Big Data Approaches to Human Chronobiology

Dimitri PERRIN[1], Craig C. JOLLEY[1] and Hiroki R. UEDA[1,2]

[1] Laboratory for Systems Biology, Center for Developmental Biology, RIKEN
2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan
[2] Department of Systems Pharmacology, Graduate School of Medicine, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Email: dperrin@cdb.riken.jp

Abstract

Human chronobiology studies are either done in laboratories or have to rely on volunteers reporting their activities using questionnaires such as the Munich ChronoType Questionnaire (MCTQ). Despite their success, both approaches have limitations, particularly in terms of their ability to handle very large sample sizes. The idea of a large-scale "human sleep project" that would rely, in part, on data being automatically acquired from a large number of participants through ad-hoc devices is very tempting but equally challenging. As a proof-of-concept for such an initiative, we looked at existing Big Data repositories and tried to analyse them in a chronobiology context. Even though the data was not optimised for such a study and mostly contains information on road traffic, telecommunications and electricity usage, we were able to extract relevant information. This is an encouraging first step, just as sleep-specific apps and devices are being released.

Keywords: Big Data, Chronobiology, Pattern identification

1 Background

It has previously been established that social constraints lead individuals to constantly switch between two 'time zones' corresponding to their work schedule and own biological clock, respectively. This 'social jet-lag' has for instance been associated with obesity [1] and depression [2]. It was recently argued that, to better understand the role and mechanisms of sleep in humans, a large-scale project involving millions of participants is necessary [3]. These participants would fill online questionnaires, and some would wear ad-hoc devices to monitor their activity. As a proof-of-concept, and a complement to such initiatives, we have obtained non-biological datasets, to try to analyse them in a chronobiology context and extract relevant patterns of activity.

2 Overview

Most of the datasets were obtained as part of the Telecom Italia Big Data challenge [4]. These correspond to measurements (road traffic, telecommunications, electricity consumption, etc.) acquired over November and December 2013, in Trento and Milan (Italy). Even though the data was not optimised for such a study and mostly contains indirect information on human activity levels, we were able to extract relevant information. Crucially, we show that there is good agreement across data types. This means it is possible to use a single data type when necessary. We therefore also looked at the electricity consumption in France between 1996 and 2013 [5], and in Japan between 2008 and 2013 (e.g. [6] for Kanto), to investigate seasonal effects and the impact of Daylight Saving Time.

References

[1] T. Roenneberg, K.V. Allebrandt, M. Merrow, and C. Vetter, "Social jetlag and obesity", Current Biology vol. 22, 939–943 (2012).

[2] C.M. de Souza, and M.P. Hidalgo, "Midpoint of sleep on school days is associated with depression among adolescents", Chronobiology International vol. 31, 199–205 (2014).

[3] T. Roenneberg, "Chronobiology: The human sleep project", Nature vol. 498, 427–428 (2013).

[4] `http://www.telecomitalia.com/tit/en/bigdatachallenge.html`.

[5] `http://clients.rte-france.com/lang/fr/visiteurs/vie/vie_stats_conso_inst.jsp`.

[6] `http://www.tepco.co.jp/en/forecast/html/download-e.html`.