

SOE 18: Social Networks

Time: Thursday 11:15–12:00

Location: MA 001

SOE 18.1 Thu 11:15 MA 001

Emergent human behaviour on Twitter modelled by a stochastic differential equation — ●ANDERS MOLLGAARD and JOACHIM MATHIESEN — Niels Bohr Institute, Copenhagen, Denmark

In the online era, humans are connected in real time on global scales. Local or seemingly local information is instantaneously shared across geographical boundaries. In particular, social media have become an important platform for the sharing of information and have allowed for detailed studies of the coherent behaviour of humans on a global scale. We have analysed data from the social-media site, Twitter, and used it to study fluctuations in tweet rates of brand names. These have been found to reveal strongly correlated human behaviour that leads to $1/f$ noise and bursty collective dynamics. Here we use a basic definition of aggregated "user interest" to model collective human dynamics by a stochastic differential equation with multiplicative noise. The model is supported by an analysis of tweet rate fluctuations and it reproduces both the bursty dynamics found in the data and the $1/f$ noise.

[1] Mathiesen, Joachim, et al. "Excitable human dynamics driven by extrinsic events in massive communities." *Proceedings of the National Academy of Sciences* 110 (2013): 17259-17262.

[2] Mollgaard, Anders, et al. "Emergent human behaviour on Twitter modelled by a stochastic differential equation." In print.

SOE 18.2 Thu 11:30 MA 001

Proposal of segmentation method of time series in the blogosphere — ●KAZUYA SATOU¹, HIDEKI TAKAYASU^{2,3}, and MISAKO TAKAYASU¹ — ¹Tokyo Institute of Technology 4259, Nagatsuta-cho, Midori-ku, Yokohama-shi, Kanagawa, 226-8502 Japan — ²Sony Computer Science Laboratories, 3-14-13, Higashigotanda, Shinagawa-ku, Tokyo, 141-0022 Japan — ³Meiji University, 4-21-1, Nakano, Nakano-ku, Tokyo, 164-8525, Japan

Time series of the words written in the blogosphere have various types of changes, and it is difficult to extract trends by conventional parametric methods.

In this presentation, we introduce a sequential segmentation method of time series based on Fisher's exact test. In this method, various

types of trends are automatically detected by comparing the p-value with that of randomized time series. Since Fisher's exact test is non-parametric, and can be tested with small number of samples, the proposed method can be potentially applied for the general time series.

In addition, we combine the method with the random diffusion model which detects large discontinuous changes in blog time series.

We discuss the results of detection accuracy of the proposed method by using artificial time series and compare with that of conventional methods. Finally, we show the segmentation results applied for real blog time series.

SOE 18.3 Thu 11:45 MA 001

Prediction of topics' survival using large-scale social data: case of comedian popularity — ●KENTA YAMADA^{1,2}, RYO TAMAOKA¹, and KIYOSHI IZUMI^{1,3} — ¹Graduate School of Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656 Japan — ²PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan — ³CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan

We proposed a new indicator analyzing large-scale textual data in blogs, which predicts future popularity of a topic after a related event. The indicator was tested about the prediction of comedians' popularity after famous comedian contests on television. There are some popular comedian contests in Japan such as the M-1 grand prix. We can universally observe clear peak and power law decaying in the number of blog entries including a champion and vice-champion name (34 samples) after contests as well as the cases in which the number of entries including the event's name follows power function after the events such as Christmas in the previous study [1]. We fitted the number of entries including the comedian's name using five days data after the contest by a power function and cumulated differences between fitted line and actual data from 6 to 12 days after the contest. We found that this index of cumulative differences has a good predictive capability for the number of future (11months later) entries about the comedian.

[1] Y. Sano, K. Yamada, H. Watanabe, H. Takayasu, and M. Takayasu, *PRE* 87, 012805 (2013).