ministry for Economy and Competitiveness (MINECO) and the European Union (FEDER funds) under grant COMMAS (ref. TIN2013-46181-C2-1-R).

A Appendix

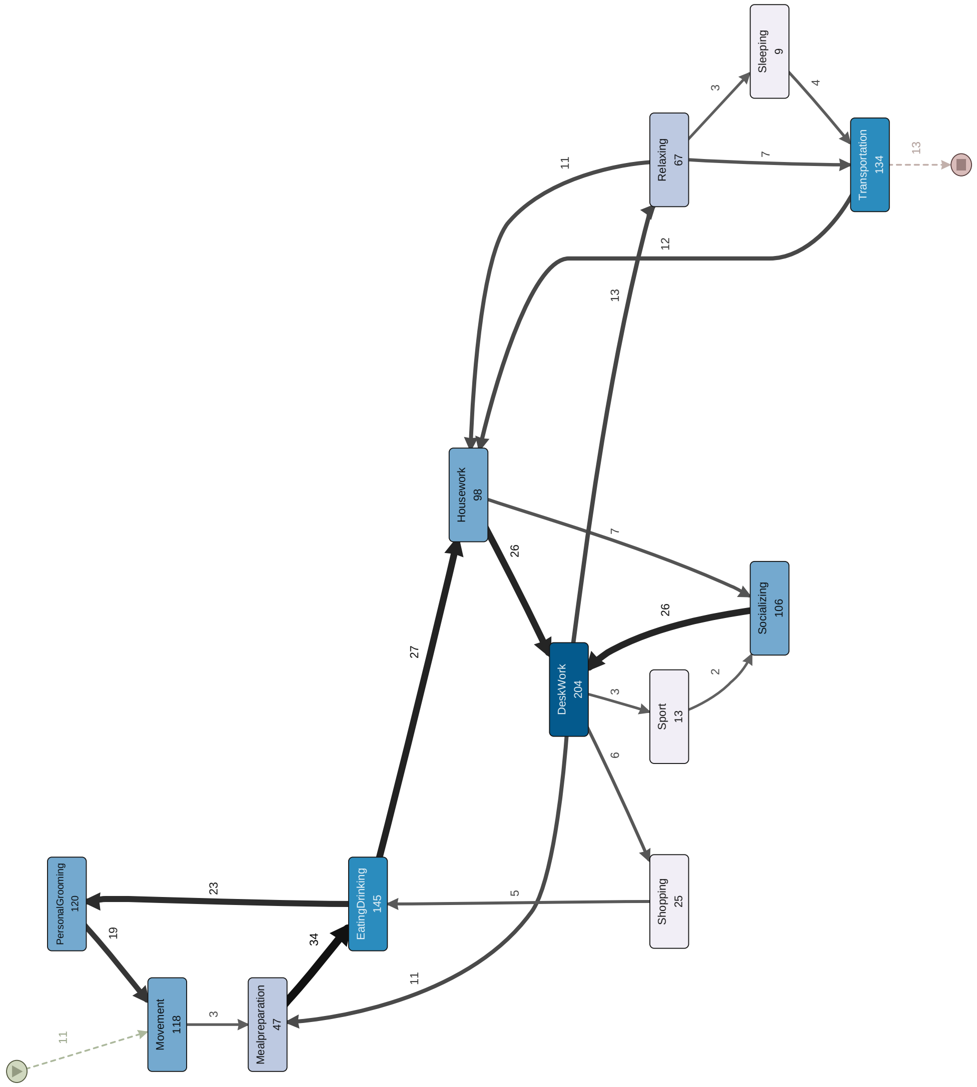


Fig. A.1. Main personal activity for all the users during the working week days (57 cases).

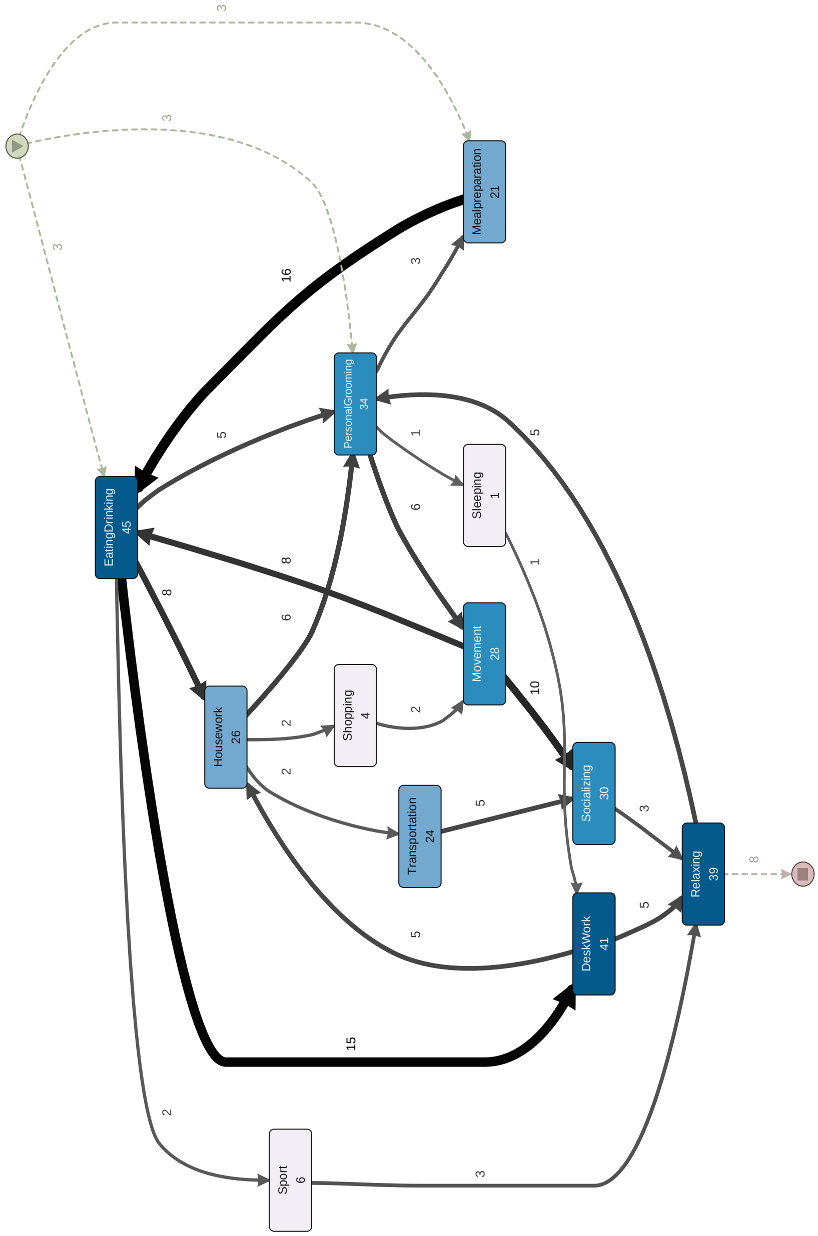
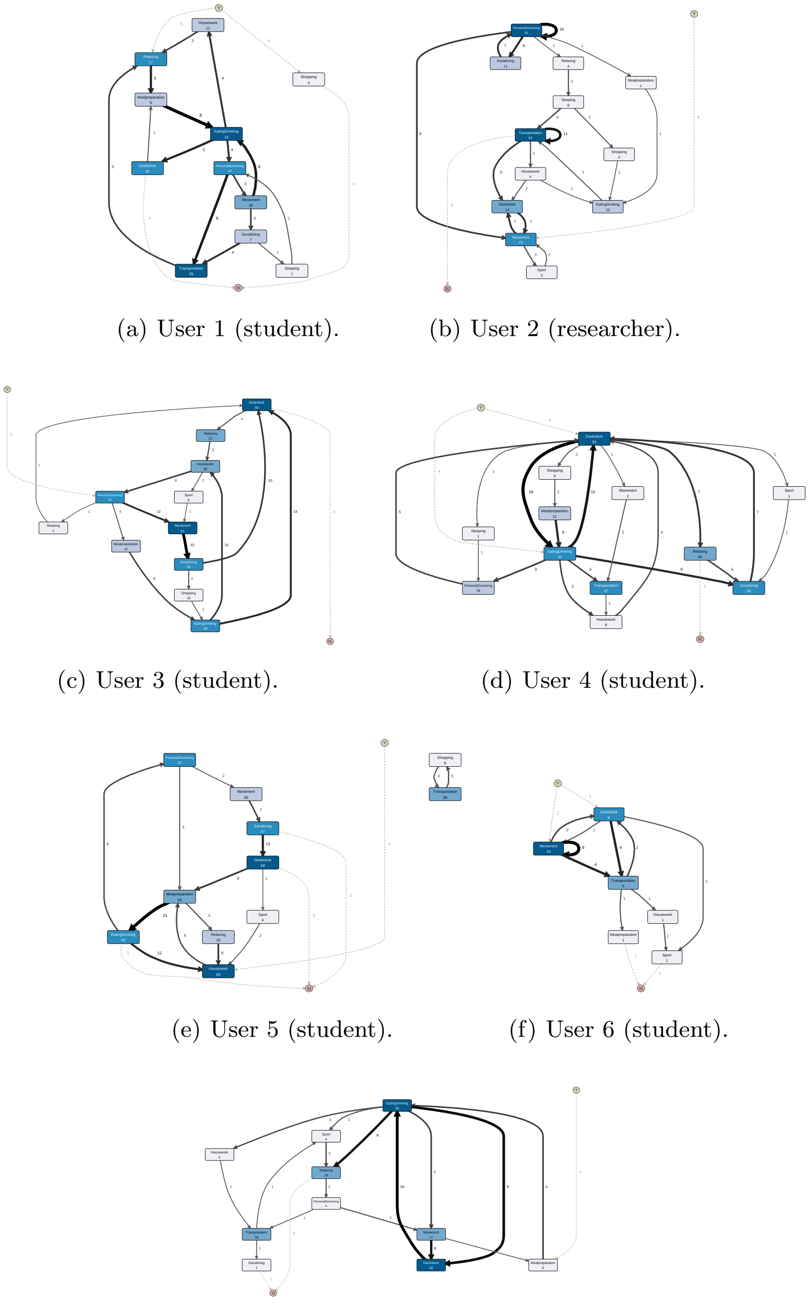


Fig. A.2. Main personal activity for all the users during the weekend days (17 cases).



(a) User 1 (student).

(b) User 2 (researcher).

(c) User 3 (student).

(d) User 4 (student).

(e) User 5 (student).

(f) User 6 (student).

(g) User 7 (worker).

Fig. A.3. Main personal activity by users.

References

1. American Heart Association. http://www.heart.org. Last Access: 29.04.2015.
2. Steven N Blair and Tim S Church. The fitness, obesity, and health equation: is physical activity the common denominator? *Jama*, 292(10):1232–1234, 2004.
3. RP J. C. Bose and Wil MP van der Aalst. Process diagnostics using trace alignment: Opportunities, issues, and challenges. *Inf. Syst.*, 37(2):117–141, 2012.
4. Disco by Fluxicon. https://fluxicon.com/disco/. Last Access: 29.04.2015.
5. M. De Leoni and W.M.P. van der Aalst. Aligning event logs and process models for multi-perspective conformance checking: An approach based on integer linear programming. In *Business Process Management, China*, pages 113–129, 2013.
6. F. Friedrich, J. Mendling, and F. Puhlmann. Process model generation from natural language text. In *Advanced Information Systems Engineering - 23rd International Conference, CAiSE 2011, London, UK, 2011. Proceedings*, pages 482–496, 2011.
7. Fosca Giannotti, Mirco Nanni, Fabio Pinelli, and Dino Pedreschi. Trajectory pattern mining. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 330–339. ACM, 2007.
8. C.W. Günther and W.M.P. van der Aalst. Fuzzy mining - adaptive process simplification based on multi-perspective metrics. In *BPM*, pages 328–343, 2007.
9. O.D. Lara and M.A. Labrador. A survey on human activity recognition using wearable sensors. *Communications Surveys & Tutorials, IEEE*, 15(3):1192–1209, 2013.
10. Sander J. J. Leemans, Dirk Fahland, and Wil M. P. van der Aalst. Discovering block-structured process models from event logs containing infrequent behaviour. In *Business Process Management Workshops - BPM 2013 International Workshops, Beijing, China, August 26, 2013, Revised Papers*, pages 66–78, 2013.
11. Zhenhui Li. Spatiotemporal pattern mining: Algorithms and applications. In *Frequent Pattern Mining*, pages 283–306. Springer, 2014.
12. Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L Littman. Activity recognition from accelerometer data. In *AAAI*, volume 5, pages 1541–1546, 2005.
13. Chaoming Song, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. Limits of predictability in human mobility. *Science*, 327(5968):1018–1021, 2010.
14. Hsiao-Ping Tsai, De-Nian Yang, and Ming-Syan Chen. Mining group movement patterns for tracking moving objects efficiently. *Knowledge and Data Engineering, IEEE Transactions on*, 23(2):266–281, 2011.
15. Wil M. P. van der Aalst. *Process Mining - Discovery, Conformance and Enhancement of Business Processes.* Springer, 2011.
16. Wil M. P. van der Aalst. Process cubes: Slicing, dicing, rolling up and drilling down event data for process mining. In *Asia Pacific Business Process Management, Beijing, China*, pages 1–22, 2013.
17. Wil M. P. van der Aalst, H. T. de Beer, and Boudewijn F. van Dongen. Process mining and verification of properties: An approach based on temporal logic. In *CoopIS, Cyprus*, pages 130–147, 2005.
18. A.J.M.M. Weijters, W.M.P. van der Aalst, and A.K. Alves de Medeiros. Process mining with the heuristics miner-algorithm. Technical Report WP 166, BETA Working Paper Series, Eindhoven University of Technology, 2006.
19. A.Y. Yang, S. Iyengar, S. Sastry, R. Bajcsy, P. Kuryloski, and R. Jafari. Distributed segmentation and classification of human actions using a wearable motion sensor network. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, pages 1–8. IEEE, 2008.
20. Yu Zheng. Trajectory data mining: an overview. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6(3):29, 2015.