

# 复杂网络局部结构的涌现： 共同邻居驱动网络演化\*

崔爱香<sup>1)</sup> 傅彦<sup>1)</sup> 尚明生<sup>1)</sup> 陈端兵<sup>1)</sup> 周涛<sup>1)2)3)†</sup>

1)(电子科技大学计算机科学与工程学院, 互联网科学中心, 成都 611731)

2)(弗里堡大学物理系, 瑞士弗里堡 CH-1700)

3)(中国科学技术大学近代物理系, 合肥 230026)

(2010年8月8日收到; 2010年9月19日收到修改稿)

在对真实网络的小世界和无标度特性进行了大量深入考量之后, 最近的研究热点开始转移到更加细致的局部结构. 实证数据显示, 大量真实网络具有幂律的低阶集团度分布. 这一普适的规律, 无法由富者愈富以及熟人推荐的网络生长机理再现. 本文提出一种由共同邻居驱动的网络演化模型, 该模型能够重现实证研究所观察到的幂律集团度分布, 暗示共同邻居驱动是复杂网络局部结构涌现形成的内在机理.

**关键词:** 复杂网络, 演化模型, 集团度分布, 共同邻居

**PACS:** 89.20.Hh, 89.75.Fb

## 1. 引言

近年来, 复杂网络的研究蓬勃发展, 已经渗透到从物理学到计算机科学, 从生命科学到社会科学的各个学科<sup>[1-11]</sup>. 其中, 对复杂网络结构演化规律的实证分析, 以及相应的建模研究, 是充分认识一切有关复杂网络的功能与应用的基础. 在复杂网络研究的早期, 研究人员主要关注网络最基本的宏观特性<sup>[12,13]</sup>, 例如小世界现象<sup>[14]</sup>、无标度特性<sup>[15]</sup>和网络连接的成团性质<sup>[16]</sup>等等. 随着对复杂网络演化模型研究的深入, 近年来, 研究的焦点开始转移到更为细致的网络局部结构, 例如对网络中模块<sup>[17]</sup>、环<sup>[18]</sup>、紧密子图<sup>[19]</sup>等结构的统计分析.

最近, 周涛等人提出了利用“集团度”刻画网络局部结构的设想<sup>[20,21]</sup>. 在无向网络中,  $m$  阶集团指由  $m$  个节点构成的完全子图 ( $m > 1$ ), 而一个节点的  $m$  阶集团度是指包含该节点的不同  $m$  阶集团的数目. 显然, 一个节点的 2 阶集团度等于该节点的

度, 3 阶集团度等于包含该节点的不同三角形的数目……从这个意义上讲, 集团度可以看作对“度”这个概念的一种推广. 周涛等人的实证研究揭示了两个普适的规律<sup>[20,21]</sup>: 1) 从计算机互联网和万维网, 再到社会网络和生物网络, 大量真实网络都具有幂律的低阶集团度分布; 2) 随着所统计的集团阶数的上升, 其相应的集团度分布的幂指数呈现下降的趋势.

如后文将要展示的, 广为人知的“富者愈富”机理<sup>[15]</sup>只能再现 2 阶集团度(也就是度)的幂律分布, 而“熟人推荐”机理<sup>[16]</sup>虽然能够生成具有很大簇系数的网络, 但是所生成的网络即便达到 10 万个节点的规模, 也基本没有 5 阶或 5 阶以上的集团. 最近, 杨涵新等人提出了一个基于随机行走的网络模型<sup>[22]</sup>, 能够生成幂律集团度分布. 遗憾的是, 该模型生成规则与真实网络演化机理之间的关联关系尚不清楚, 并且所得到的集团度分布指数随阶数呈上升趋势, 与真实情况恰恰相反. 本文提出一种全新的网络演化机理——共同邻居驱动

\* 国家自然科学基金重大研究计划(批准号:90924011), 国家自然科学基金重点项目(批准号:10635040)和国家自然科学基金(批准号:60973069)资助的课题.

† 通讯联系人. E-mail: zhutou@ustc.edu

机理,并讨论了为何该机理广泛存在于从互联网到社会网络的大量真实系统中.数值模拟显示,该机理能够重现实证研究所观察到的幂律集团度分布,且指数变化规律与真实观察不谋而合.本文的结果暗示共同邻居驱动是复杂网络局部结构涌现形成的内在机理.

## 2. 网络演化模型

在复杂网络的研究中,各种各样的机理被提出来模拟真实网络的演化生长.从目前的研究现状来看,并不存在某种或某几种“放之四海皆准”的演化机理,与此相反,不同类型的网络可能有着完全不同的生长机理,即便它们表现出了非常相似的统计特性.举个例子来说,无标度网络结构的涌现就被公认为具有多种可能的原因<sup>[23]</sup>.即便如此,寻找可能表征部分真实网络的演化规律并建立相应模型,一直是推动复杂网络演化模型研究的根本动力<sup>[24-29]</sup>.

对网络真实演化过程的实证研究有助于我们了解网络生长可能的内在驱动力. Kossinets 和 Watts 研究了在校大学生之间的熟人关系网络<sup>[30]</sup>,发现两名原本没有社会关系的学生在将来是否会建立关系很大程度上受到他们当前共同熟人数目的影响:拥有越多共同熟人的学生在将来具有更大的可能性成为熟人. Liben - Nowell 和 Kleinberg<sup>[31]</sup>以及周涛、吕琳媛和张翼成<sup>[32]</sup>研究了大量刻画节点接近性的定量指标,发现两个节点的共同邻居数目越多,它们之间存在直接连接的可能性就越大.这些实证结果都暗示我们,共同邻居可能是诱导两个节点建立连接的原因.这也非常直观,因为在社会网络中拥有很多共同朋友的人当然有较大的可能性通过这些朋友的引介互相认识而成为朋友,而在万维网中,如果一个网页的管理员发现在该网页能够链接到的很多网页中都出现了某特定网页的链接,那么他很可能会去访问这个特定网页并将其加入为自己网页的新链接.

本文充分考虑共同邻居的驱动效应,其模型的具体规则如下:

1) 初始时网络中包含  $m_0$  个两两相互连接的节点.

2) 每次引入一个新节点和  $m$  条边,其中  $m \leq m_0$ .

3) 新引入的节点首先采用优先连接的机理与网络中已存节点建立一条连边.连边建立的概率正比于此连边所指向的节点的度.

4) 基于共同邻居驱动的机理添加其余的  $m - 1$  条边,考虑网络中所有未连接的节点对,其产生连边的概率正比于这对节点的共同邻居数  $| \Gamma_i \cap \Gamma_j |$ ,即两个未连接的节点共同邻居数越多产生连接的概率越大,其中  $\Gamma_i$  和  $\Gamma_j$  分别是节点  $i$  和  $j$  的邻点集.注意,这  $m - 1$  条边往往与新节点并无关系.

为了通过比较体现本模型的优越之处,我们同时参考了 Barabási - Albert (BA) 模型<sup>[15]</sup> 和 Holme - Kim (HK) 模型<sup>[16]</sup> 的结果.这两个模型的规则简介如下. BA 模型最初也是由  $m_0$  个两两相互连接的节点开始,每一个时间步加入一个新节点,与网络中已存在的  $m$  个节点相连,并且优先与度大的节点建立连接,即新节点和网络中已存节点建立连边的概率与已存节点的度成正比关系. HK 模型可生成具有可调簇系数的无标度小世界网络,是 BA 模型的一种改进. HK 模型的规则与 BA 模型类似,只是加入  $m$  条边的过程与 BA 模型有所不同.首先, HK 模型按照 BA 优先连接的规则加入一条边,以后的  $(m - 1)$  条边,要么以概率  $(1 - p)$  都以 BA 优先连接的规则加入,要么以概率  $p$  每次都随机选择第一条边所连接的节点的一个邻居节点,并建立连接.显然,当  $p = 0$  时, HK 模型退化到 BA 模型.

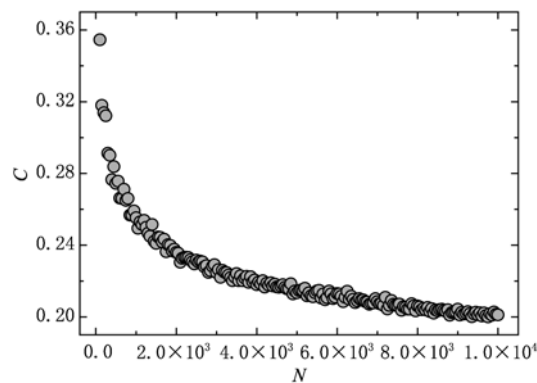


图1 本文所提模型中簇系数随网络规模的变化

### 3. 数值实验结果

数值实验表明,当  $m_0$  和  $m$  远小于网络最终规模  $N$  的时候,初始条件对最终的结果没有影响,所以本文在针对三个模型的数值实验中均取  $m_0 = m = 3$ . 当生成的网络规模  $N = 10000$  时,我们提出的模型与 BA 及 HK 模型都有 29994 条边. BA 网络及 HK 网络都是小世界网络,但是 BA 网络集聚系数随网络规模增大会迅速减小<sup>[33]</sup>,而 HK 模型在  $p$  较大时具有较大的簇系数<sup>[16]</sup>. 与之比较,图 1 和图 2 分别给出了本文所提模型的簇系数和平均距离随网络规模的变化规律,其中当网络规模为  $10^4$  时,簇系数在 0.20 左右,而平均距离的增长呈网络规模的对数趋势,表现出明显的小世界效应.

图 3 给出了本文所提模型 2 至 5 阶集团度分布,可以看到,这些分布都具有幂律性质. 我们利用最大似然估计<sup>[34]</sup>的方法得到了度分布的指数,分别为 1.93, 1.46, 1.16 和 0.99,呈现出随着阶数上升递减的趋势,这 and 实际观察<sup>[21]</sup>是完全符合的. 然而,不管是 BA 网络(如图 4 所示)还是 HK 网络(如图 5

所示),当集团阶数提高时,分布形式很快偏离幂律,而且不存在 5 阶集团. 数值实验说明二阶集团度分布(也就是度分布)的幂函数律与大的簇系数并不能保证更高阶的幂律集团度分布. 通过比较数值试验结果和真实网络的特征<sup>[21]</sup>,我们认为优先连接<sup>[15]</sup>和熟人推荐<sup>[16]</sup>机理并不能很好地刻画真实网络的局部结构,本文所提模型能够重现实证研究所观察到的幂律集团度分布,暗示共同邻居驱动是网络局部结构涌现形成的重要机理.

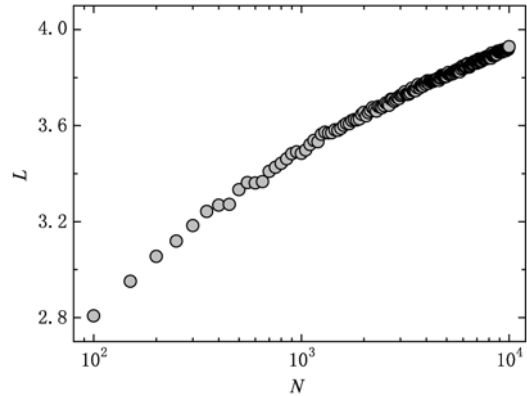


图 2 本文所提模型中平均距离随网络规模的变化

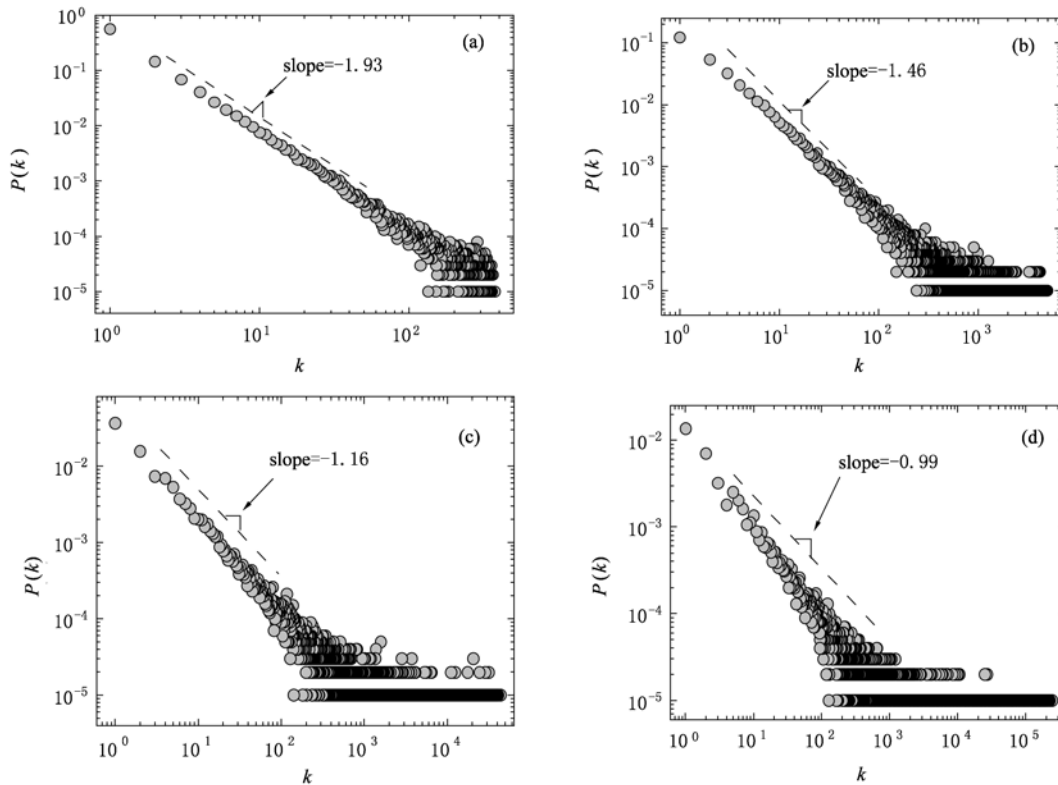


图 3 (a) — (d) 分别对应于本文所提模型的 2 至 5 阶集团度分布. 分布指数分别为 1.93, 1.46, 1.16 和 0.99, 呈现递减趋势, 和实证观察符合

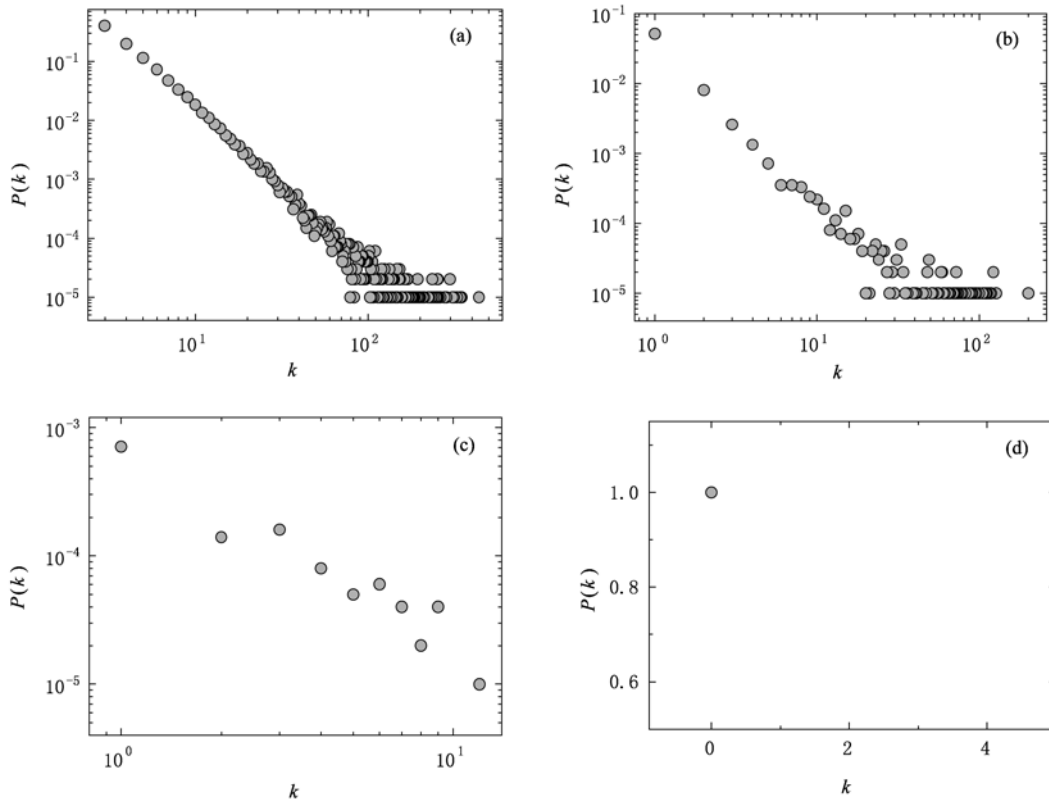


图4 (a)–(d)分别对应于BA网络2至5阶集团度分布

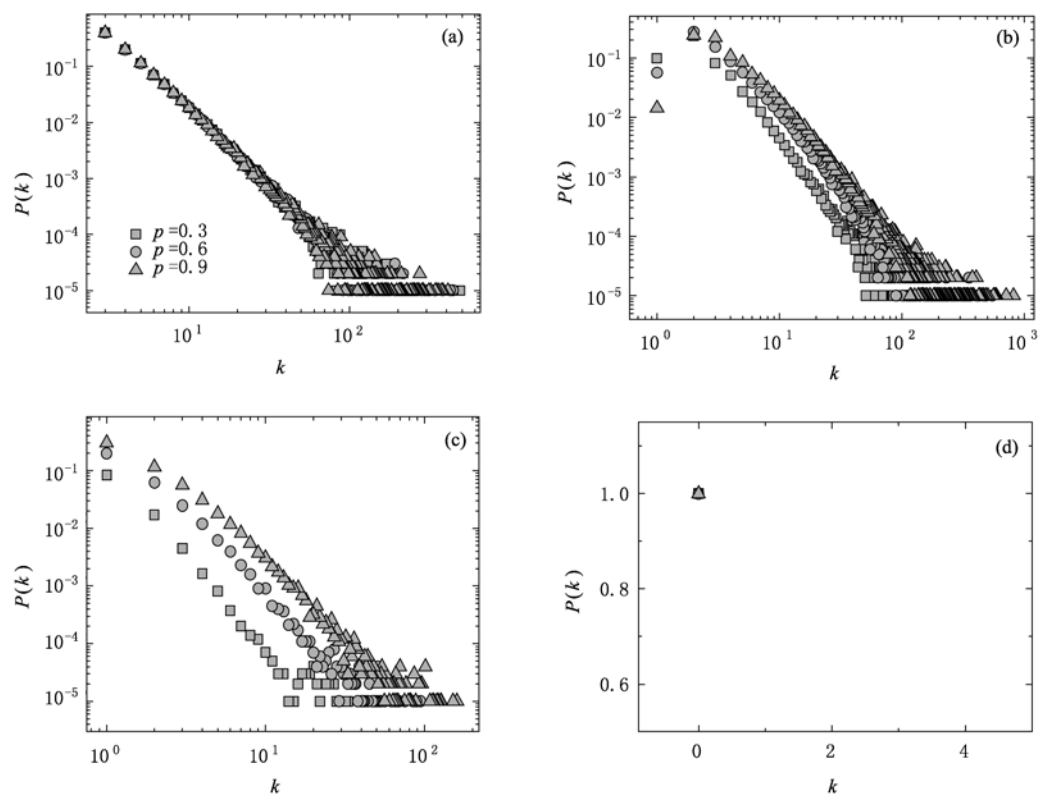


图5 (a)–(d)分别对应于HK网络2至5阶集团度分布,圆圈、正方形和三角形分别代表 $p$ 值为0.3,0.6和0.9三种情况

## 4. 结 论

在复杂网络演化模型研究的早期,绝大部分模型都只关注网络的主要宏观性质<sup>[15,16]</sup>,例如无标度特性和小世界效应.就无标度特性而言,已知的机理就数不胜数<sup>[23]</sup>,例如富者愈富机理<sup>[15]</sup>,好者变富机理<sup>[35]</sup>,优化设计驱动<sup>[36]</sup>,哈密顿动力学驱动<sup>[37]</sup>,聚生机理<sup>[38]</sup>,稳定性限制驱动<sup>[39]</sup>,等等.仅仅从宏观性质出发,难以给出对不同演化机理的可信的评估.因此,深入挖掘网络细致的统计性质,特别是局部结构特性,并利用这些统计性质对已知演化机理进行更严格更准确的评估,是网络演化建模研究向前发展的必然趋势.集团度是典型的用于刻画节点局部环境的指标,有趣的是,大量真实网络都表现

出幂律的集团度分布,但目前还没有简单的机理能够再现这种分布形式,特别是分布指数随集团阶数下降的性质.本文中作为参考的优先连接机理<sup>[15]</sup>和熟人推荐机理<sup>[16]</sup>,亦无法再现幂律的集团度分布.本文考虑了真实网络中广泛存在的“共同邻居驱动”机理,相应的演化模型能够很好再现所观测到的幂律集团度分布,且分布指数也随着集团阶数的增长而下降.

本文提供了研究网络局部结构形成机理的范例,所提出的“共同邻居驱动”的演化机理,符合我们对真实网络的认知<sup>[30-32]</sup>,具有广泛的解释力,很可能揭示了大量真实网络演化生长的最重要的机理.进一步地,通过在此模型上讨论传播、交通、同步、博弈等问题,可以观察到局部结构对网络动力学性质的影响.

- [1] Wang X F, Chen G R 2006 *Complex Networks Theory and its Applications* (Beijing: Tsinghua University Press) (in Chinese) [汪小帆、李翔、陈关荣 2006 复杂网络理论及其应用(北京:清华大学出版社)]
- [2] He D R, Lin Z H, Wang B H 2008 *Complex Systems and Complex Networks* (Beijing: Higher Education Press) (in Chinese) [何大韧、刘宗华、汪秉宏 2008 复杂系统与复杂网络(北京:高等教育出版社)]
- [3] Wu J S, Di Z R 2004 *Prog. Phys.* **24** 18 (in Chinese) [吴金闪、狄增如 2004 物理学进展 **24** 18]
- [4] Fang J Q 2006 *Sci. Techn. Rev.* **24** 67 (in Chinese) [方锦清 2006 科技导报 **24** 67]
- [5] Zhou T, Bai W J, Wang B H, Liu Z J, Yan G 2005 *Physics* **34** 31 (in Chinese) [周涛、柏文洁、汪秉宏、刘之景、严钢 2005 物理 **34** 31]
- [6] Li J, Wang B H, Jiang P Q, Zhou T, Wang W X 2006 *Acta Phys. Sin.* **55** 4051 (in Chinese) [李季、汪秉宏、蒋品群、周涛、王文旭 2006 物理学报 **55** 4051]
- [7] Xu D, Li X, Wang X F 2007 *Acta Phys. Sin.* **56** 1313 (in Chinese) [许丹、李翔、汪小帆 2007 物理学报 **56** 1313]
- [8] Liu H K, Zhang X L, Cao L, Wang B H, Zhou T 2009 *Sci. China Ser. G* **39** 935 (in Chinese) [刘宏鲲、张效莉、曹崑、汪秉宏、周涛 2009 中国科学 G 辑 **39** 935]
- [9] Zhang P P, He Y, Zhou T, Su B B, Chang H, Zhou Y P, Wang B H, He D R 2006 *Acta Phys. Sin.* **55** 60 (in Chinese) [张培培、何阅、周涛、苏蓓蓓、常慧、周月平、汪秉宏、何大韧 2006 物理学报 **55** 60]
- [10] Weng W G, Ni S J, Shen S F, Yuan H Y 2007 *Acta Phys. Sin.* **56** 1938 (in Chinese) [翁文国、倪顺江、申世飞、袁宏永 2007 物理学报 **56** 1938]
- [11] Liu H K, Zhou T 2007 *Acta Phys. Sin.* **56** 106 (in Chinese) [刘宏鲲、周涛 2007 物理学报 **56** 106]
- [12] Albert R, Barabasi A L 2002 *Rev. Mod. Phys.* **74** 47
- [13] Dorogovtsev S N, Mendes J F F 2002 *Adv. Phys.* **51** 1079
- [14] Watts D J, Strogatz S H 1998 *Nature* **393** 440
- [15] Barabasi A L, Albert R 1999 *Science* **286** 509
- [16] Holme P, Kim B J 2002 *Phys. Rev. E* **65** 026107
- [17] Alon U 2002 *Nat. Rev. Gene* **8** 450
- [18] Bianconi G, Caldarelli G, Capocci A 2008 *Phys. Rev. Lett.* **100** 118701
- [19] Vazquez A, Dobrin R, Sergi D, Eckmann J P, Oltvai Z N, Barabasi A L 2004 *Proc. Natl. Acad. Sci. USA* **101** 17940
- [20] Zhou T, Xiao W K, Ren J, Wang B H 2007 *Complex Syst. Complex Sci.* **4** 10 (in Chinese) [周涛、肖伟科、任捷、汪秉宏 2007 复杂系统与复杂性科学 **4** 10]
- [21] Xiao W K, Ren J, Qi F, Zhu M X, Yang H F, Jin H Y, Wang B H, Zhou T 2007 *Phys. Rev. E* **76** 037102
- [22] Yang H X, Wang B H, Liu J G, Han X P, Zhou T 2008 *Chin. Phys. Lett.* **25** 2718
- [23] Caldarelli G 2007 *Scale-free networks: complex webs in nature and technology* (New York: Oxford University Press) p112
- [24] Barabási A L 2009 *Science* **325** 412
- [25] Xing C M, Liu F A 2010 *Acta Phys. Sin.* **59** 1608 (in Chinese) [邢长明、刘方爱 2010 物理学报 **59** 1608]
- [26] He M H, Zhang D M, Wang H Y, Li X G, Fang P J 2010 *Acta Phys. Sin.* **59** 5175 (in Chinese) [何敏华、张端明、王海艳、李小刚、方频捷 2010 物理学报 **59** 5175]
- [27] Wang G Z, Cao Y J, Bao Z J, Han Z X 2009 *Acta Phys. Sin.* **58** 3597 (in Chinese) [王光增、曹一家、包哲静、韩祯祥 2009 物理学报 **58** 3597]

- [28] Ou Y M, Fei Q, Yu M H 2008 *Acta Phys. Sin.* **57** 6763 (in Chinese) [ 欧阳敏、费奇、余明晖 2008 物理学报 **57** 6763 ]
- [29] Tan S, Dai G Z, Wang L, Fan M 2007 *Acta Phys. Sin.* **56** 6326 (in Chinese) [ 覃森、戴冠中、王林、范明 2007 物理学报 **56** 6326 ]
- [30] Kossinets G, Watts D J 2006 *Science* **311** 88
- [31] Liben – Nowell D, Kleinberg J M 2007 *J. Am. Soc. Inf. Sci. Tech.* **58** 1019
- [32] Zhou T, Lü L, Zhang Y C 2009 *Eur. Phys. J. B* **71** 623
- [33] Klemm K, Eguiluz V M 2002 *Phys. Rev. E* **65** 057102
- [34] Goldstein M L, Morris S A, Yen G G 2004 *Eur. Phys. J. B* **41** 255
- [35] Garlaschelli D, Capocci A, Caldarelli G 2007 *Nature Physics.* **3** 813
- [36] Valverde S, Cancho R F, Sole R V 2002 *Euro. Phys. Lett.* **60** 512
- [37] Baiesi M, Manna S S 2003 *Phys. Rev. E* **68** 047103
- [38] Kim B J, Trusina A, Minnhagen P, Sneppen 2005 *Eur. Phys. J. B* **43** 369
- [39] Perotti J I, Billoni O V, Tamarit F A, Chialvo D R, Cannas S A 2009 *Phys. Rev. Lett.* **103** 10870

## Emergence of local structures in complex network: common neighborhood drives the network evolution\*

Cui Ai-Xiang<sup>1)</sup> Fu Yan<sup>1)</sup> Shang Ming-Sheng<sup>1)</sup> Chen Duan-Bing<sup>1)</sup> Zhou Tao<sup>1)2)3)†</sup>

1) ( *Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 611731, China* )

2) ( *Department of Physics, University of Fribourg, Fribourg 1700, Switzerland* )

3) ( *Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China* )

( Received 8 August 2010; revised manuscript received 10 September 2010 )

### Abstract

After extensive study on the small-world and scale-free properties of networks, the research focus is shifting to detailed local structures. Empirical analysis shows that many real networks exhibit the power-law clique-degree distribution. This general regularity cannot be produced by the rich-get-richer mechanism. In this paper, we propose a common-neighborhood-driven model in which the observed power-law clique-degree distribution can be well reproduced, indicating that the common-neighborhood-driven mechanism is an essential factor leading to the emergence of local structures.

**Keywords:** complex networks, evolving model, clique-degree distribution, common neighborhood

**PACS:** 89.20.Hh, 89.75.Fb

\* Project supported by the Major Research Plan of the National Natural Science Foundation of China (Grant No. 90924011), the Key Program of the National Natural Science Foundation of China (Grant No. 10635040), and the National Natural Science Foundation of China (Grant No. 60973069).

† Corresponding author. E – mail: zhutou@ustc.edu