系统科学学院月报

2016 年第 1 期（总第 16 期）　　2016 年 4 月 15 日

科学研究

◆ 一、吴金闪副教授参加第一次统计物理学计算方法会议

1 月 9 日-1 月 10 日，第一次统计物理学计算方法会议在桂林召开，我院吴金闪副教授应邀参会，并做题为“相干态表象下的量子主方程用于计算非平衡定态”的学术报告。

◆ 二、近期科研成果汇总

1. Zhang, Peng, Wang, Xiang, Wang, Futian, Zeng, An*, & Xiao, Jinghua (2016). Measuring the robustness of link prediction algorithms under noisy environment. Scientific reports, 6.

摘要：链路预测算法的核心是估算网络中任意两点之间存在连边的可能性。至今，很多高精度的链路预测算法已经被提出。然而，这些算法的精确性都是在无噪音信息下的网络中估算的。当网络可观测的连边信息不再准确的情况下，这些链路预测算法的预测精度会如何变化是一个重要的问题，至今尚未解决。在本文中，我们系统的研究了现有的 22 种链路预测算法精度受网络中缺失连边，虚假连边，

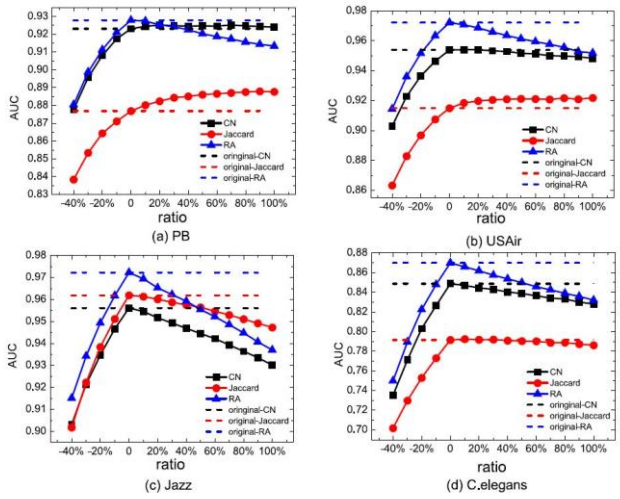

Figure 1. The dependence of link prediction algorithms' accuracy (AUC) on the fraction of missing and noisy links in four real-world networks. The four networks are PB, USAir, Jazz and C.elegans. The link prediction algorithms are *CN*, *Jaccard* and *RA*. Dashed lines represent the AUC of these prediction algorithms in the clean network (i.e. link prediction AUC based on the network without any noisy or missing links). *ratio* < 0 represents the missing link case and *ratio* > 0 stands for the noisy link case.

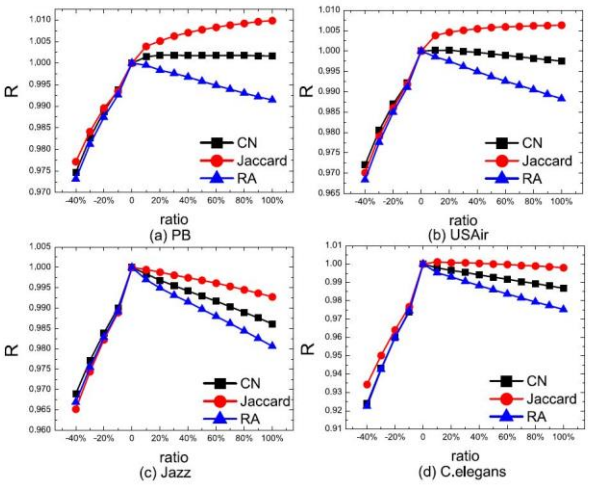

Figure 2. The dependence of the robustness of the algorithms (R) on the fraction of missing and noisy links in four real-world networks.

重连连边的影响。我们发现缺失连边比虚假和重连连边对预测精度影响更大，我们还提出了一个指标来定量化这种影响。在这些链路预测算法中，我们发现一些算法虽然精度略低，但是它们能在存在噪音信息的网络中表现稳定。

Abstract: Link prediction in complex networks is to estimate the likelihood of two nodes to interact with each other in the future. As this problem has applications in a large number of real systems, many link prediction methods have been proposed. However, the validation of these methods is so far mainly conducted in the assumed noise-free networks. Therefore, we still miss a clear understanding of how the prediction results would be affected if the observed network data is no longer accurate. In this paper, we comprehensively study the robustness of the existing link prediction algorithms in the real networks where some links are missing, fake or swapped with other links. We find that missing links are more destructive than fake and swapped links for prediction accuracy. An index is proposed to quantify the robustness of the link prediction methods. Among the twenty-two studied link prediction methods, we find that though some methods have low prediction accuracy, they tend to perform reliably in the "noisy" environment.

原文链接：http://www.nature.com/articles/srep18881

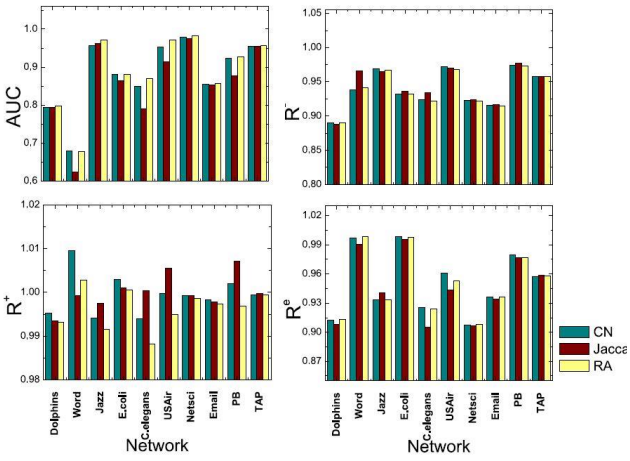

Figure 3. The *AUC* and robustness of link prediction algorithms in ten real networks. R^-, R^+ and R^e are respectively the robustness of the algorithms with missing links, noisy links and swapped links. The fraction c changed links here is 40%.

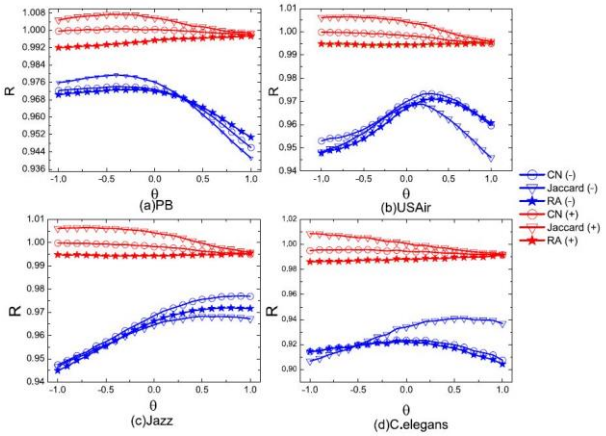

Figure 4. The effect of different biased disturbance on four real-world networks. The tunable parameter θ is used to control the preference of the noisy links and missing links connecting to nodes with different degree. The red curves are the results of R^+ while the blue curves are the results of R^-.

2. Fu, Ling, Shen, Zhesi, Wang, Wenxu, Fan, Ying*, & Di, Zengru (2016). Multi-source localization on complex networks with limited observers. EPL (Europhysics Letters), 113(1), 18006.

简介：基于有限观察者的复杂网络多源头定位。复杂网络上扩散过程的源头定位是一个重要问题，多个源头的存在给复杂网络源头定位带来了更大的困难。本文通过部分节点的被扩散时间，利用逆向回溯的方法得到扩散过程的多个源头。

摘要：Source localization is a significant task in the contagion process. In this paper, we study the problem of locating multiple sources in complex networks with limited observations. We propose a backward diffusion-based source localization method and apply it on several networks, finding that multiple sources can be located with high accuracy even when the fraction of observers is small and the time delay along the links are not known exactly. By comparing different observer placement strategies, we find that choosing small-degree nodes as observers is better than the other strategies.

原文链接：http://iopscience.iop.org/article/10.1209/0295-5075/113/18006/meta

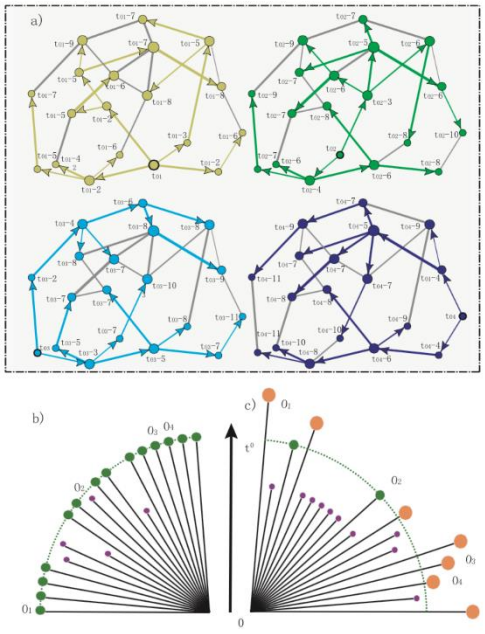

Fig. 2: (Color online) Illustration of the backward diffusion-based multi-source localization method. (a) Backward diffusions from each observer; the number marked at each node is $\Gamma(i,o_k)=t_{o_k}-\Delta_{i,o_k}$. (b) For a source node s_m, all values of $\Gamma(s_m,i)$, $\forall i\in V$ are arranged in a circle with radius proportional to $\Gamma(s_m,i)$. All the points are bounded in a circle whose radius is t^0, that is $\Gamma^{\max}_{s_m}=t^0$. (c) For a normal node $j\notin S$, some values of $\Gamma(j,i)$ will exceed the boundary, resulting in $\Gamma^{\max}_{j}>t^0$. The sources can be then identified as the nodes whose $\Gamma^{\max}_{s}$ are minimal.

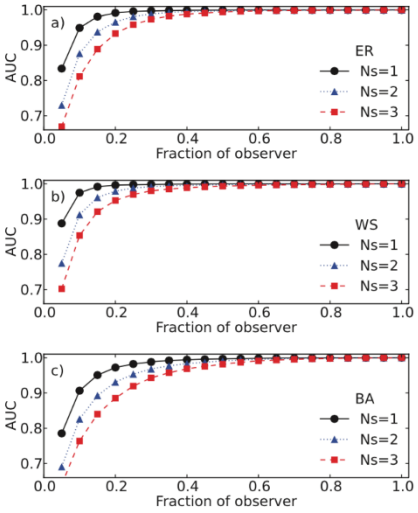

Fig. 3: (Color online) Performance of multiple source localization on model networks. (a) ER random network. (b) WS small-world network; the rewiring probability is 0.1. (c) BA scale-free network. The network size used here is $N=100$, and the average degree is $\langle k\rangle=4$. The time delay along the links follows a Gaussian distribution $N(1,0.25^2)$ and only the mean delay of all links is used to identify the sources.

3. Zhesi Shen, Shinan Cao*, Wen-Xu Wang*, Zengru Di, and H. Eugene Stanley (2016). Locating the source of diffusion in complex networks by time-reversal backward spreading, Phys. Rev. E 93, 032301.

简介：基于逆向回溯方法的复杂网络源头推断。本文主要研究了复杂网络上扩散过程的源头定位问题。定位扩散过程的源头是在众多动力学研究中有重要的意义，比如疾病传播、谣言扩散等。本文通过观察部分节点的被扩散时间，逆向回溯得到传播扩散的源头，并提出基于此的可定位性条件。

摘要：Locating the source that triggers a dynamical process is a fundamental but challenging problem in complex networks, ranging from epidemic spreading in society

and on the Internet to cancer metastasis in the human body. An accurate localization of the source is inherently limited by our ability to simultaneously access the information of all nodes in a large-scale complex network. This thus raises two critical questions: how do we locate the source from incomplete information and can we achieve full localization of sources at any possible location from a given set of observable nodes. Here we develop a time-reversal backward spreading algorithm to locate the source of a diffusion-like process efficiently and propose a general locatability condition. We test the algorithm by employing epidemic spreading and consensus dynamics as typical dynamical processes and apply it to the H1N1 pandemic in China. We find that the sources can be precisely located in arbitrary networks insofar as the locatability condition is assured. Our tools greatly improve our ability to locate the source of diffusion in complex networks based on limited accessibility of nodal information. Moreover, they have implications for controlling a variety of dynamical processes taking place on complex networks, such as inhibiting epidemics, slowing the spread of rumors, pollution control, and environmental protection.

原文链接：http://journals.aps.org/pre/abstract/10.1103/PhysRevE.93.032301

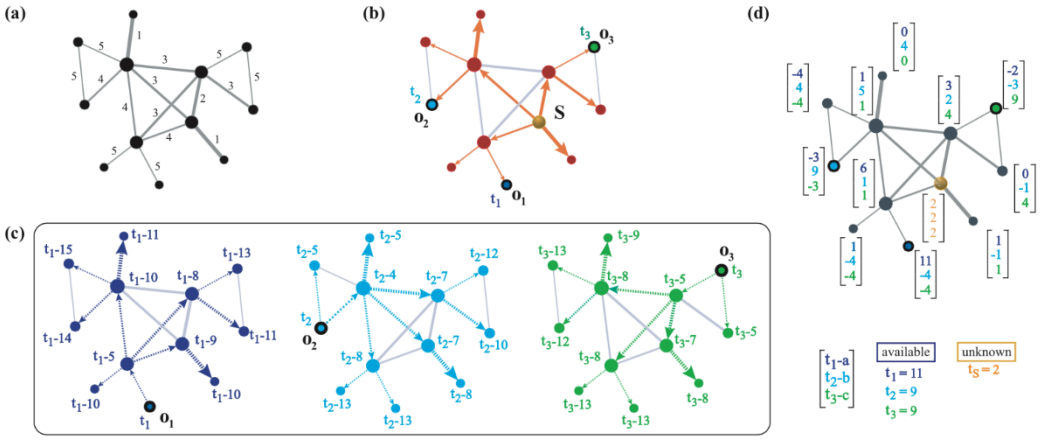

FIG. 1. Time-reversal backward spreading for locating the source. (a) A network topology with link weights (time delay). (b) The diffusion paths from the source S and the observers o_1, o_2, and o_3. The arrival time only at the three observers, namely, t_1, t_2, and t_3 can be accessed. (c) Implement TRBS along weighted shortest paths from o_1, o_2, and o_3, respectively, and the reversed arrival time at each node stems from each observer, respectively. (d) the vector $\mathbf{T}$ consisting of the reversed arrival time from each of the observers. The elements of $\mathbf{T}_s$ of the source are identical, which is the key to distinguishing the source from the other nodes. If the observers provide sufficient information of the source, the revered arrival time from observers are the original time t_s of the diffusion from the source, enabling the recovery of t_s. The source S is in yellow and the three observer nodes are in dark blue, light blue, and green with black boundary. The actual diffusion from S is marked by orange solid lines with arrows and the TRBS from the observers are, respectively, marked by colored dotted lines with arrows. The color of numbers in the vector in (d) corresponds to the observer with the same color.

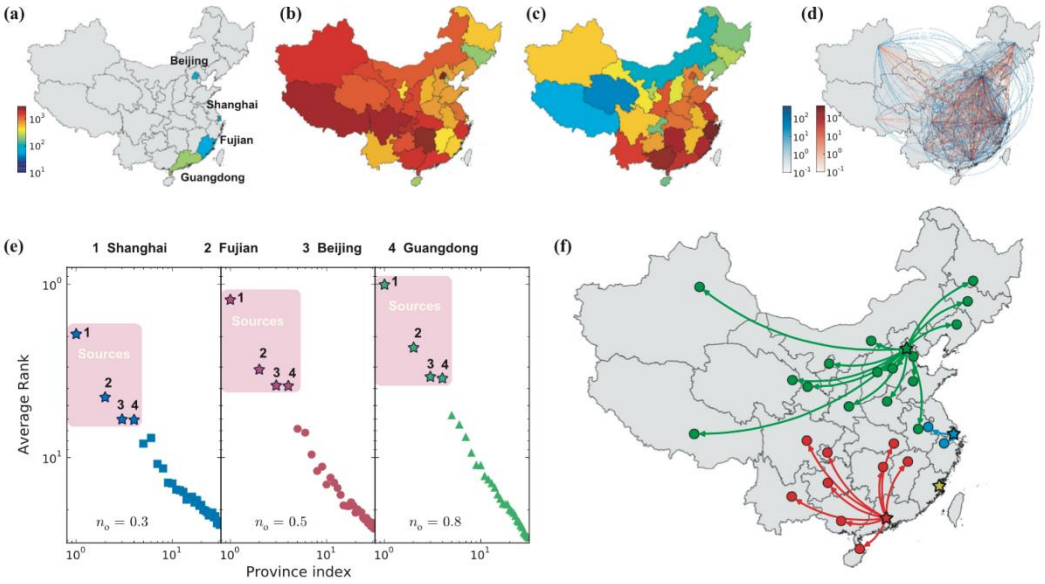

FIG. 5. Locate the sources of H1N1 pandemic in China. (a) The earliest outbreak of H1N1 in June 2009 in four provinces—Beijing, Shanghai, Fujian and Guangdong—which are the sources of the epidemic spreading in China. The epidemic outbreaks occur at the four locations nearly simultaneously. (b) The outbreak of H1N1 all over China in Oct. 2009. (c) The number of patients in China in December 2009. The color bar in (a), (b), and (c) denote the number of patents. (d) China airline and train networks with weighted links. The color bars capture the passenger flux of airlines and trains per day, respectively. The mixture of the airline and train networks is used as the propagation network of the H1N1 virus. (e) The average ranks of different provinces corresponding to the probabilities of being the sources of the epidemic spreading calculated by our algorithm. The four actual sources are of the highest four ranks with respect to different fraction n_o of observers and there is a clear gap between the sources and the other provinces. (f) The most probable paths of spreading from the sources uncovered by using the estimated time delays along links. The results in (e) are obtained by randomly choosing 100 independent configurations of observers with different fractions.

4. Li Xiaomeng, Chen Qinghua, Fang Fukang* & Zhang Jiang (2016). Is online education more like the global public goods? Futures, Available online 6 January 2016.

简介：发近些年随着技术进步，在线课程成为教育的新模式。但是较少有人从定量的角度研究在线教育对世界范围内教育平等的贡献。本文认为目前在线教育在消费中的少竞争性和排他性使其类似全球公共物品。我们通过建立教育产出模型进行分析，并指出在线教育的发展能够帮助发展中国家和发达国家提高人力资本产出，而这种增长源自在线教育作为全球公共物品的基本属性。在进一步分析这种增长的可行性和可持续性的基础上，我们建议成立在线教育组织等机构来帮助行业逃离“搭便车”的决策陷阱，这在公共物品供给中是常见的问题。本文提供了一种全新的视角来审视在线教育的意义，指出在互联网及“慕课”时代，这种基于不同比较优势的教育合作，能够帮助减少国家间教育资源的差距并在一定程度上实现协同发展。

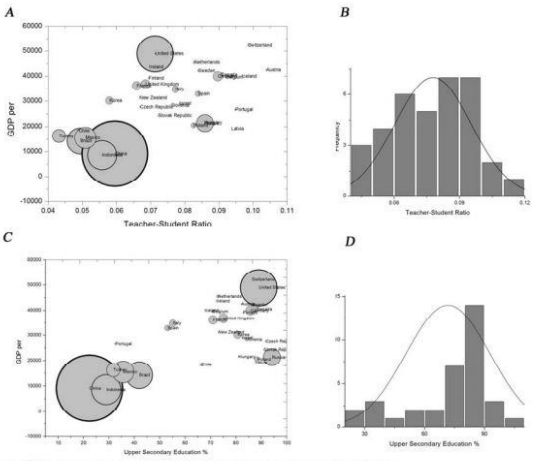

Figure 1 The gap between countries in educational resources and outcomes.

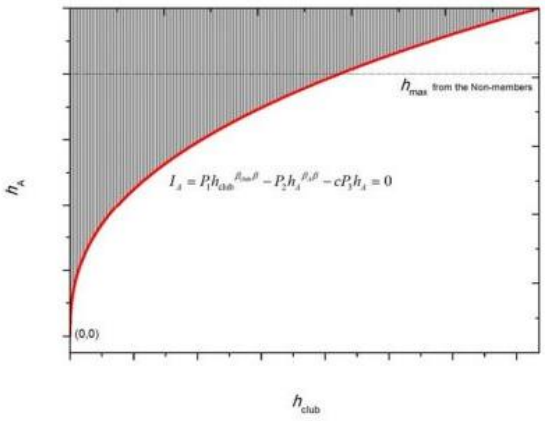

Figure 3 Decision Function of country *A*.

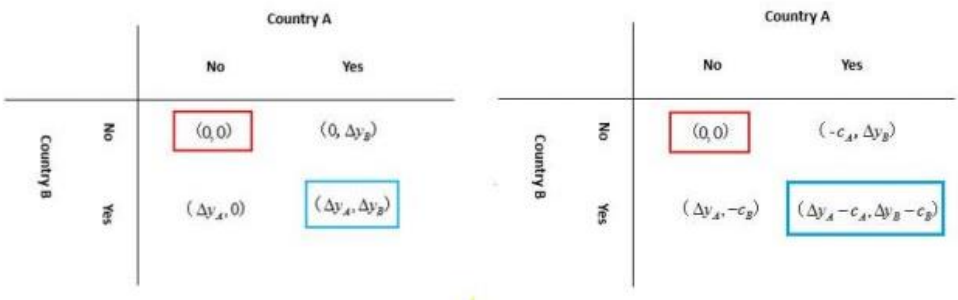

Figure 2 The Pay-Off Matrix and Strategies Analysis.

Abstract：Nowadays with the possibility of synchronous e-learning, online courses are becoming a new paradigm in education. But there are few studies systematically discussing the interplay between online courses and global education equality quantitatively. Here we indicate that online education is more like the global public goods with on-rival and non-excludable in consumption. We establish an educational outcome model and indicate that popularity of online education is likely to give rise to the human capital augmentation, both for developing and developed countries. And this growth stems from the characteristics of global public goods. To analyze the possibility of the potential growth, we suggest that an online education club would help countries to escape from the dominant strategies of "Free ride", which is the common dilemma in public goods supply. So the cooperation would bring coordinated growth to countries with different advantages in educational resources. And this win-win situation offers an underlying driving force to the development of human capital and economics. This paper provides a view offering the deep understanding on reducing the gap of education outcome between countries, and get a coordinated development in the Internet or MOOCs Era. It gives more supports to the development of online education in the future.

原文链接：http://www.sciencedirect.com/science/article/pii/S0016328715001354

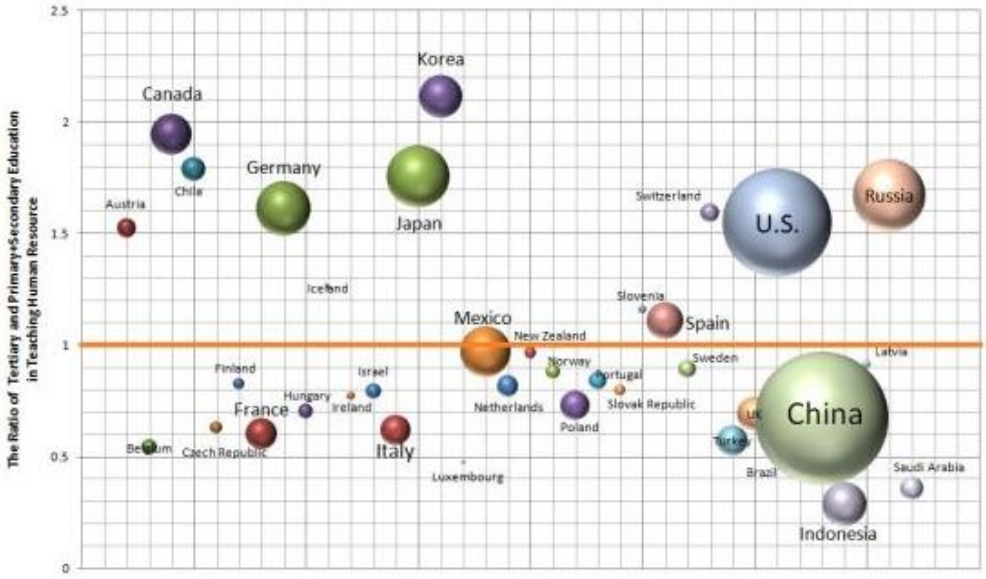

Fig 4. The elasticity coefficients of education resources in Tertiary Education and Primary + Secondary Education.

5. Lin, Zhenquan*, Hou, Shanci, & Wu, Jinshan (2016). The correlation between editorial delay and the ratio of highly cited papers in Nature, Science and Physical Review Letters. Scientometrics, 1-8.

摘要：Ideally, in a reviewing process, it is generally easier for referees to make faster and more reliable decisions for high quality papers, which ideally and on average will later attract more citations. Therefore, it is possible that the editorial delay time—the time between dates of submission and acceptance or publication—is correlated to the number of received citations, as has been weakly confirmed by previous studies. In this study, we propose a different measure for this correlation. Instead of directly calculating the correlation coefficient between the editorial delay and the number of citations, we define a ratio of above median highly cited papers and perform correlation analysis between editorial delay and that ratio for all academic papers published in Nature, Science and Physical Review Letters. Using this alternative measure, we find that on average papers with shorter editorial delay do have larger probabilities of becoming highly cited papers.

原文链接：http://link.springer.com/article/10.1007/s11192-016-1936-z

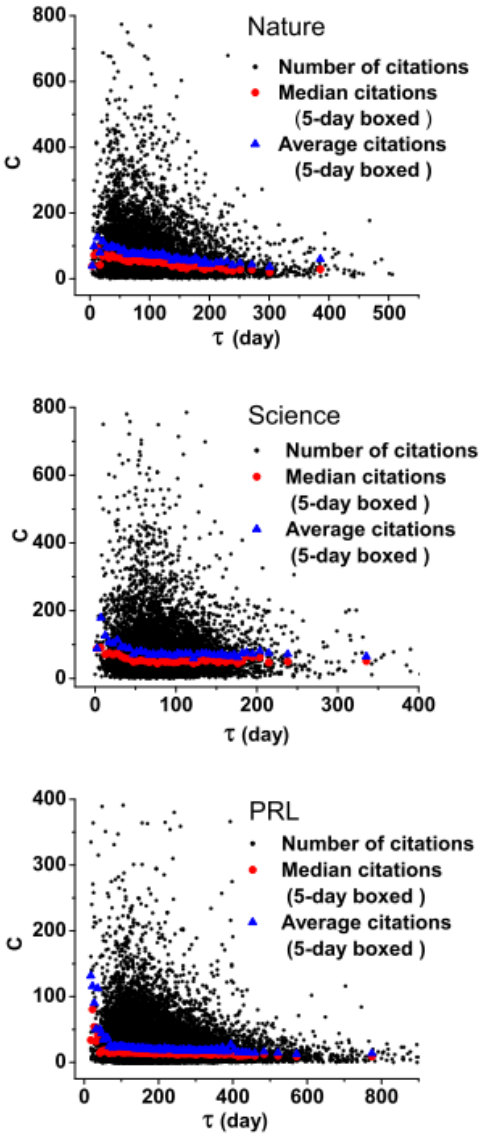

Fig. 1 Plots of the number of citations C in the 3-year period after the publication year versus the editorial delay τ of each sample paper in the journals *Nature*, *Science* and *PRL*. The average number of citations $\bar{C}$ and the median number of citations C_m of papers in each 5-day editorial delay box versus the average editorial delay $\bar{\tau}$ are also plotted

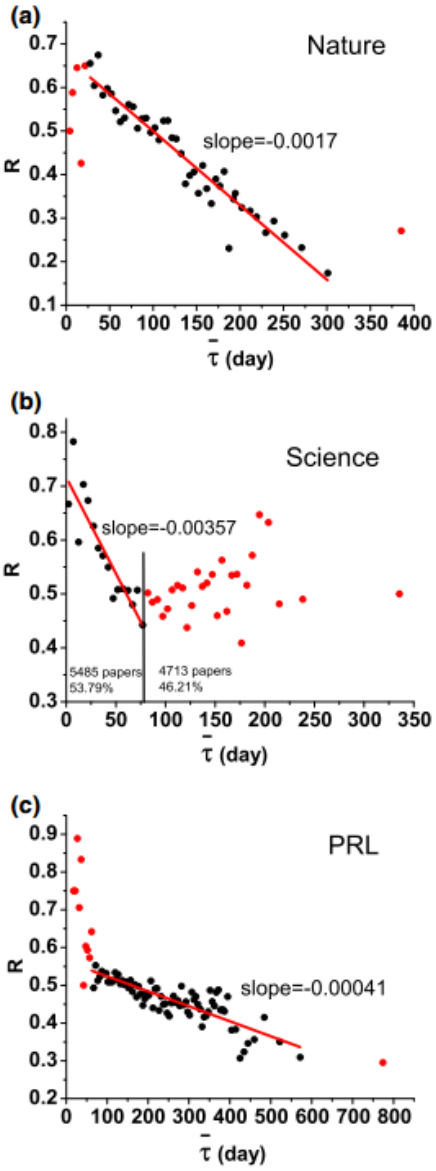

Fig. 2 The ratios of highly cited papers R versus the average editorial delay $\bar{\tau}$ of each 5-day *box plotted* for papers from: **a** *Nature*, **b** *Science*, and **c** *PRL*

◆ 三、学术讲座

	主讲人	学术头衔	主持人	题目	时间	地点
1	Haviland Wright	University of Colorado, University of Denver, Ph.D.	狄增如	计学在交叉学科中的应用	3 月 25 日	教 8 楼 207
2	张帆	北京师范大学天文系，副教授		终于等到你——引力波发现趣谈	3 月 18 日	教二楼 108
3	Hans-F Graf	剑桥大学地理学系，教授	崔雪锋	The ENSO Continuum-Predictability and change of the background state	3 月 17 日	217
4	Zhang Ke	James Cook University in Australian，博士后	狄增如	Dynamics of coupled social-ecological systems: learning from the past	3 月 14 日	217
5	武斌	北京邮电大学理学院，讲师	曾安	Evolution of cooperation	1 月 8 日	217
6	廖旭红	北京师范大学认知神经科学与学习国家重点实验室，博士后	崔晓华	基于静息态功能磁共振的人脑功能网络研究简介	1 月 6 日	217