

复杂网络研究的机遇与挑战

周涛¹, 张子柯², 陈关荣³, 汪小帆⁴, 史定华⁵, 狄增如⁶, 樊瑛⁶,
方锦清⁷, 韩筱璞², 刘建国⁸, 刘润然², 刘宗华⁹, 陆君安¹⁰, 吕金虎¹¹,
吕琳媛², 荣智海¹, 汪秉宏¹², 许小可¹³, 章忠志¹⁴

(1. 电子科技大学互联网科学中心 成都 611731; 2. 杭州师范大学阿里巴巴复杂科学研究中心 杭州 310036;

3. 香港城市大学电子工程系 中国香港 九龙城区; 4. 上海交通大学电子信息与电气工程学院 上海 闵行区 200240;

5. 上海大学数学系 上海 宝山区 200444; 6. 北京师范大学系统科学学院 北京 海淀区 100875; 7. 中国原子能研究院 北京 房山区 102413;

8. 上海理工大学管理学院 上海 杨浦区 200093; 9. 华东师范大学物理系 上海 闵行区 200241; 10. 武汉大学数学与统计学院 武汉 430072;

11. 中国科学院数学与系统科学研究院 北京 海淀区 100190; 12. 中国科学技术大学物理学院 合肥 230026;

13. 大连民族学院信息与通讯工程学院 辽宁 大连 116605; 14. 复旦大学计算机科学技术学院 上海 杨浦区 200433)

【摘要】该文是2013年在杭州师范大学组织下所召开的复杂网络研讨会基础上的总结和拓展,包含了与会多名学者共同讨论修订后所认可的目前复杂网络研究面临的最主要的10个挑战,这些挑战既是当前复杂网络前沿研究的提炼,又结合了大数据发展的宏观背景。该文旨在为对复杂网络研究感兴趣的青年学者们提供具有参考意义的研究方向和建议。

关键词 大数据; 复杂网络; 功能; 结构

中图分类号 N94

文献标志码 A

doi:10.3969/j.issn.1001-0548.2014.01.001

The Opportunities and Challenges of Complex Networks Research

ZHOU Tao¹, ZHANG Zi-ke², CHEN Guan-rong³, WANG Xiao-fan⁴, SHI Ding-hua⁵, DI Zeng-ru⁶, FAN Ying⁶,
FANG Jin-qing⁷, HAN Xiao-pu², LIU Jian-guo⁸, LIU Run-ran², LIU Zong-hua⁹, LU Jun-an¹⁰, LÜ Jin-hu¹¹, LÜ
Lin-yuan², RONG Zhi-hai¹, WANG Bing-Hong¹², XU Xiao-ke¹³, and ZHANG Zhong-zhi¹⁴

(1. Web Sciences Center, University of Electronic Science and Technology of China Chengdu 611731; 2. Alibaba Research Center for Complexity Sciences, Hangzhou Normal University Hangzhou 310036; 3. Department of Electronic Engineering, City University of HongKong Kowloon HongKong China;

4. School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University Minhang Shanghai 200240; 5. Department of Mathematics, Shanghai University Baoshan Shanghai 200444; 6. School of Systems Science, Beijing Normal University Haidian Beijing 100875; 7. China Institute of Atomic Energy Fangshan Beijing 102413; 8. School of Management, University of Shanghai for Science and Technology Yangpu Shanghai 200093;

9. Institute of Theoretical Physics and Department of Physics, East China Normal University Minhang Shanghai 200241; 10. School of Mathematics and Statistics, Wuhan University Wuhan 430072; 11. Academy of Mathematics and Systems Science, Chinese Academy of Sciences Haidian Beijing 100190;

12. School of Physical Science, University of Science and Technology of China Hefei 230026; 13. School of Information and Communication, Dalian Nationalities University Dalian Liaoning 116605; 14. School of Computer Science, Fudan University Yangpu Shanghai 200433)

Abstract This article is an edited report from a workshop on complex networks organized by the Hangzhou Normal University in 2013, where ten challenging research problems in the field of network science were identified through collective discussions. These ten issues are formulated from the active complex networks research in the present big-data era, aiming to provide important research directions along with sensible suggestions and recommendations to young scholars with potential contributions into future research on complex networks.

Key words big data; complex networks; function; structure

复杂网络的相关研究进入中国已经超过十年了,同时一年一度的全国复杂网络大会也已经成功地召开了九届。在第九届全国复杂网络大会中,有接近600名学者赴杭州参会,盛况空前。在过去的十

年中,复杂网络的很多分支方向受到来自不同研究领域学者们的广泛关注,并极大地推动了复杂网络和复杂性科学的发展。随着信息技术的飞速发展,使得可供研究的数据也越来越丰富——科学研究也

收稿日期: 2013-11-15; 修回日期: 2014-01-02

作者简介: 周涛(1983-),男,教授,主要从事网络科学与工程、信息挖掘、人类动力学和群集动力学等方面的研究。

进入了大数据时代^[1-2], 基于海量数据的定量化分析有可能重塑社会学、心理学、管理学等多个学科的研究范式, 与此同时助推复杂性科学和多个学科的结合和发展。因此, 如何适应并推动大数据时代下复杂网络的研究和发展, 为复杂网络领域的年轻一代指明研究方向成为了当务之急。

通过集中研讨和频繁的邮件交流, 我们根据学科特点和时代背景, 拟定了大数据时代下复杂网络研究所亟需解决的十大问题。这些问题旨在为对复杂网络研究感兴趣的青年学者们提供具有参考意义的研究方向和建议。当然, 复杂网络的研究并不局限于这些方向, 更广阔的领域有待广大学者们在实际研究工作中更深入地加以发掘和补充。

1 从大数据到好网络

随着我们能够收集的数据规模和种类的不断增大, 如何从大数据构建合适的网络也变得日益重要。如何获得高质量的网络结构数据? 如何科学分析数据质量? 基于对不完整的网络结构数据所做的分析在多大程度上能够推广到整个网络?

在这方面的另一个重大挑战是抽样问题。2005年《美国科学院院刊》发表的一篇文章^[3], 证明从任何严格幂律分布的网络抽样所得的子网络都不是严格幂律的。而实际上所用的数据大多都是抽样的结果。反过来, 抽样得到了一个具有幂律度分布的网络, 原网络的分布形式如何——这个问题也没有严格清晰的答案。如何针对所要研究的问题设计合适的抽样算法, 以及如何从抽样网络中预测原始网络的性质仍然面临许多挑战, 因为不同抽样算法下抽样网络和原始网络的关系也不同。与网络采样相对应的一个问题是能否以及如何从不完整数据推得原始实际网络的结构, 特别是如何通过恢复丢失连边和去除虚假连边来还原原始实际网络^[4-5]。

2 网络科学的基本理论

目前关于复杂网络的基本理论遇到了很多困难, 需要有系统且严格的理论去描述网络科学中的概念及规律^[6]。当前最突出的问题就是幂律分布的定义问题^[7]。现在判断一个分布是否为幂律分布, 大多是观察双对数坐标下的分布图是否成一条直线。其实幂律分布可以进一步细分为严格幂律尾部、具有幂律行为和广义幂律关系等, 具体是从概率分布、概率密度函数还是补分布来判断是否为幂律, 都需要数学论证。在这种严格的科学定义没有有效

解决的情况下, 各种相关的研究和讨论就不是严格的。

3 从网络科学到网络工程

针对具体的实际系统和实际问题, 如实际交通系统的优化、传染病的防控等, 提出具体的解决方案, 是网络科学产生社会经济影响的必要途径。在这方面, 虽然暂时可能无法达到理论与实际的统一, 以形成普适性的研究方法和成果, 但是对于实际系统的理解是一个认识逐渐深化的过程。需要在这方面进行逐步的探索, 如实际遇到的超大规模网络、智能电网、移动互联网、物联网等的研究^[8-9]。特别地, 复杂网络分析在社会科学和生命科学领域应该大有可为!

网络科学与网络工程之间的连接桥梁是数学模型的建立、分析与求解。真实网络首先在自然、社会和工程中存在。为了从科学的角度来认识并描述这些网络, 人们开始建立各种各样的网络模型。有了模型, 数学分析、计算、仿真便可以进行。这些科学研究结果被用来解释和预测真实网络的本质、特性和行为, 并建议改造、利用和控制真实网络的策略。目前的研究尚停留在从真实网络到理论模型建立和分析这个阶段, 要回到网络工程去验证和实现科学研究的结果, 还有很长的一段路要走。

4 复杂网络算法分析与设计

算法问题涉及非常广泛, 包含网络结构和动力学的基本分析工具^[10]、社区挖掘问题^[11]、图的匹配问题^[12]、图的模体识别和子图挖掘问题^[13]等。可以说, 算法是进行网络建模和分析的技术基础, 也是包括信息推荐、人工智能优化等一系列具有极为广泛科学意义和应用价值的网络研究工具和手段^[14]。

复杂网络的算法问题, 特别是算法复杂性问题、快速近似算法问题、并行计算问题、分布式存储问题等等, 是在大数据背景下的新挑战——如何快速、准确地处理包含数千万甚至数亿节点的超巨网络? 基于大数据的有效算法分析与设计将成为未来复杂性科学研究的技术基石之一。

5 网络动力学的普适性和差异性

传播、同步、交通、博弈等动力学具备各自的特征, 同时又具有很多共同的特点和相似性。特别地, 它们都是受到信息流或物质流驱动的动力学。已有的研究暗示, 各种相异的网络动力学之间可能

存在着很大的相通性^[15]。在跨越具体动力学种类的层面上, 高屋建瓴地讨论不同类型动力学的普适性和差异性, 特别是建立网络流动力学的全景框架, 具有十分重要的理论和实践意义^[16]。

6 网络信息挖掘和预测

本文主要涉及两个问题, 一个是从动力学表征挖掘网络结构, 这是个反问题; 一个是预测问题, 也就是如何从现有观察中出发, 挖掘网络中的缺失信息, 并进行网络结构、功能和演化趋势预测。在网络结构挖掘反问题的研究上, 一些研究人员注意到了压缩感知的重要性, 该方法最大程度反映了动力学和结构的耦合关系, 具有广泛的应用价值^[17]。在预测问题这个研究方向上, 较成熟也最有活力的研究是链路预测^[18-20]。另外, 关于包含社交关系的人类行为预测^[21], 生物网络中结构链接和功能链接之间的耦合预测等^[22], 都是近期受到广泛关注的发展方向。

7 网络控制、观测与镇定

在很大程度上, 解释现象的目的是为了预测, 随之进行控制。从广义上来说, 科学研究的最终目的就是研究对象进行控制。现代控制论的提出本身是复杂性科学兴起的早期标志之一。近年来, 随着网络可控性问题的提出和探讨, 对复杂网络系统的可控性和控制方法的研究已经成为一个热点问题。复杂网络乃至各种复杂系统是否可控、能否和如何实现控制、如何精确控制、如何实行最小成本控制 and 如何选择最佳控制点, 以及自发和自反馈可控性等问题, 都是非常关键而且有着广泛科学意义和应用价值的研究问题^[23-24]。

作为可控性的对偶概念, 在网络环境下复杂系统的可观测性研究也必不可少^[25]。事实上, 网络可观测性与网络建模、识别、预测、利用, 以及反过来对网络实施控制等任务都息息相关。

动态网络的镇定和同步、一致性问题紧密地联系在一起, 在执行时也在一定程度上依赖于网络的可控性和可观性。对于多尺度和异质的复杂网络来说, 许多经典的理论问题需要作不平凡的推广, 如对多种不同特性网络组成的超网络来说, 存在同时具有不同尺度和不同性质(如连续与离散、光滑与脉冲)的混合型Lyapunov函数吗? 如何去构造?

8 含有时空信息的网络以及网络的网络

现实的网络大多数是随时间和空间持续变化

的, 结构完全固定的网络非常少见。如实际的社会网络中, 人与人之间的联系与交互是遵循一定时空统计规律出现, 而不是一直保持不变的。这使得现实的社会网络传播过程中, 在连边上有效传播的出现总是暂态进行, 并且总的来说, 距离较远的连接发生感染的可能性较小, 但是往往感染后危害性却较大(导致疾病的远距离扩散)。在这种含有时间和空间的新型网络上的动力学过程中, 可能会呈现出与静态网络和非空间网络极为不同的规律。因此, 需要探索这种随时空演化的动态网络上的动力学特性, 以及节点、连边的活跃特性与动力学的关联规律。这是一个重要的科学问题, 也有望促进复杂网络研究的理论成果与真实应用的结合^[26-28]。

目前网络科学研究主要针对的是单个网络, 而事实上许多网络都不是孤立存在的, 而是与其它网络之间存在着相互依赖、合作或竞争等关系^[29], 即“网络的网络”, 如物联网。随着数据获取能力的不断增强, 可以对网络的开展从理论到应用的深入研究。

以往的研究挖掘出了众多的具有一定普适性的网络统计特性, 如无标度特性、小世界特性、社团性等, 并研究得到了对这些统计特性的产生规律和动力学效应的基本认知。同时研究也发现, 现实网络的结构极为丰富。以往的研究更多关注了这些网络在统计意义上的共性问题, 而实际网络中, 即使两种网络有着高度相似的统计特性, 也有可能出现迥异的动力学现象。对于不同网络, 例如无向网络、含权网络、有向网络、多部分网络、时空网络乃至网络的网络等, 它们的普适性和差异性也是一个值得注意的基本问题。总的来说, 复杂网络的研究热点正在经历从挖掘不同类型网络普适性到发现不同类型网络独特性的转型中。

9 网络中微观个体的角色界定

现实网络大多是高度异质性的, 不同节点和连边在网络动力学演化过程中所扮演的角色往往差异巨大。这使得如何有效地衡量节点、连边的重要性, 以及节点、连边在相应动力学过程中所扮演的角色和影响, 成为一个重要的研究问题。具体来说, 需要针对不同的动力学, 判断何种类型的节点以及何种类型的边分别扮演的不同角色; 以及挖掘在给定的动力学规则下, 哪些节点或连边扮演最关键的作用^[30-32]。除了网络动力学之外, 不同的节点和连边在保持网络结构在演化过程中的连通性和稳定性方

面所起到的作用也不尽相同^[33-34]。

10 从微观到宏观结构的自组织演化问题

复杂性科学研究的核心理念之一就是深刻理解连接微观基本单元到宏观复杂结构和统计规律的涌现机制。可以说,演化和动力学是分不开的——从本质上来说,演化自身就是一种动力学行为。当前复杂网络研究中的众多问题都与该问题相关,如宏观动力学过程、宏观结构和宏观统计规律的涌现等。网络动力学的一个区别于经典大系统理论的显著特征是涌现。网络动力现象不是各个节点动力系统行为的简单叠加,而是具有繁殖、变化和增长能力的交织与融合,最后导致本质性的突变和新行为的涌现。

此外,很多学者都注意到,直接找到从微观个体到全网的宏观结构比较困难,两个端点之间尚需要一个中间过渡,也就是以模体(motif)^[35]和群落(community)^[11]为代表的中尺度结构。这些中尺度结构的分布、在动力学过程中的作用,以及从微观到中观、从中观到宏观的演化和涌现过程,都值得高度关注^[36]。

2013年5月11日,由杭州师范大学阿里巴巴商学院信息经济研究所、阿里研究中心、阿里巴巴复杂科学研究共同组织的“大数据时代下复杂网络的机遇与挑战”研讨会在杭州市香格里拉大酒店举行,本文系该会议讨论基础上的补充和延伸。会议由周涛主持并最终定稿,成稿期间,张子柯、陈关荣、汪小帆和史定华参与了撰写与修订,其余作者均在会议期间对本文的内容提出了实质性的贡献。本文修改过程中,周银座、刘闯、潘黎明都提出了宝贵意见,在此一并致谢。

参 考 文 献

[1] MARX V. The big challenges of big data[J]. *Nature*, 2013, 498(7453): 255-260.

[2] 周涛. 大数据: 商业革命和科学革命[J]. 半月谈: 时事资料手册, 2013, 98(4): 14-17.
ZHOU Tao. Big data: Commercial revolution and scientific revolution[J]. *China Comment: Current data handbook*, 2013, 98(4): 14-17.

[3] STUMPF M P H, WIUF C, MAY R M. Subnets of scale-free networks are not scale-free: sampling properties of networks [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2005, 102(12): 4221-4224.

[4] GUIMERA R, SALES-PARDO M. Missing and spurious

interactions and the reconstruction of complex networks[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(52): 22073-22078.

[5] ZENG A, CIMINI G. Removing spurious interactions in complex networks[J]. *Physical Review E*, 2012, 85(3): 036101.

[6] 史定华. 无标度网络: 基础理论和应用研究[J]. *电子科技大学学报*, 2010, 39(5): 644-650.
SHI Ding-hua. Scale-free networks: basic theory and applied research[J]. *Journal of University of Electronic Science and Technology of China*, 2010, 39(5): 644-650.

[7] 史定华. 网络度分布理论[M]. 北京: 高等教育出版社, 2011.
SHI Ding-hua. *Theory of network degree distributions*[M]. Beijing: Higher Education Press, 2011.

[8] COSTA L D F, OLIVEIRA J O N, TRAVIESO G, et al. Analyzing and modeling real-world phenomena with complex networks: a survey of applications[J]. *Advances in Physics*, 2011, 60(3): 329-412.

[9] 汪小帆, 李翔, 陈关荣. 网络科学导论[M]. 北京: 高等教育出版社, 2012.
WANG Xiao-fan, LI Xiang, CHEN Guan-rong. *Network science: an introduction*[M]. Beijing: Higher Education Press, 2012.

[10] GÓMEZ S, ARENAS A, BORGE-HOLTHOEFER J, et al. Discrete-time Markov chain approach to contact-based disease spreading in complex networks[J]. *EPL (Europhysics Letters)*, 2010, 89(3): 38009.

[11] FORTUNATO S. Community detection in graphs[J]. *Physics Reports*, 2010, 486(3): 75-174.

[12] XUAN Q, WU T J. Node matching between complex networks[J]. *Physical Review E*, 2009, 80(2): 026103.

[13] AGGARWAL C C, WANG H. *Managing and mining graph data*[M]. New York: Springer, 2010.

[14] LÜ L, MEDO M, YEUNG C H, et al. Recommender Systems[J]. *Physics Reports*, 2012, 519(3): 1-49.

[15] 吕琳媛, 陆君安, 张子柯, 等. 复杂网络观察[J]. *复杂系统与复杂性科学*, 2010, 7(2-3): 173-186.
LÜ Lin-yuan, LU Jun-an, ZHANG Zi-ke, et al. Looking into complex networks[J]. *Complex Systems and Complexity Science*, 2010, 7(2-3): 173-186.

[16] BARZEL B, BARABÁSI A L. Universality in network dynamics[J]. *Nature Physics*, 2013(9): 673-681.

[17] WANG W X, YANG R, LAI Y C, et al. Predicting catastrophes in nonlinear dynamical systems by compressive sensing[J]. *Physical Review Letters*, 2011, 106(15): 154101.

[18] 吕琳媛. 复杂网络链路预测[J]. *电子科技大学学报*, 2010, 39(5): 651-661.
LÜ Lin-yuan. Link prediction on complex networks[J]. *Journal of University of Electronic Science and Technology of China*, 2010, 39(5): 651-661.

[19] 吕琳媛, 周涛. 链路预测[M]. 北京: 高等教育出版社, 2013.
LÜ Lin-yuan, ZHOU Tao. *Link prediction*[M]. Beijing: Higher Education Press, 2013.

[20] LÜ Lin-yuan, ZHOU Tao. Link prediction in complex

- networks: a survey[J]. *Physica A*, 2011, 390(6): 1150-1170.
- [21] AIELLO L M, BARRAT A, SCHIFANELLA R, et al. Friendship prediction and homophily in social media[J]. *ACM Transaction on the Web*, 2012, 6(2): 9.
- [22] HONEY C J, SPORNS O, CAMMOUN L, et al. Predicting human resting-state functional connectivity from structural connectivity[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(6): 2035-2040.
- [23] LIU Y Y, SLOTINE J J, BARABÁSI A L. Controllability of complex networks[J]. *Nature*, 2011(473): 167-173.
- [24] 陈关荣. 复杂动态网络环境下控制理论遇到的问题与挑战[J]. *自动化学报*, 2013, 39(4): 312-321.
CHEN Guan-rong. Problems and challenges in control theory under complex dynamical network environments[J]. *Acta Automatica Sinica*, 2013, 39(4): 312-321.
- [25] LIU Y Y, SLOTINE J J, BARABÁSI A L. Observability of complex systems[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(7): 2460-2465.
- [26] HOLME P, SARAMÁKI J. Temporal networks[J]. *Physics Reports*, 2012, 519(3): 97-125.
- [27] BARTHÉLEMY M. Spatial networks[J]. *Physics Reports*, 2011, 499(1): 1-101.
- [28] 周涛, 韩筱璞, 闫小勇, 等. 人类行为时空特性的统计力学[J]. *电子科技大学学报*, 2013, 42(7): 481-540.
ZHOU Tao, HAN Xiao-pu, YAN Xiao-yong, et al. Statistical mechanics on temporal and spatial activities of human[J]. *Journal of University of Electronic Science and Technology of China*, 2013, 42(7): 481-540.
- [29] GAO J X, BULDYREV S V, STANLEY H E, et al. Networks formed from interdependent networks[J]. *Nature Physics*, 2012, 8(1): 40-48.
- [30] KITSACK M, GALLOS L K, HAVLIN S, et al. Identification of influential spreaders in complex networks[J]. *Nature Physics*, 2010, 6(11): 888-893.
- [31] LÜ L, ZHANG Y C, YEUNG C H, et al. Leaders in social networks, the delicious case[J]. *PLoS ONE*, 2011, 6(6): e21202.
- [32] CHEN D B, LÜ L, SHANG M S, et al. Identifying influential nodes in complex networks[J]. *Physica A*, 2012, 391(4): 1777-1787.
- [33] CSERMELY P. Weak links: stabilities of complex systems from proteins to social networks[M]. Berlin: Springer-Verlag, 2006.
- [34] CHENG X Q, REN F X, SHEN H W, et al. Bridgeness: a local index on edge significance in maintaining global connectivity[J]. *Journal of Statistical Mechanics: Theory and Experiment*, 2010: P10011.
- [35] MILO R, SHEN-ORR S, ITZKOVITZ S, et al. Network motifs: simple building blocks of complex networks[J]. *Science*, 2002, 298(5594): 824-827.
- [36] 陈娟, 陆君安. 复杂网络中尺度研究揭开网络同步化过程[J]. *电子科技大学学报*, 2012, 41(1): 8-16.
CHEN Juan, LU Jun-an. Mesoscales reveal synchronization processes in complex networks[J]. *Journal of University of Electronic Science and Technology of China*, 2012, 41(1): 8-16.

编辑 蒋晓