Taushanov, Z. & A. Berchtold (2016)

Markovian-Based Clustering of Internet Addiction Trajectories

in G. Ritschard & M. Studer (eds), Proceedings of the International Conference on Sequence Analysis and Related Methods, Lausanne, June 8-10, 2016, pp 179-181.

Markovian-Based Clustering of Internet Addiction Trajectories

Zhivko Taushanov and André Berchtold

Abstract A hidden Markov clustering procedure is applied to a sample of n=185 longitudinal Internet Addiction Test trajectories collected in Switzerland. The best solution has 4 groups. This solution is related to the level of emotional wellbeing of the subjects, but no relation is observed with age, gender and BMI.

1 Introduction

Excessive Internet use is an emerging health issue in the medical literature. The most current tool to quantify the degree of addiction to Internet is the Internet Addiction Test (IAT) developed by Young [7]. However, this scale based on 20 items is quite long and it was not validated for agreement between successive measurements performed on the same subjects, so it is difficult to distinguish between real changes of the subjects and changes due to the IAT itself.

We consider data taken from the *ado@internet.ch* study [6], a longitudinal study about the use of Internet among youth in the Swiss canton of Vaud. Data were collected at 5 occasions every 6 months from Spring 2012 (T0) to Spring 2014 (T4). A convenience sample of n=185 adolescents having answered to all 5 waves is used in the present study (67% females, mean age at T0: 14.1 years). Our goals are 1) to classify the trajectories of Internet addiction into meaningful categories, 2) to test whether this classification is related to other characteristics such as age, gender, wellbeing (measured by the WHO-5 index) and Body Mass Index (BMI). We hypothesize that if significant relationships exist with other characteristics, then the IAT could be considered as a reliable measure of the evolution of Internet addiction.

Zhivko Taushanov
University of Lausanne, ISS & NCCR LIVES, e-mail: zhivko.taushanov@unil.ch

André Berchtold
University of Lausanne, ISS & NCCR LIVES, e-mail: andre.berchtold@unil.ch

2 Model

The clustering of longitudinal continuous data is still an open question. In this paper, we use a specific class of Markovian Models called Hidden Mixture Transition Distribution model (HMTD) for that purpose. These models consist in a latent and an observed levels [4]. The visible level is a Mixture Transition Model (MTD) model introduced by Raftery in 1985 as a modelling of high order Markov chains [5] and developed later by Berchtold [1],[2] and Berchtold and Raftery [3]. Here we use a Gaussian version of the MTD model where the mean of the Gaussian distribution is a function of past observations. The latent level of the model is a homogeneous Markov chain. Each state of the chain is associated to a different component (or model) at the visible level, and the transition matrix is used to determine which component best represents the current observation.

To use the HMTD model as a clustering tool, we fix the hidden transition matrix to the identity matrix. Each sequence of successive observations is then associated to only one component of the model. The dependence order for the Gaussian distributions is fixed to one, and the number of components varies from 2 to 5. We use the Bayesian Information Criterion (BIC) to select the best model.

3 Results

The best model identified by the BIC is the 4 components model. Figure 1 shows the IAT trajectories associated to each group. We clearly differentiate between one group with average volatility and IAT level (group 1); one group with relatively low scores and low variance (2); one group with very low variability and low and constantly diminishing IAT score (3); one erratic group with high variability (4). When comparing the classification with covariates, no significant relationship does appear with age, gender and BMI. However a significant relationship appears with the WHO-5 measured on each wave.

4 Discussion

The HMTD model is able to classify sequences of continuous longitudinal data into as many groups as required. In our example, the four resulting groups differ in terms of average value and of variability. The relationship observed between the IAT and emotional wellbeing suggests that both concepts are linked and that a higher risk of Internet addiction is related to a poorer wellbeing.

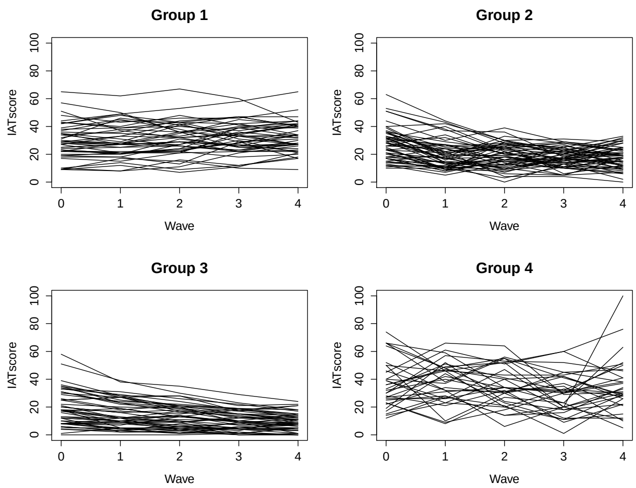

Fig. 1 IAT sequences associated with each cluster in the 4-groups solution.

References

1. Berchtold A. (2001) Estimation in the Mixture Transition Distribution Model. *Journal of Time Series Analysis* 22:379-397
2. Berchtold, A.(2003) Mixture transition distribution (MTD) modeling of heteroscedastic time series. *Computational statistics and data analysis* 41:399-411
3. Berchtold, A. and Raftery, A. (2002) The mixture transition distribution model for high-order Markov chains and non-Gaussian time series. *Statistical Science* 17:328-356.
4. Bolano, D., Berchtold, A. (2016) General framework and model building in the class of Hidden Mixture Transition Distribution models. *Computational Statistics and Data Analysis* 93:131-145
5. Raftery, A. (1985) A Model for High-Order Markov Chains. *Journal of the Royal Statistical Society* 47:528-539
6. Suris, J.-C., Akré, C., Berchtold, A., Fleury-Schubert, A., Michaud, P.-A. and Zimmermann, G. (2012) Ado@Internet.ch: Usage d'internet chez les adolescents vaudois. Lausanne: Institut universitaire de médecine sociale et préventive. (Raisons de santé, 208).
7. Young, K.S. (1998) Internet Addiction: The Emergence of a New Clinical Disorder. *CyberPsychology & Behavior* 1:237-244