Continuous sleep profiles clustering with a novel 2–step functional data approach

Zuzana Rošťáková[1], Roman Rosipal[1], Georg Dorffner[2]

[1]Department of Theoretical Methods,
Institute of Measurement Science,
Slovak Academy of Sciences,
Bratislava, Slovakia

zuzana.rostakova@savba.sk
roman.rosipal@savba.sk

[2]Section of Artificial Intelligence and Decision Support,
Center for Medical Statistics, Informatics and Intelligent Systems,
Medical University of Vienna,
Vienna, Austria

georg.dorffner@meduniwien.ac.at

Introduction

The influence of sleep on humans daily life behavior is already well–known. In our long–term research we aim to identify specific temporal sleep profiles which reflect important physiological aspects of sleep. Instead of the classical Rechschaffen and Kales sleep model we focus on the probabilistic sleep model (PSM), which characterizes the sleep process by probability values of 20 sleep states, which we call sleep microstates. By considering the probability values of a given sleep microstate as a function of sleep time we obtain a curve. In this work we study functional data clustering methods for detecting groups of subjects with similar curve-profiles. When curves are not synchronized in time, classical clustering techniques like k–means may assign some curves into incorrect clusters. The second major problem is associated with methods operating on cluster–typical night profiles. Profiles computed on misaligned data are in general poor cluster representatives. Therefore, to overcome these problems of clustering sleep microstates curves we apply our recently developed 2–step iterative approach which iteratively combines the cluster analysis and curve registration steps.

Curve alignment problem

Let X and Y represent curves defined over the time interval T. To register (in time align) curves X, Y means to find a strictly increasing warping function $h : T \to \mathbb{R}$ which minimizes a chosen similarity criterion, for example

$$\int_T \left(X(t) - (Y \circ h)(t)\right)^2 dt$$

In this work we use the Self–Modelling Time Warping (SMTW, [1]) algorithm, which good performance has been proved in practice.

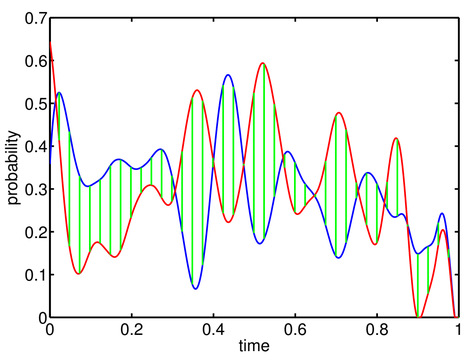


Figure 1: Example of the two misaligned curves with similar overall profile. Green lines between corresponding time points indicate that the squared area between the curves is quite large (equal to 0.04).

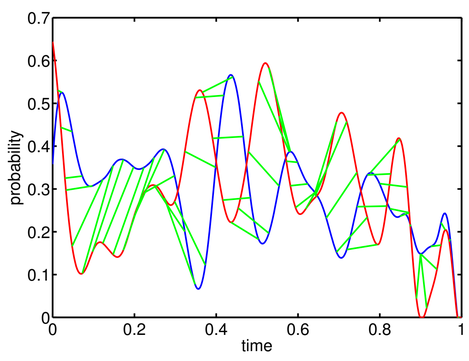


Figure 2: Points on the curves from Fig. 1 connected according to the overall curves dynamics regardless to the real time. The Dynamic Time Warping algorithm (see below block) was used here to match the corresponding curve points.

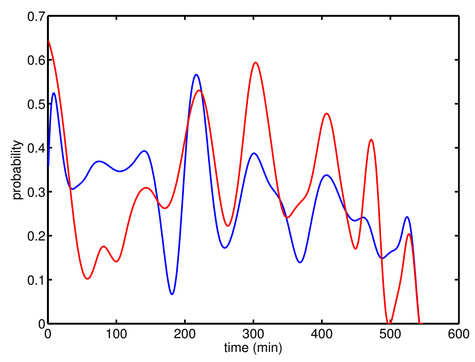


Figure 3: The curves from Fig. 1 synchronized in time by the Self–Modelling Time Warping algorithm. Similarity between the curves profiles is more evident and the square area between the curves decreased to 0.01.

2–step approach

Let $\{X_1, \ldots, X_N\}$ be a set of misaligned probabilistic sleep curves defined over the time interval T and represents the set which we would like to divide into K clusters. We apply an iterative clustering approach consisting of the following steps:

1. Assign the curves into K clusters using the distance matrix

   $$M_{dtw} = \{dtw\,(X_i, X_j)\}_{i,j=1,\ldots,N}$$

   as an input for the k–medoids algorithm. The dtw measure defined by eq. (1) takes into account possible curves time misalignment and therefore is more appropriate than the point-to-point Euclidean distance.

2. Separately align curves in each cluster and denote the aligned curves as $X_1^\star, \ldots, X_N^\star$. For curves alignment we use the SMTW method, although an arbitrary registration algorithm could be chosen and this can be done according to a given structure of curves. To guarantee that the warping function h is strictly increasing and to avoid alignment of distant time segments an additional restriction is considered within the chosen registration method.

3. Compute the average similarity within the formed clusters $C_1, \ldots, C_K$

   $$L = \sum_{i=1}^{K} \frac{1}{|C_i|} \sum_{j:X_j^\star \in C_i} \int_T \left(X_j^\star(t) - \mu_i(t)\right)^2 dt \qquad \mu_i(t) = \frac{1}{|C_i|} \sum_{j:X_j^\star \in C_i} X_j^\star(t) \qquad t \in T$$

4. If the number of iterations exceeds 100 or $L < \varepsilon$, where ε is a small given constant, stop. Otherwise repeat the algorithm with the registered curves $X_1^\star, \ldots, X_N^\star$.

Dynamic Time Warping

Dynamic Time Warping (DTW, [2]) is a member of a wider area of registration methods. For discrete observations of two curves

$$X_1 = (X_1(t_1), \ldots, X_1(t_{n_1}))$$
$$X_2 = (X_2(s_1), \ldots, X_2(s_{n_2}))$$

the goal of DTW is to find the optimal warping path

$$w = \left\{(i_l, j_l),\ \begin{array}{l} i_l \in \{1, \ldots, n_1\}, \\ j_l \in \{1, \ldots, n_2\}, \end{array}\ l = 1, \ldots W\right\}$$

which minimizes the sum of distances between matched points

$$Q_{X_1,X_2}(w) = \sum_{(i_l,j_l)\in w} d\left(X_1(t_{i_l}), X_2(t_{j_l})\right)$$

where W is the length of the warping path w and d is Euclidean distance.
Instead of curve alignment we use DTW as a similarity measure of misaligned curves

$$dtw(X_1, X_2) = \min_w Q_{X_1,X_2}(w) \qquad (1)$$

References

[1] D. Gervini and T. Gasser. Self-modeling warping functions. *J. R. Statist. Soc. B*, 2004.

[2] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. In *IEEE Trans. Acoust. Speech Signal Process*, volume 26, pages 43–49, 1978.

[3] G. Klösch, B. Kemp, T. Penzel, A. Schlögl, P. Rappelsberger, E. Trenker, G. Gruber, J. Zeitlhofer, B. Saletu, W.M. Herrmann, S.L. Himanen, D. Kunz, M.J. Barbanoj, J. Röschke, A. Varri, and G. Dorffner. The SIESTA project polygraphic and clinical database. *Medicine and Biology Magazine*, 20(3):51–57, 2001.

[4] R. Rosipal, A. Lewandowski, and G. Dorffner. In search of objective components for sleep quality indexing in normal sleep. *Biological Psychology*, 94(1):210–220, 2013.

[5] D.J. Greene, A. Barnea, K. Herzberg, A. Rassis, M. Neta, A. Raz, and E. Zaidel. Measuring attention in the hemispheres: the lateralized attention network test (LANT). *Brain Cogn*, 66(1):21–31, 2008.

[6] A. Lewandowski, R. Rosipal, and G. Dorffner. Extracting more information from EEG recordings for a better description of sleep. *Computer method and programs in biomedicine*, 108(3):961–972, 2012.

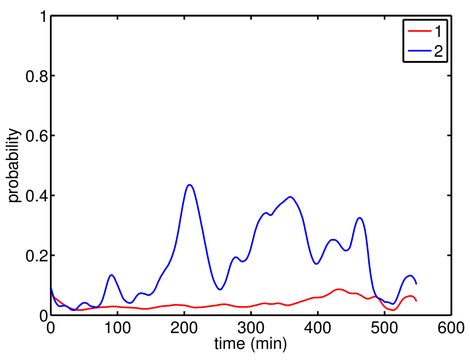
Acknowledgement

This research was supported by the Ministry of Health of the Slovak Republic, grant number MZ 2012/56-SAV-6 and by the Slovak Research and Development Agency, grant number APVV-0668-12.

Results

To compare our 2–step approach with standard k–means raw data clustering we consider two sleep datasets. The PSM was trained and applied to each dataset separately. As a preprocessing step, probabilistic sleep curves of each sleep microstate were aligned considering the sleep latency and they were smoothed with ??? .

The first dataset consists of the PSG recordings of 146 healthy subjects each spending two consecutive nights in the sleep lab. This cohort represents a subset of healthy sleepers from the SIESTA database [3]. Except of that, the SIESTA database contains results of a set of questionnaires about sleep and awakening quality, tests for assessment of memory and motor activity, and the physiological blood pressure and pulse rate measures [4].



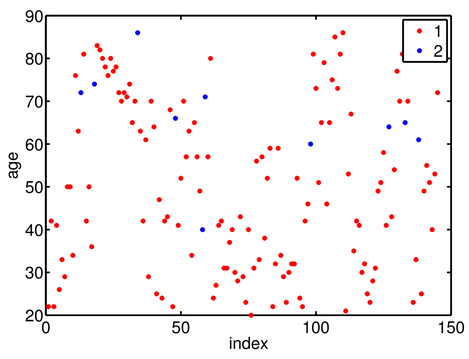
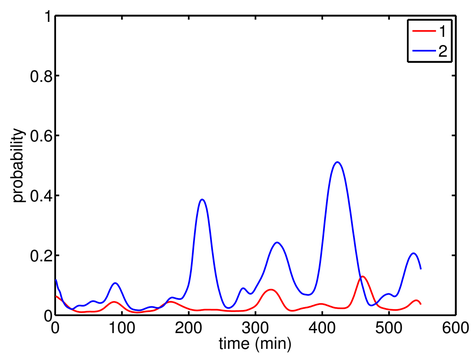


Figure 4: Cluster analysis of a microstate similar to the *Wake* stage using the k–means algorithm. Cluster representatives are depicted on the left. Because of only few subjects assigned into the blue cluster, the difference in *age* between clusters is not evident, although significant, *p–value* = 0.02 (right).



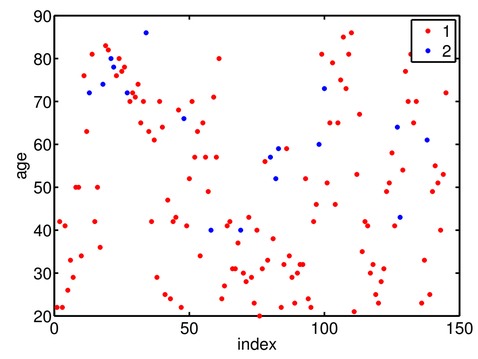
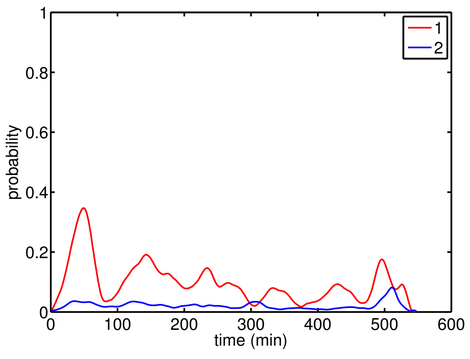


Figure 5: Cluster analysis of a microstate similar to the *Wake* stage using the 2–step approach. Cluster representatives are depicted on the left. The difference in *age* between formed clusters is highly significant, *p–value* < 0.001 (right).



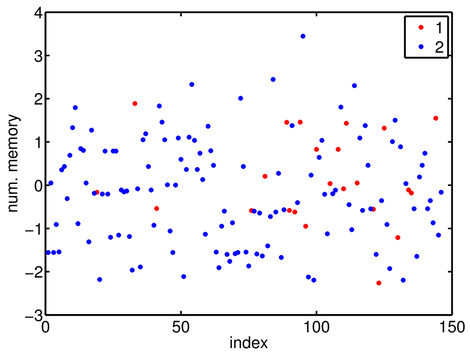
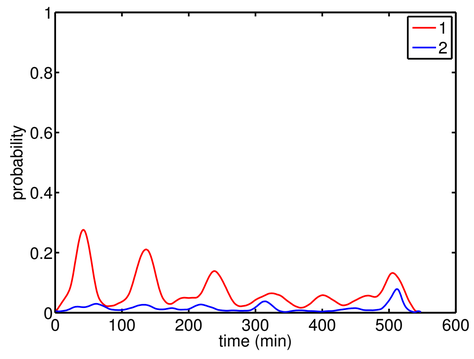


Figure 6: Cluster analysis of a microstate similar to the *S2* stage (25%) and *SWS* (74%). Clusters were formed by the k–means algorithm (left). The difference in *numeric memory test* between clusters is not significant, *p–value* = 0.25 (right).



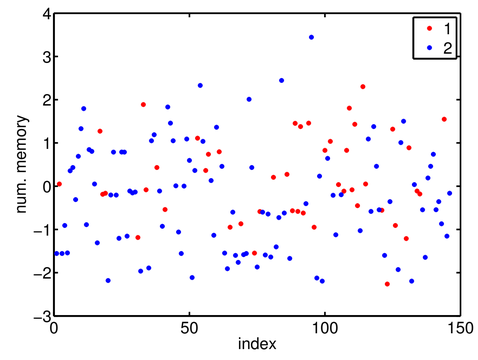
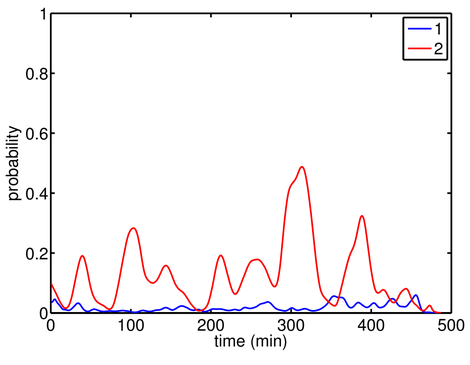


Figure 7: Cluster analysis of a microstate similar to the *S2* stage (25%) and *SWS* (74%) using the 2–step approach (left). The red cluster representative follows the typical pattern for slow wave sleep. The difference in *numeric memory test* between clusters is significant, *p–value* = 0.017 (right).

The second database includes 21 PSG recordings of patients after ischemic stroke together with results of a battery of tests for assessment of motor activity, working memory and attention. Some results obtained by the 2–step approach are depicted in figures below. Considering only the k–means raw data clustering, these relationships remained hidden.



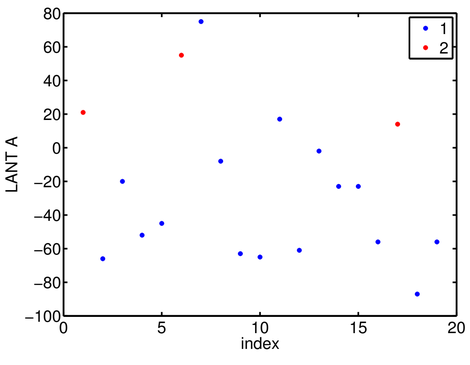
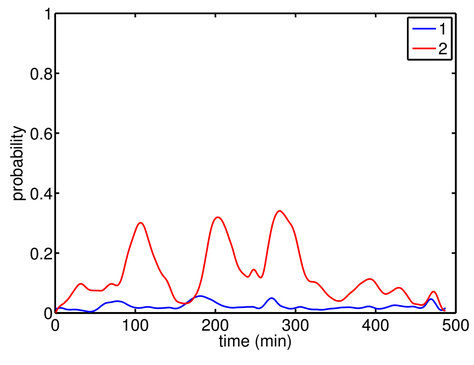


Figure 8: Cluster analysis of a microstate similar to the *Wake* stage (45%) and *S1* stage (51%). On the left cluster representatives are plotted, on the right the structure of clusters is applied to the LANT A test [5]. The ANOVA test rejects the null hypothesis that the LANT A results in both groups are equal, *p-value* = 0.02.



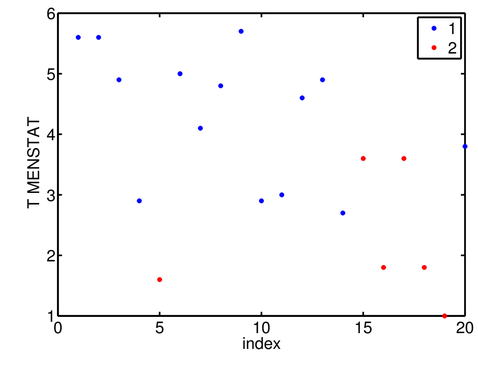


Figure 9: Cluster analysis of a microstate similar to the *S2* stage (53%) and *SWS* (30%). On the left cluster representatives are plotted, on the right the structure of clusters is applied to the results of the TMENSTAT - the questionnaire about mood, mental energy and fatigue. The ANOVA test rejects the null hypothesis that the values in the clusters are equal, *p-value* < 0.001.

Conclusion

Curves time misalignment forms a difficult problem in the curves clustering process. The k–means clustering based on point–to–point distance usually detects clusters with only few subjects and one big, usually nonhomogeneous, subset. Because of an improper clustering relationships between the cluster specific night profiles and daily test scores remain hidden. In contrast to the k–means clustering, the proposed 2–step clustering method shows an improvement to reveal existing relationships between the structure of the sleep process and daily measures. Working with the database of healthy sleep subjects we were able to detect relationship between the majority of sleep microstates and age or the level of day-time drive and drowsiness. Considering the database of patients after ischemic stroke, the observed connection between microstates representing the *Wake* or *S2* sleep stages and the results of LANT seems promising.