Extraction of Abnormal Values of Word Frequencies in the Blogosphere and Influence on Stock Price

Emiko Takagi[*1], Kenta Yamada[2], Hideki Takayasu[3,4], and Misako Takayasu[1,4]

[1]Department of Mathematical and Computing Science, School of Computing, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Yokohama 226-8502, Japan

[2]Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

[3]Sony Computer Science Laboratories, 3-14-13 Higashigotanda, Shinagawa-ku, Tokyo 141-0022, Japan

[4]Institute of Innovative Research, Tokyo Institute of Technology, 4259, Nagatsuta-cho, Yokohama 226-8502, Japan

E-mail: [*]takagi.e.ab@m.titech.ac.jp

A number of researchers are now trying to search for the relation between financial markets and big data in the society. For example, abnormally large fluctuation of trading volumes in financial markets are shown to be partially explained by the information flow of news [1,2]. It is pointed out that Social Networking Service (SNS) is a new media reflecting the voice of many people, and the information in SNS also affects markets even if it is not true [3]. Chong *et al*. analyzed data of invester's postings for 89 days and reported that investor sentiment correlates with market directions [4]. These studies demonstrate that we cannot ignore the influence of SNS as well as public news for the quantification of the social evaluation of companies. By utilizing the data of major blog portal sites in Japan compiled in over the past 10 years, Here, we pay attention to sudden increase of word frequency of various attributes.

We pick up those words whose frequencies increase considerably higher than the normal periods using a method based on the random diffusion model. Random diffusion model was originally introduced to describe diffusion properties of random walkers on a certain network [5], and it is known to be applicable for abnormal values of word frequency in the blogosphere [6]. In the cases that the words are related to companies, we check corresponding stock prices and analyze how both movement correlate. Then, we try to measure social impact of news which cannot be measured only by the stock prices.

References

[1] R. Hisano *et al*, PLoS one, 8, e64846, 2012.

[2] J.-P. Bouchaud *et al*, Quantitative Finance 4, 176-190, 2004.

[3] M. Takayasu *et al*, PLoS one, 10, e0121443, 2015.

[4] C. Oh *et al*, 32nd International Conference on Information Systems, 2011.

[5] S. M et al, Phys. Rev. Lett., 100, 208701, 2008.

[6] Y. Sano *et al*, Phys. Rev. E, 87, 012805, 2013.