

Empirical analysis on temporal statistics of human correspondence patterns

Nan-Nan Li^a, Ning Zhang^a, Tao Zhou^{b,c,*}

^a Business School, University of Shanghai for Science and Technology, Shanghai 200093, PR China

^b Department of Modern Physics, University of Science and Technology of China, Hefei 230026, PR China

^c Department of Physics, University of Fribourg, 1700 Fribourg, Switzerland

Recently, extensive empirical evidence shows that the timing of human behaviors obeys non-Poisson statistics with heavy-tailed interevent time distribution. In this paper, we empirically study the correspondence pattern of a great Chinese scientist, named *Hsue-Shen Tsien*. Both the interevent time distribution and response time distributions deviate from the Poisson statistics, showing an approximate power-law decaying. The two power-law exponents are more or less the same (about 2.1), which strongly support the hypothesis in [A. Vázquez, J.G. Oliveira, Z. Dezsö, K.-I. Goh, I. Kondor, A.-L. Barabási, Phys. Rev. E 73 (2006) 036127] that the response time distribution of the tasks could in fact drive the interevent time distribution, and both the two distributions should decay with the same exponent. Our result is against the claim in [A. Vázquez, J.G. Oliveira, Z. Dezsö, K.-I. Goh, I. Kondor, A.-L. Barabási, Phys. Rev. E 73 (2006) 036127], which suggests the human correspondence pattern belongs to a universality class with exponent 1.5.

1. Introduction

Being a joint interest of sociology, psychology and economics, human behaviors have been extensively investigated during the last few decades. However, up to now, most of the academic reports on human behaviors are based on clinical records and laboratorial data, and most of the corresponding hypotheses and conclusions are only qualitative. Therefore, the establishment of a quantitative theory on human behaviors becomes one of the main scientific interests nowadays [1,2]. Traditionally, the individual activity pattern is generally simplified as a Poisson process, leading to an exponential interevent time distribution [3]. In such conditions, the probability density function of the recorded time intervals, τ , between two consecutive events has an exponential form as $p(\tau) = \lambda e^{-\lambda\tau}$.

However, with the prompt development of computer technology recently, more and more empirical studies show that many human activities deviate from the Poisson process [4,5]: our activity patterns follow non-Poisson statistics, characterized by bursts of rapidly occurring events separated by long periods of inactivity. Those heavy-tailed distributions, well approximated by a power-law $p(\tau) \sim \tau^{-\gamma}$, are observed in many real systems, including e-mail communication [4], commercial transactions [6], web browsing [7,8], movies-on-demand [9], online games [10], human correspondence patterns [11], short-messages communication [12], computer operations [13], and so on. The discovery of this phenomenon has opened a new research direction namely *human dynamics*. So far, the underlying mechanism leading to those non-Poisson statistics is not clear. Barabási and his colleagues [4,6] have recently proposed that the bursty nature of human dynamics is a consequence of a queuing process driven by human decision making: whenever an individual is presented

* Corresponding author at: Department of Modern Physics, University of Science and Technology of China, Hefei 230026, PR China.
E-mail address: zhutou@ustc.edu (T. Zhou).

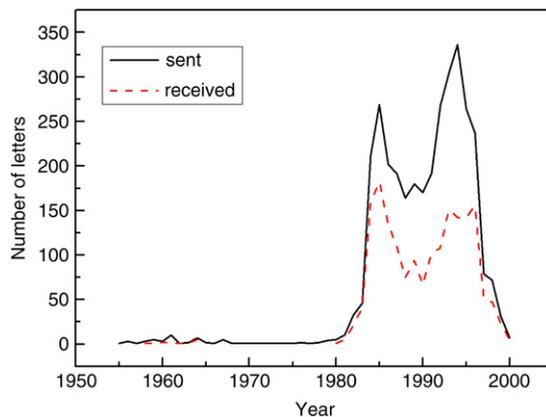


Fig. 1. Historical record of the number of letters sent and received every year.

with multiple tasks and chooses among them based on some perceived priority parameter, the waiting time of the various tasks will be Pareto distributed. Based on the analytical solution of the Barabási queuing model, Vázquez et al. [6] claimed the existence of two universality classes for human dynamics, whose characteristic power-law exponents are 1 and 1.5, respectively. The e-mail communication, web browsing and library loans belong to the former, while the surface mail communication and computer operations belong to the latter.

In the nascent stage of the studies on human dynamics, the empirical explorations are of essential significance, which contribute to the understanding of a complete scenario about real human activity patterns. In this paper, we provide an empirical analysis on the human correspondence patterns of a great Chinese scientist, named *Hsue-Shen Tsien*. Tsien (born in December 11, 1911) is a major figure in the missile and space programs of both the *United States* and the *People's Republic of China*. During the 1940s Tsien was a co-founder of the *Jet Propulsion Laboratory* at the *California Institute of Technology*. When he returned to China he led the *Chinese rocket program*, and became known as the *Father of Chinese Rocketry*. Tsien also contributed a lot to China's Higher Education. He was the first Chairman of the Department of Mechanics of University of Science and Technology of China (USTC). Tsien issued a total of seven monographs and over 300 papers.

The heavy-tailed distributions of interevent and response times are observed. Our result is against the hypothesis of universality classes [6], while it strongly supports another claim of Ref. [6], that is, the interevent time distribution is driven by the response time distribution, thus those two distributions are governed by the same power-law exponent.

2. Empirical results

The database used here contains a record of 3335 letters sent by Hsue-Shen Tsien in day resolution, 1819 of which have indicated the receiving time of the corresponding letters to which they replied. Therefore, we have 1819 response times in total. Fig. 1 gives the historical record of the number of letters sent and received every year. The interevent time, τ , representing the time interval between two consecutive letters sent by Xuesen Qian, is expressed in day resolution. Fig. 2 shows the statistical result about the interevent time distribution of these letters, consisted of 3334 samples in total. The distribution can be well approximated by a power-law form, with the best fitting exponent,¹ $\gamma = 2.1 \pm 0.1$.

Next, we investigate the distribution of response times. Here, the response time, τ_ω , is defined as the time interval between the receive of a letter and the sending time of the corresponding reply. As shown in Fig. 3, the probability that a letter is replied in τ_ω days can be approximately described by a power law $p(\tau_\omega) \sim \tau_\omega^{-\gamma_\omega}$ with $\gamma_\omega = 2.1 \pm 0.1$.

3. Conclusion and discussion

In this paper, we empirically studied the temporal statistics of the correspondence of a great Chinese scientist. Both the distributions of interevent times and response times exhibit heavy-tailed nature, which can be well approximately by a power-law function with exponent 2.1. Based on the analytical solution of the Barabási queuing model [4], Vázquez et al. [6] claimed the existence of two universality classes for human dynamics, whose characteristic power-law exponents are 1 and 1.5, respectively. The e-mail communication, web browsing and library loans belong to the former, while the surface mail communication belongs to the latter. However, our result is against this hypothesis, at least, it can be considered as

¹ A simple but very accurate method to estimate the power-law exponent is the *maximum likelihood estimation* [14], which has already been applied in the empirical investigation of human dynamics [9]. However, due to the very heavy tails in both $p(\tau)$ and $p(\tau_\omega)$, as well as the drooped head in $p(\tau_\omega)$, both the two distributions can not pass the Kolmogorov-Smirnov test with threshold quantile 0.9. Actually, only part of data can be well fitted by a power law, and fitting ranges for $p(\tau)$ and $p(\tau_\omega)$ are [1, 100] and [5, 100] respectively. Here, the best fitting exponents, γ and γ_ω , as well as their errors are obtained by least square method.

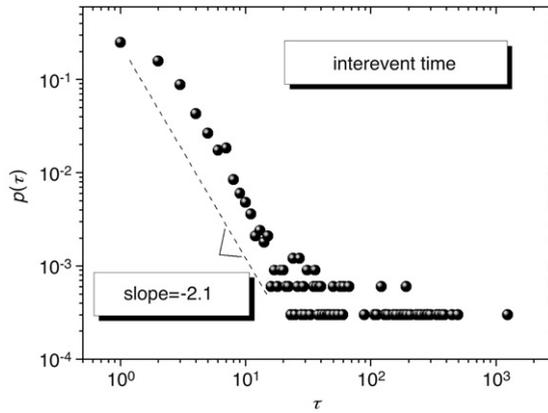


Fig. 2. The distribution of the interevent times for the letters sent by Xuesen Qian, which can be well approximated by a power-law function with exponent 2.1.

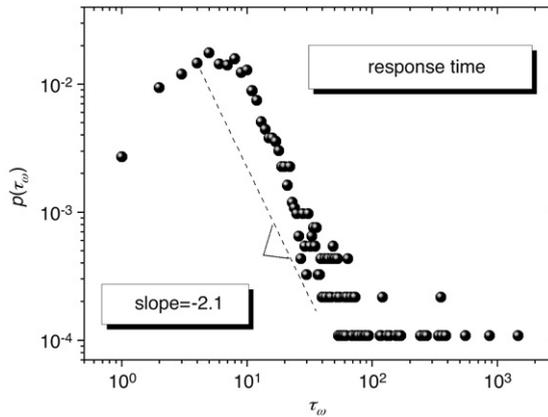


Fig. 3. Response time distribution of the letters, which can be well approximated by a power-law function with exponent 2.1.

a complementary result that suggests a more complicated description of human dynamics, rather than a simple scenario with only two universality classes. Actually, the exponents shall be different for different individuals, since besides the priority of letters, there are countless ingredients (such as the memories [15] and the interests [16]) affecting the human correspondence patterns, and for most of them, we do not know their impact. In the current specific case, as stated by Kentsis [17], the semantic content may play a crucial role. For example, in Einstein’s correspondence history, a huge number of letters waiting for a long time before being replied are relevant to some scientific problems, that is to say, they are not of low priorities, but hard to reply to/answer immediately. In contrast, instead of idiographic scientific problems, most of Tsien’s letters are about administration and management, which can be and need to be replied to fairly quickly. This may be the main reason the exponent reported here is higher than that in Refs. [6,11]. However, though the heavy tails extensively exist in temporal statistics of human correspondence, we believe that different people should have different exponents that reflect their different behavioral patterns.

Furthermore, similar to the phenomena in the statistics on e-mail [4], we found that the exponents of distributions of interevent times and response times are almost the same. This result strongly supports the hypothesis in Ref. [6] that the response time distribution of the tasks could in fact drive the interevent time distribution, and both the two distributions should decay with the same scaling exponent.

Acknowledgements

We would like to thank Zhu-Jun Deng, Yan Qin and Ji-Ming Li for collecting the data. This work is supported by the National Natural Science Foundation of China under Grant No. 10635040, the 973 Program 2006CB705500, Shanghai Leading Academic Discipline Project (No.T0502), and the Natural Science Foundation of Shanghai (06ZR14144).

References

[1] A.-L. Barabási, IEEE Control Syst. Mag. 27 (4) (2007) 33.

- [2] T. Zhou, X.-P. Han, B.-H. Wang, Towards the understanding of human dynamics, in: M. Burguete, L. Lam (Eds.), Science Matters, World Scientific Publishing, Singapore, 2008.
- [3] F.A. Haight, Handbook of the Poisson Distribution, John Wiley and Sons, New York, 1967.
- [4] A.-L. Barabási, Nature 435 (2005) 207.
- [5] K.-I. Goh, A.-L. Barabási, Europhys. Lett. 81 (2008) 48002.
- [6] A. Vázquez, J.G. Oliveira, Z. Dezső, K.-I. Goh, I. Kondor, A.-L. Barabási, Phys. Rev. E 73 (2006) 036127.
- [7] Z. Dezső, E. Almaas, A. Lukács, B. Rácz, I. Szakadát, A.-L. Barabási, Phys. Rev. E 73 (2006) 066132.
- [8] G.-S. Zhao, N. Zhang, T. Zhou, Acta Phys. Sinica (in press).
- [9] T. Zhou, H.A.-T. Kiet, B.J. Kim, B.-H. Wang, P. Holme, Europhys. Lett. 82 (2008) 28002.
- [10] T. Henderson, S. Nhatti, Modelling user behavior in networked games, in: Proc. 9th ACM Int. Conf. on Multimedia, ACM Press, New York, 2001, p. 212.
- [11] J.G. Oliveira, A.-L. Barabási, Nature 437 (2005) 1251.
- [12] W. Hong, X.-P. Han, T. Zhou, B.-H. Wang, arXiv:0802.2577.
- [13] S.K. Baek, T.Y. Kim, B.J. Kim, Physica A 387 (2008) 3660.
- [14] M.L. Goldstein, S.A. Morris, G.G. Yen, Eur. Phys. J. B 41 (2004) 255.
- [15] A. Vázquez, Physica A 373 (2007) 747.
- [16] X.-P. Han, T. Zhou, B.-H. Wang, New J. Phys. 10 (2008) 073010.
- [17] A. Kentsis, Nature 441 (2006) E5.