

◆ 九、近期科研成果汇总

1. An Zeng, Chi Ho Yeung, Matus Medo, Yi-Cheng Zhang, Modeling mutual feedback between users and recommender systems. [Journal of Statistical Mechanics, P07020 (2015)]

简介：我们在各种网站上常常接触并使用推荐系统。在我们享受它带来的便捷时，它也在潜移默化的影响着我们的选择习惯。前人对推荐系统的研究主要集中在如何提高推荐系统的精确性和多样性，而不同推荐系统在长期使用下会对商品形成如何的关注度分布尚不清楚。本文提出了一个基于用户-商品二分网的演化模型来研究推荐系统对用户选择行为的长期影响。虽然大家广泛认为推荐系统能为用户拓宽视野，结果显示在长期演化中用户的选择范围被推荐系统大大缩小。另外，本文还发现此模型能涌现出迟滞现象，使得推荐系统的这种不利效果很难通过自身恢复。最终，我们发现在演化早期牺牲推荐系统的少量推荐精度可以有效减弱它产生的这种长期不利影响。

Recommender systems daily influence our decisions on the Internet. While considerable attention has been given to issues such as recommendation accuracy and user privacy, the long-term mutual feedback between a recommender system and the decisions of its users has been neglected so far. We propose here a model of network evolution which allows us to study the complex dynamics induced by this feedback, including the hysteresis effect which is typical for systems with non-linear dynamics. Despite the popular belief that recommendation helps users to discover new things, we find that the long-term use of recommendation can contribute to the rise of extremely popular items and thus ultimately narrow the user choice. These results are supported by measurements of the time evolution of item popularity inequality in real systems.

We show that this adverse effect of recommendation can be tamed by sacrificing part of short-term recommendation accuracy.

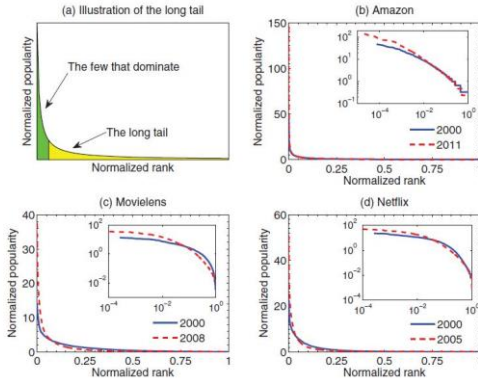


Figure 1. (a) A simple illustration of the long tail phenomenon, and (b)–(d) its presence in real data produced by various e-commerce systems. The normalized item popularity is used to remove the size effect⁴. The same plots in log-log scale are shown in the respective insets. Results from different time periods show that popular items become more popular compared to the niche items at the tail, contrary to the belief of a thicker tail obtained through recommendation algorithms.

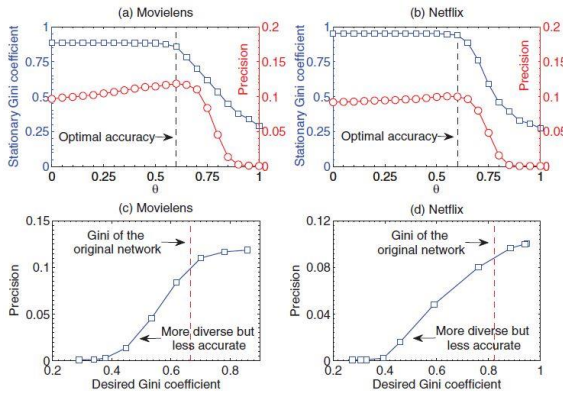


Figure 5. The dependence of the stationary Gini values and recommendation precision on θ in (a) Movielens and (b) Netflix data. Panel (c) and (d) show the relation between the desired stationary Gini coefficient and the short-term recommendation precision.

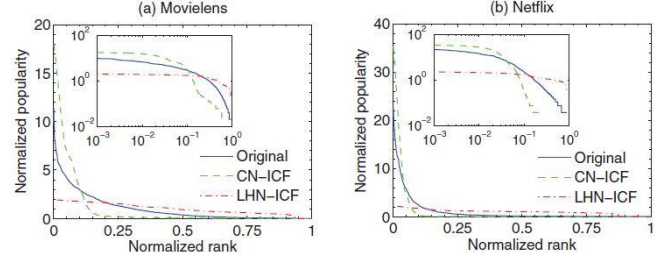


Figure 2. The normalized item popularity versus the normalized item rank in (a) Movielens and (b) Netflix. The same plots in log-log scale are shown in the respective insets.

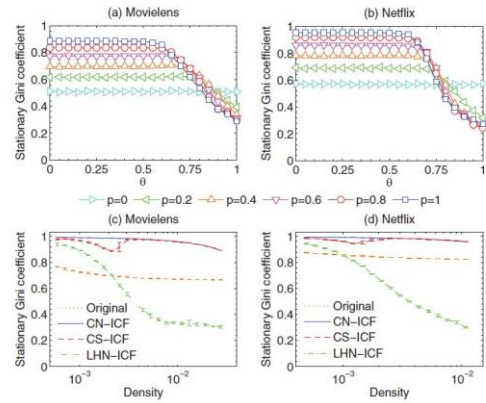


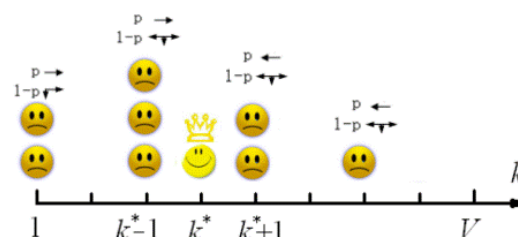
Figure 3. (a), (b) The effect of the ICF's parameter θ on the stationary Gini coefficient: links which are not drawn based on recommendation are drawn based on preferential attachment. (c), (d) The effect of data density on the stationary Gini coefficient for different recommendation methods (here all links are drawn based on recommendation). The curve labeled original is the Gini coefficient of the network after density modification and before rewiring.

2. Qinghua Chen, Zi-Gang Huang, Yougui Wang and Ying-Cheng Lai (2015) Multi-agent model and mean field theory of complex auction dynamics. New Journal of Physics, 17: 093003. doi:10.1088/1367-2630/17/9/093003 (<http://iopscience.iop.org/article/10.1088/1367-2630/17/9/093003>)

简介：长期以来，学者对于最小唯一拍卖的研究主要立足于完美信息下的 Nash 均衡框架，从结果上很难对真实的投标分布情况做出好的解释，特别是在实证投标分布中的增长部分。此外，从逻辑上看，完美信息的 Nash 均衡所需要的假设很强，实际情况难以满足。所以需要寻找另外的框架来进行讨论这个问题。

文章提出了一个具有学习效应的多主体模型，物理上可以看作多个粒子在一定随机驱动情况下的有界范围内的随机行走。模型框架如下图所示，每个个体以一定的概率 p 向胜利者前进一步（多步也可以，或者某一个分布）； $1-p$ 的概率进行随机行走。个体的投标是离散的并被限制在一个有限区间（不能比 1 小，不能比标的物品的价格 V 多）。

无论初始的投标分布如何，经过足够长的时间后系统会达到一个相当稳定的动态平衡。我们进行了平均场分析得到了和多个体模拟一致的结果，还证明了尾段的



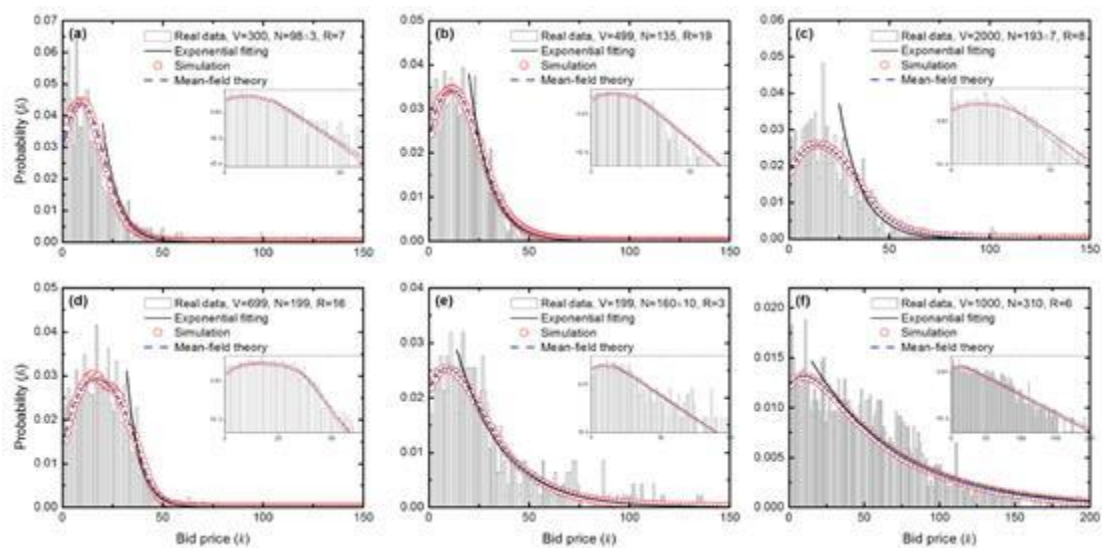
指数下降现象（这个现象在一些实证工作的文章中被提及而没有解决）。理论和实证结果的比较如下图，图中灰色柱状图为实际分布，红圈为多个体模型模拟的结果，蓝色虚线为平均场分析结果，黑色实线是对实证数据尾部的指数拟合。相比以前的模型，文章提出的向胜利者学习的机制显得更加合理（向更高收益者学习是人类的基本行为模式），并且该假设从实际数据的分析中获得支持（感谢 **Filippo Radicchi** 所提供的数据）。最后还值得突出说明的是，模型只有一个参数（学习率 p ），但获得了更好的整体拟合效果。

注：和传统的“价高者得之”的拍卖方式不同，最小唯一拍卖是一种比较新颖的拍卖方式，所有唯一出价中最小的获胜。这种拍卖中的投标分布会呈现倒 J 形模式，并且这种特征非常显著，如何解释这个分布的形成吸引了学者的关注。我院陈清华副教授和王有贵教授及其合作者近几年对这种分布的形成进行了深入的研究，提出了几种可能的机制模型，相比以前的工作与实际分布的拟合效果更好，而机制更加合理。具体可以参考如下已经发表的文献：

- a. Yinan Zhao, Qinghua Chen, Yougui Wang (2013) Bid distribution derived from consistent mixed strategy in lowest unique bid auction. *International Journal of Modern Physics C* 25: 1440002.
- b. C Zhou, H Dong, R Hu, Q Chen (2015) Smarter than Others? Conjectures in Lowest Unique Bid Auctions. *PLoS ONE* 10(4): e0122923.

doi:10.1371/journal.pone.0122923

(<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0122923>)



3. Qikai Niu, An Zeng, Ying Fan and Zengru Di, Robustness of centrality measures against network manipulation, *Physica A* 438, 124 (2015).

简介: Node centrality is an important quantity to consider in studying complex networks as it is related to many applications ranging from the prediction of network structure to the control of dynamics on networks. In the literature, much effort has been devoted to design new centrality measurements. However, the reliability of these centrality measurements has not been fully assessed, particularly with respect to the fact that many real networks are facing different kinds of manipulations such as addition, removal or rewiring of links. In this paper, we focus on the robustness of classic centrality measures against network manipulation. Our analysis is based on both artificial and real networks. We find that the centrality measurements are generally more robust in heterogeneous networks. Biased link manipulation could more seriously distort the centrality measures than random link manipulation. Moreover, the top part of the centrality ranking is more resistant to manipulation.

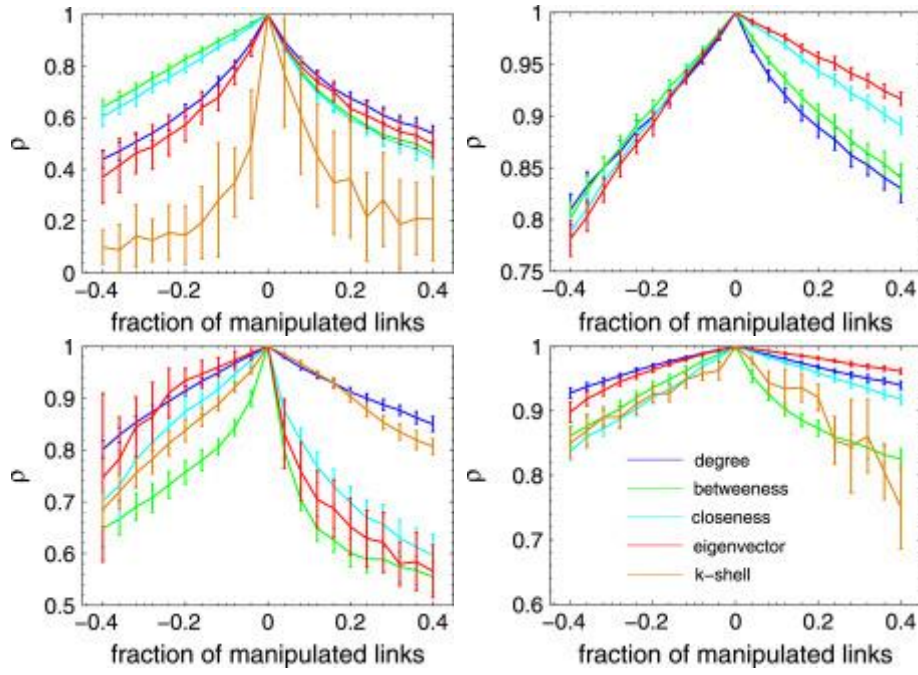


Fig. 1 (Color online) The influence of adding and removing links on the Spearman correlation ρ between the true measure C_T and manipulated measure C_M . We denote the number of links in the original network as E and the fraction of manipulated links as f . When f is positive, fE links are added to the original network. When f is negative, fE links are removed from the original network. For SW networks, $N=500$, $\langle k \rangle = 12$, $p=0.2$. For BA networks, $N=500$, $\langle k \rangle = 12$, $M=12$ (M is the number initial nodes in the BA model). The results are obtained from 50 independent realizations.

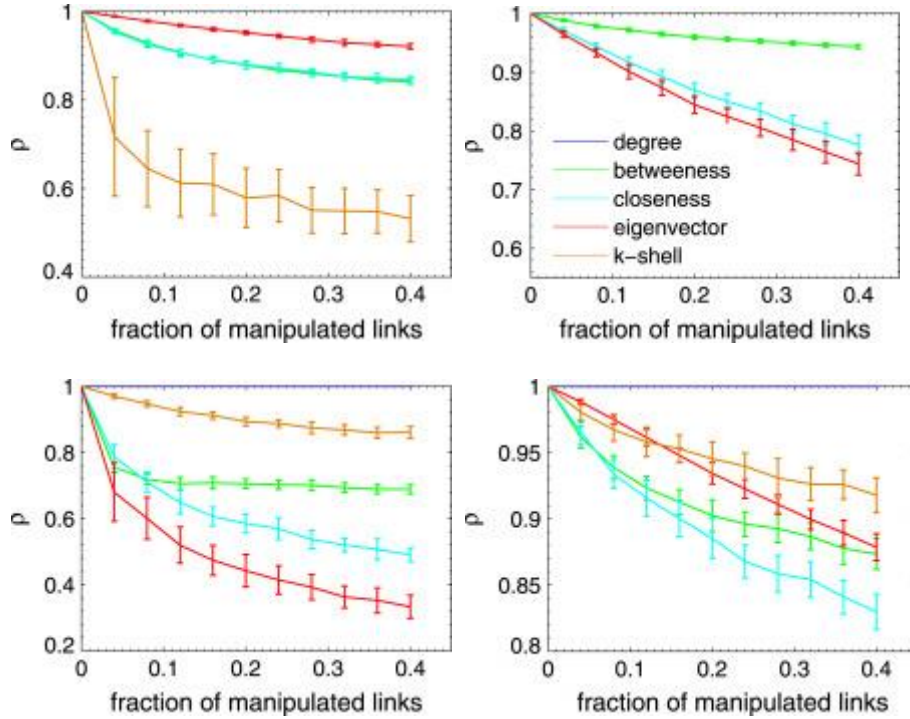


Fig. 2 (Color online) The influence of rewiring links on the Spearman correlation ρ between the true measure C_T and manipulated measure C_M . For SW networks, $N=500$, $\langle k \rangle = 12$, $p=0.2$. For BA networks, $N=500$, $\langle k \rangle = 12$, $M=12$. The results are obtained from 50 independent realizations.

4. An Zeng, Chi Ho Yeung, Matus Medo, Yi-Cheng Zhang, Modeling mutual feedback between users and recommender systems. [Journal of Statistical Mechanics, P07020 (2015)]

简介： Recommender systems daily influence our decisions on the Internet. While considerable attention has been given to issues such as recommendation accuracy and user privacy, the long-term mutual feedback between a recommender system and the decisions of its users has been neglected so far. We propose here a model of network evolution which allows us to study the complex dynamics induced by this feedback, including the hysteresis effect which is typical for systems with non-linear dynamics. Despite the popular belief that recommendation helps users to discover new things, we find that the long-term use of recommendation can contribute to the rise of extremely popular items and thus ultimately narrow the user choice. These results are supported by measurements of the time evolution of item popularity inequality in real systems. We show that this adverse effect of recommendation can be tamed by sacrificing part of short-term recommendation accuracy.

我们在各种网站上常常接触并使用推荐系统。在我们享受它带来的便捷时，它也在潜移默化的影响着我们的选择习惯。前人对推荐系统的研究主要集中在如何提高推荐系统的精确性和多样性，而不同推荐系统在长期使用下会对商品形成如何的关注度分布尚不清楚。本文提出了一个基于用户-商品二分网的演化模型来研究推荐系统对用户选择行为的长期影响。虽然大家广泛认为推荐系统能为用户拓宽视野，结果显示在长期演化中用户的选择范围被推荐系统大大缩小。另外，本文还发现此模型能涌现出迟滞现象，使得推荐系统的这种不利效果很难通过自身恢复。最终，我们发现在演化早期牺牲推荐系统的少量推荐精度可以有效减弱它产生的这种长期不利影响。

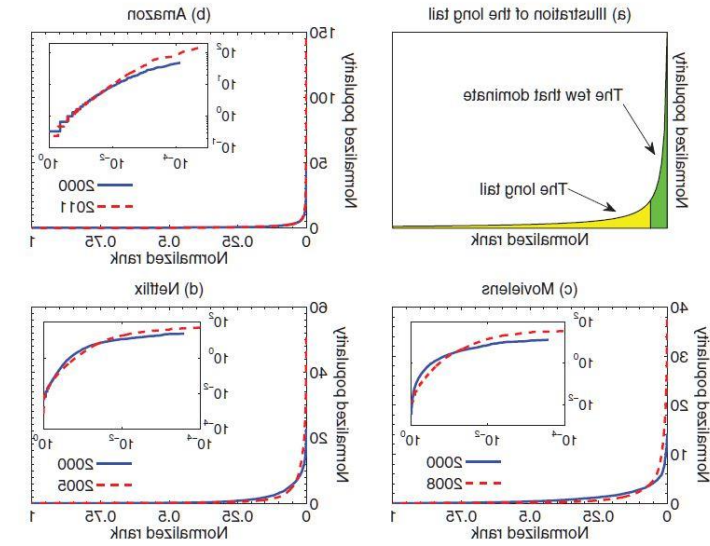


Figure 1. (a) A simple illustration of the long tail phenomenon. (b)–(d) its presence in real data produced by various e-commerce systems. The normalized item popularity is used to remove the size effect. The same plots in log-log scale are shown in the respective insets. Results from different time periods show that popular items become more popular compared to the niche items at the tail, contrary to the belief of a thicker tail obtained through recommendation algorithms.

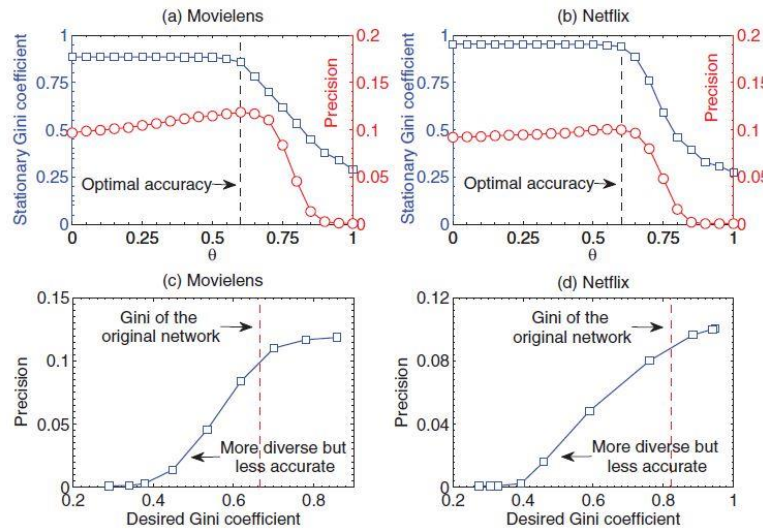


Figure 5. The dependence of the stationary Gini values and recommendation precision on θ in (a) Movielens and (b) Netflix data. Panel (c) and (d) show the relation between the desired stationary Gini coefficient and the short-term recommendation precision.

5. Liangzhu Guo,Xiaodan Lou,Peiteng Shi, Jun Wang, Xiaohan Huang, Jiang Zhang:
Open Flow Distances on Open Flow Networks; Physica A, Vol 437, 1, 235 – 248.

简介： Open flow network is a special weighted directed graph in which weighted links are flows, and the flows are in balance. We define a new set of distance metrics,

which measure the average length of particles flow from i to j . Based on the distances, we discuss the calculation of trophic levels of species on energetic food webs, the centrality of nodes, and the industrial clustering problem on input-output networks, etc. We also compare the new distances with old distances on graph.

开放流网络是一种特殊的加权有向网，其中加权连边表示流量，同时节点满足流平衡。在这种网络上，我们定义了一组流距离，即流子沿连边流动从 i 到 j 经历的平均路径长度。在此基础上，我们讨论了食物网的营养级计算、投入产出网上的节点中心度、产业聚类等问题，我们还比较了新的距离与其它网络距离。

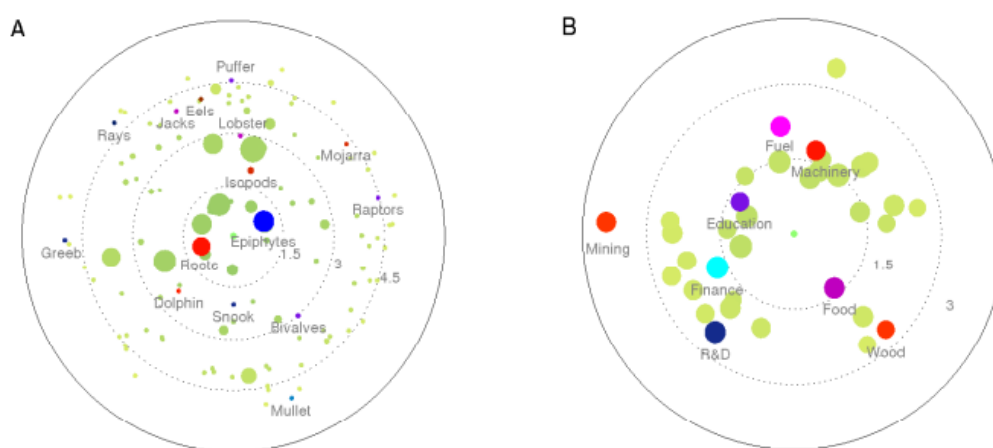


FIG. 2. Trophic levels of species in Baydry community (a) and industrial sectors of U.S. input-output network in 2000 (b). The polar radii, i.e., the distances between every node and the center are proportional to nodes' trophic levels, the polar angles are randomly assigned, and the sizes of nodes are proportional to the logarithmic volumes of the total throughflow for each node (f_i). The colors are assigned randomly.

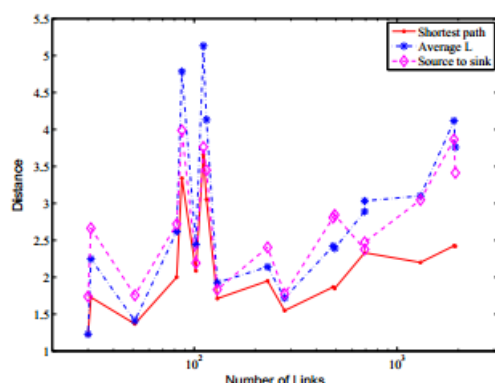


FIG. 3. Three kinds of distances for all collected energetic food webs. All food webs are sorted according to their number of edges in an increasing order

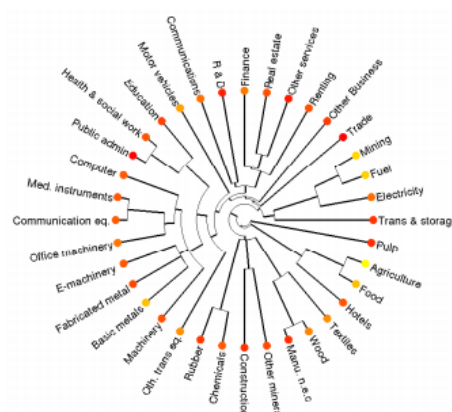


FIG. 4. Hierarchical clustering of different industrial sections in U.S. Colors represent node centrality. All the sector names are abbreviated, and full names can be referred to the Supplementary Material.

6. Hongmei Lei,Ying Chen,Ruiqi Li, Deli He, Jiang Zhang: Maximum Entropy for the International Division of Labor; PLoS ONE 2015, 10(7): e0129955.

简介：As a result of the international division of labor, the trade value distribution on different products substantiated by international trade flows can be regarded as one country' s strategy for competition. Each country wants to diversify their investments on different products as well as make profits as possible as they can. We build a model based on maximum entropy principle to reproduce the distribution curves of countries. The results show that almost all countries' export share distributions can be explained by the maximum entropy model if the constraints are properly selected.

我们可以将每个国家在不同产品上的出口份额看作是一种面向国际市场的劳动分工策略。每个国家都在尽量多样化自己的出口多样性的同时牟取最大的经济利益。本文提出了一个最大化熵模型以解释各个国家在不同产品上的出口份额分布曲线。在合适地选择了最大化熵的约束条件后，我们成功地用一个单参数最大熵模型较好地拟合了 100 多个国家的出口分布曲线。

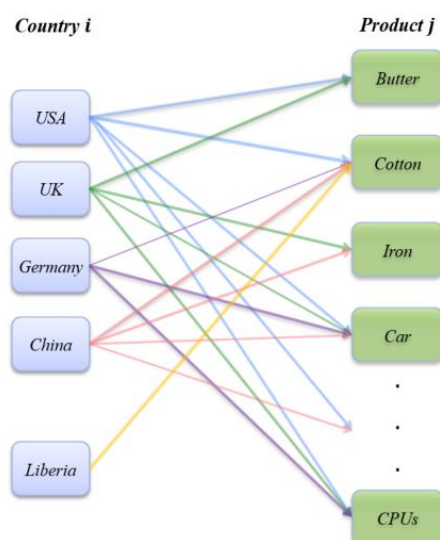


Fig 1. Country-product bipartite network. Country i 's diversification D_i is defined as the degree of country i . The product j 's ubiquity U_j is the degree of j . The weight w_{ij} of the link from country i to product j denoted the trade flow between them. The ratio p_{ij} between w_{ij} and F_i , the total trade values of country i , is defined as the strategy of i .

doi:10.1371/journal.pone.0129955.g001

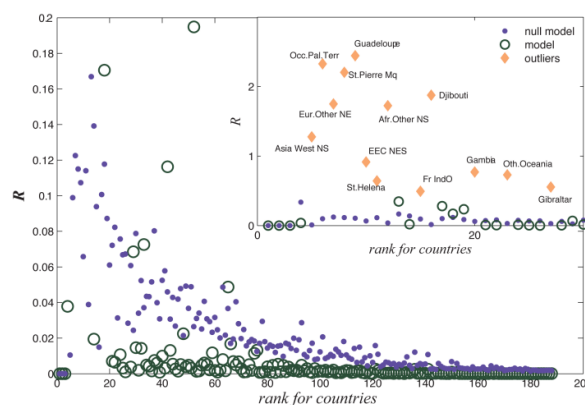


Fig 4. Relative entropies between the theoretical and the empirical results measured by the Kullback-Leibler Distance for all countries. Blue dots are the relative entropies for the Null model and the red hollow circles are those for the theoretical model. Insets show the outliers with very large errors (larger than 0.5).

doi:10.1371/journal.pone.0129955.g004

7. Xintong Li,Xinran Wang,Jiang Zhang, Lingfei Wu: Allometric scaling, size distribution and pattern formation of natural cities; Palgrave Communications, 1, 15017 (2015)

简介： In this paper, we treated connected clusters of nighttime light as natural cities. We then study the allometric scaling laws, Zipf laws, and fractals on these natural cities. A concise model based on geometric matching mechanism is built to reproduce all the observed patterns.

在这篇文章中，我们把夜光形成的连同区域看作一种自然城市。我们系统性地探索了这些自然城市的异速标度律以及尺度分布。最后，我们构造了一个基于“几何匹配”的随机几何图模型成功地浮现出了所有观察到的实证现象。

8. Jiang Zhang, Xintong Li, Xinran Wang, Wen-

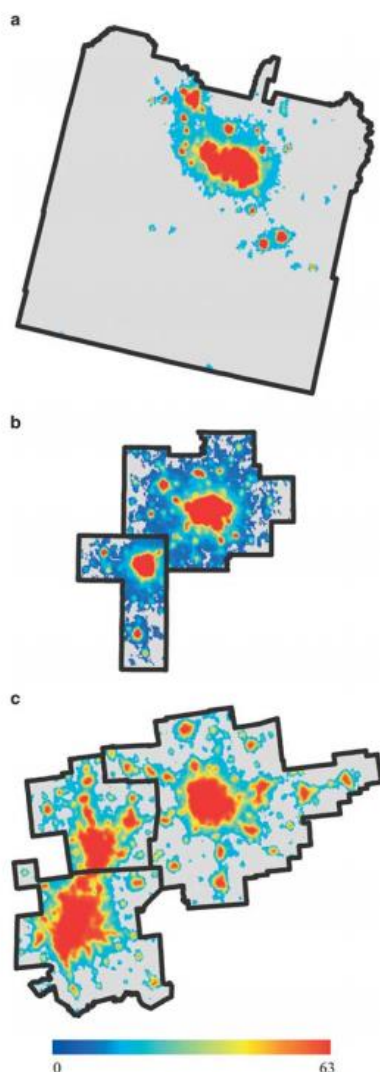


Figure 5 | The comparison between MSAs and natural cities. The nighttime light is plotted inside the administrative boundaries of MSAs including (a) Boise City-Mountain Home-Ontario (ID-OR); (b) Omaha-Council Bluffs-Fremont (NE-IA) and Lincoln-Beatrice (NE); and (c) Columbus-Marion-Zanesville (OH), Dayton-Springfield-Sidney (OH), and Cincinnati-Wilmington-Maysville (OH-KY-IN). Only the nighttime light in (c) is cut off by the low light threshold with $DN_{th} = 4$.

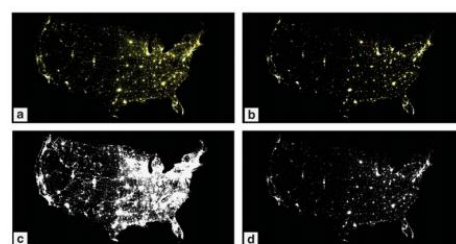


Figure 1 | The nighttime light in natural cities of Contiguous United States (CONUS). The visualization is based on satellite image data (2009) with different thresholds (DN_{th}). The nighttime light cut off by $DN_{th} = 0$ (a), 30 (b). The spatial distribution of natural cities with $DN_{th} = 0$ (c), 30 (d).

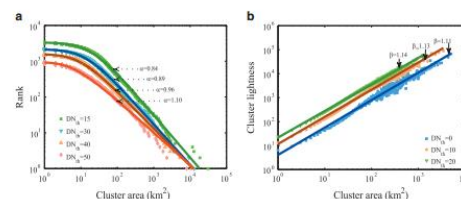


Figure 2 | Examples of rank-size distribution and allometry relationship of natural cities. (a) Rank-size distribution of natural cities in China (2009) with different low light thresholds. (b) Allometric scaling between total lightness and area of natural cities in US (CONUS region, 2009) with different thresholds.

xu Wang, Lingfei Wu: Scaling behaviours in the

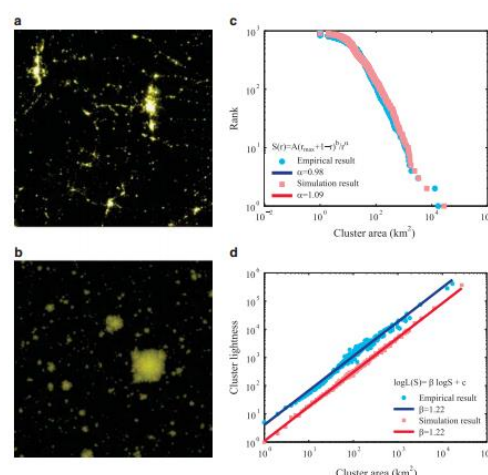


Figure 7 | The comparison between the model results and empirical data with $DN_{th} = 4$ in 2009. (a) Satellite image of nighttime light clusters (or natural cities as defined above) in our ROI from CONUS. (b) The clusters (or natural cities) grown by the model. (c) The size distributions of natural cities for nighttime light image and model simulation. (d) The allometric scaling between total lightness and size of natural cities for nighttime light image and model simulation.

简介: In many networked systems (cities, online communities), links (or interactions) grow faster than nodes, as well, the diversity of nodes grow slower than nodes. We build a simple random network model based on geometric matching mechanism to reproduce both phenomena. The extensive model is further applied to model the distribution of natural cities.

在很多网路系统（在线社区、城市）中，连边（相互作用）总是会比节点以更快速度的增长。与此同时，节点类别多样性会比节点以更慢的速度增长。本文章提出了一个简单的随机几何网络增长模型，同时给出了连边超线性生长和多样性亚线性生长的现象。通过改进该模型，我们还可以模拟城市系统的生长和分布。

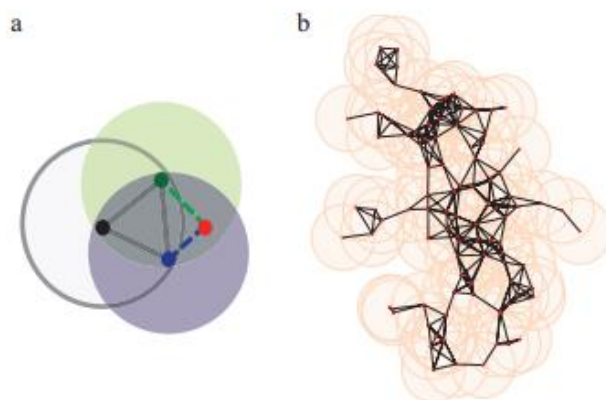


Figure 1 | An illustration of the basic model in $d = 2$ -dimensional space. (a) The filled disks represent existing nodes, the red disk represents a newly added node that will survive. The dark lines represent existing links, and the dashed lines represent the newly added links. The shaded areas represent the interaction regions of existing nodes. (b) A simulation of the basic model after $t = 3625$ steps and $N(t) = 100$ nodes. The shaded area represents $V(t)$ in two-dimensional space. In all simulations $r = 1$.

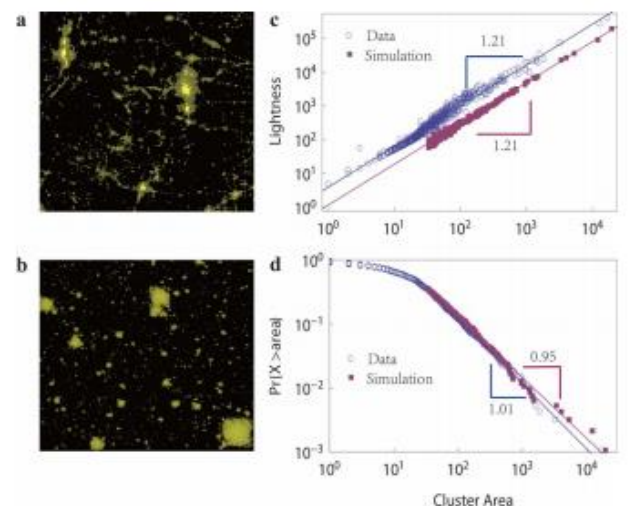


Figure 3 | Comparison of modified SCA model with nighttime light data. (a) Satellite image of a nighttime light distribution. We suppose that each connected cluster in the image is a natural city. (b) The clusters grown by the model using the modified rules ($\varepsilon = 0.03, \alpha = 1.5$). (c) The scaling between the total area and the total light intensity of these clusters both for the nighttime light image (blue circles) and for the image generated by the model (purple squares). In the model, we take the total number of edges of each cluster to represent its total light intensity. (d) Area distributions of nighttime light clusters (blue circles) and the model's clusters (purple squares). In (c) and (d), only clusters with sizes larger than 33 in simulation are shown for comparison. In all simulations, $r = 1$.

9. 曹雪薇, 熊婉婷, 王有贵. 经济金融中的自组织临界性[J]. 现代物理知识, 2015, 03:38-42. 2015-06-25

简介: 复杂性科学在最近二三十年取得了突破性进展,目前从这个学科中发展起来的许多范式和方法都已经渗透到多门传统学科并与其不断融合促成了各种门类的交叉学科和方向。其中自组织临界性就是复杂性科学领域发展出来的一种独特方法,也是理论物理学家对理解复杂系统的一般规律的一个最重要的贡献。在理解和解释由大量个体通过相互作用所导致的复杂系统时空长程相关特性问题上,自组织临界性理论提供了一种非常新颖和通用的机制方法,其应用领域涵盖了物理学、地学、生物学、经济学等多门学科以及它们的交叉方向。本文重点介绍自组织临界模型在经金融系统中的可能应用及其发展前景。

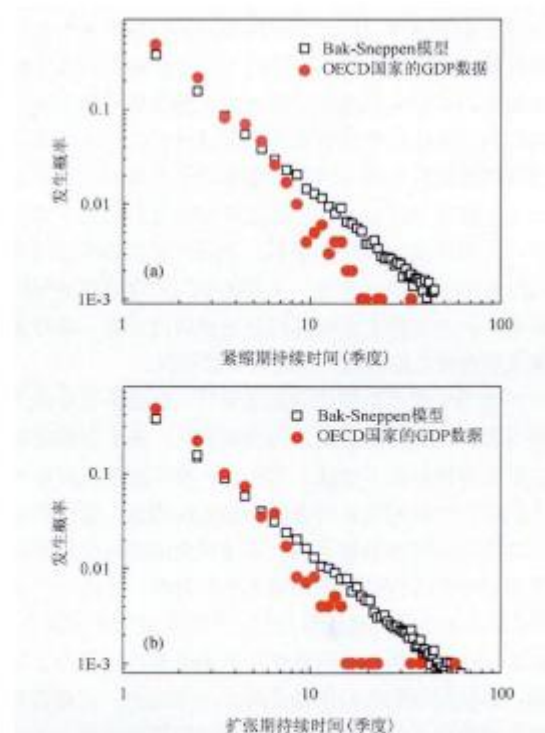


图4 由 Bak-Sneppen 模型和 OECD 数据得到的紧缩期和扩张期持续时间的幂律分布特性的比较。(引自: N. Xi, P. Ormerod, Y. Wang, Technological innovation, business cycles and self-organized criticality in market economies[J], EPL 97(6): 68005 (2012).)

10. 狄增如, 樊瑛. 多元主体合作共治的组织与模式创新[J]. 工程研究-跨学科视野中的工程, 2015, 02:182-187. 2015-06-25

简介: 当前的信息时代和网络社会背景下,社会治理生态具有"复杂性",需要从思维范式、管理模式和体制机制方面创新社会治理。本文基于多元主体合作共治的视角,分别从宏观制度设计及转型途径、多主体行为及其相互影响的微观层次、基层社区管理中多元主体合作共治的组织与模式三个方面,分析了各自层面关注的问题及可能的研究途径,以期探讨多元主体合作共治的治理框架与结构体系。在大数据技术基础上加深对多元主体行为的认识是建立新型社会治理组织与模式的基础;在宏观层面,建立促进社会组织发展和壮大的正反馈机制,是驱动社会治理走向新模式的关键;而在基层社区层次,则需要建立维持新型社区治理模式

稳定和适应性运行的负反馈机制,关注多元主体之间的信息传播与互动模式,特别是基于现代信息网络技术和综合集成方法的互动、协调与决策机制。

11. 张煜霞, 王有贵. 基于 SFC 视角看影子银行的本质[J]. 金融经济, 2015, 16:54-57. 2015-08-25

简介：美国金融危机的爆发揭示出主流经济学忽略货币及金融体系的弊端，而后凯恩斯经济学派发展的存量流量一致方法弥补此缺陷；影子银行被公认为是导致本次危机的罪魁祸首，现有国内外文献大部分集中在对其定义及范围、形式及特征、成因及影响和风险及监管，很少涉及影子银行的本质。运用存量流量一致方法分析影子银行的作用，既验证存量流量一致方法对金融危机的适用性，又回答了影子银行的本质是什么。

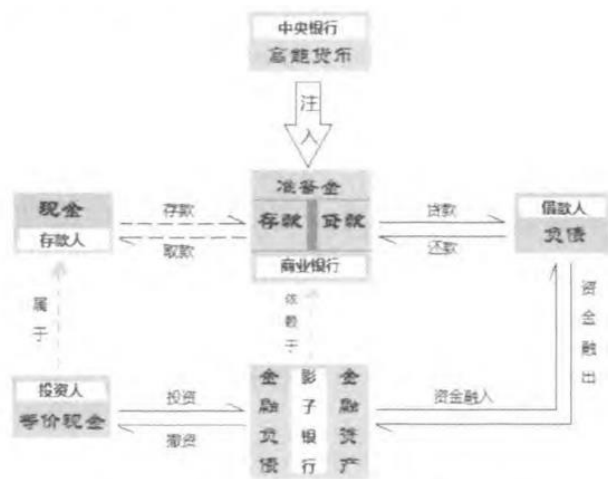


图2 关于影子银行的存量流量分析