

Internet 网络的关联性研究^{*}

张国强^{1,2+}, 张国清¹

¹(中国科学院 计算技术研究所,北京 100080)

²(中国科学院 研究生院,北京 100049)

Research on Internet Correlation

ZHANG Guo-Qiang^{1,2+}, ZHANG Guo-Qing¹

¹(Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100080, China)

²(Graduate School, The Chinese Academy of Sciences, Beijing 100049, China)

+ Corresponding author: Phn: +86-10-62565533 ext 8837, E-mail: guoqiang@ict.ac.cn

Zhang GQ, Zhang GQ. Research on Internet correlation. *Journal of Software*, 2006,17(3):490-497.
<http://www.jos.org.cn/1000-9825/17/490.htm>

Abstract: Network correlation plays a key role in the research of Internet topology. Current researches focus on the clustering, mixing and rich-club characteristics separately. This paper makes a deep study on the three network correlation characteristics. First, it points out the possible inconsistency between mean clustering and clustering coefficient, which are two main metrics for measuring the clustering characteristic. Then it shows that the local clustering coefficient is highly correlated with the nodes' degree. After a deep study of the PFP (positive-feedback preference) model, the intrinsic mechanism governing the rise of rich-club phenomenon is discovered and verified. The work is extended by exploring the relationships between these network correlation characteristics.

Key words: autonomous system; network topology; network correlation; clustering; rich-club phenomenon

摘要: 网络的关联性在 Internet 网络拓扑的研究中具有重要作用。目前的研究分别集中于聚集特性、mixing 特性和 rich-club 现象。深入研究了这 3 种网络关联特征:在指出刻画网络聚集特性的两个衡量参数——平均聚集系数与聚集系数可能存在不一致性的同时,发现 AS(autonomous system)网络的局部聚集系数和节点度高度相关;揭示并验证了 PFP(positive-feedback preference)模型中 rich-club 现象的内在形成机制。在此基础上,对这些网络关联特征之间的关联关系进行了研究。

关键词: 自治系统;网络拓扑;网络关联性;聚集特性;富人俱乐部现象

中图法分类号: TP393 文献标识码: A

“结构决定功能”。在 Internet 中,网络的结构决定了网络的路由特性^[1,2]和安全特性^[3]。幂律^[4-7]的发现让人们对这个错综复杂的人造网络有了新的认识,也使原来用于描述网络的随机模型^[8]或层次模型^[9,10]变得不合时宜。在随机模型中,网络的节点是固定的,节点的度分布符合二项分布,在满足 $Np(N$ 是节点个数, p 是连接概率)趋向

* Supported by the Next Generation Internet Collaborative Project between China and Japan (IPv6-CJ) under Grant No.20032050 (下一代互联网中日 IPv6 合作项目)

Received 2005-06-13; Accepted 2005-10-11

定值的情形下,近似服从泊松分布.在该分布下,高度节点存在的可能性以指数速度衰减.因此,可以基本忽略高度节点的存在,而将网络近似看作一种均匀的状态.但是,幂律的存在极大地提高了高度节点存在的可能性,许多实际网络的幂律指数 r 满足 $2 < r < 3$.在这种情况下,网络节点度不存在有穷的方差,这一特征导致了幂律网络的很多特性和随机网络相悖.例如,高度节点的存在极大地削弱了网络的鲁棒性,一个恶意的攻击者只需选择攻击网络很少的一部分节点,就能使网络迅速瘫痪.

网络的节点度分布只是刻画网络拓扑特性的冰山一角.通过对网络拓扑结构的深入研究^[11-17],表明仅仅用节点度分布来区分网络是远远不够的,满足同样幂律分布的网络可以呈现截然不同的结构.除去度分布以外,实际的网络往往呈现很强的关联性,不同的实际网络具有不同的关联特性,而经典的随机网络则不具备任何关联性.为了更细粒度地对纷繁复杂的网络进行区分,人们分别研究了不同的网络关联特征,主要包括聚集 (clustering) 特性^[11]、mixing 特性^[12]和 rich-club 特性^[13,18].聚集系数用于刻画一个节点邻居之间的亲疏程度,实际的网络通常具有很高的聚集系数,而许多网络拓扑模型并不能模拟这一点.Newman^[12]研究了节点之间的连接倾向性,引入了 assortative mixing 和 dissortative mixing 的概念.他发现,社会关系网络通常具有 assortative mixing 特性,即高度节点倾向于和其他高度节点连接;而信息网络和神经网络则呈现 dissortative mixing 特性,即高度节点倾向于和低度节点连接.Zhou^[13,18]对 Internet 的 AS 层拓扑做了详细研究发现,除了 dissortative mixing 特性之外,Internet 拓扑中还存在 rich-club 的现象,即 Internet 高度节点之间存在很强的集团性.

网络拓扑结构关联性的研究面临两个挑战:一方面对于各个关联特性本身的研究尚不成熟,尤其是这些关联特性的形成机理;另一方面,上述所有这些研究都是从某个方面独立地研究网络的关联特性,但对于这些特性之间的关系,尚未有深入的探讨.

然而,网络的关联性是研究一个复杂网络的关键,清楚而透彻地了解一个网络的关联性,能使我们把网络当成是一个有机的整体去研究而不是独立地从某个方面去研究.本文全面地研究了网络的各种关联特征和它们之间的关联关系,主要贡献在于:

- (1) 指出了两种用于刻画聚集特性的参数——平均聚集系数(mean clustering)和聚集系数(clustering coefficient),它们之间可能存在不一致性,建议采用聚集系数(clustering coefficient)来刻画全网络的关联特征;
- (2) 通过对实际的 AS 网络拓扑和不同网络模型的模拟结果进行比较,发现 AS 网络的局部聚集系数 $C(k)$ 和节点度 k 高度相关,而网络模型中仅 PFP(positive-feedback preference)模型能模拟这一特征;
- (3) 提出了一种新方法直观地刻画网络的 mixing 特性;
- (4) 发现并验证了 rich-club 现象产生的内在机制,即新增内部边的两个端点存在特定的关联关系(这里同时采用优先附着函数进行选择)而导致了 rich-club 现象的产生;
- (5) 从整体的角度去研究 clustering, mixing 和 rich-club 这 3 个特征之间存在的关联性.通过解释 AS 网络中低度节点为何具有很高的局部聚集系数,指出了这种关联关系的存在性,为未来的研究指明了方向.

本文第 1 节介绍网络关联性的概念.第 2 节阐述 AS 网络的局部聚集系数和节点度 k 的相关性.第 3 节提出一种用于直观刻画各种不同网络的 mixing 特征的新方法.第 4 节揭示 rich-club 形成的内在机制.第 5 节以一个有机的整体去研究 3 种关联特征的关联关系.最后进行总结,并对未来的研究作出展望.

1 网络的关联性

网络的关联性主要研究的是网络中节点与节点之间的连接概率和节点度的关系.在无关联的网络中,如 E-R 模型,节点和节点之间的连接概率与这两个节点本身的度无关,而在关联网络中则不然.

Dorogovtsev^[11]探讨了关联网络中的聚集系数的计算,利用节点联合度分布 $P(k_1, k_2)$ 可以完整地刻画一个网络的关联特性.这里, $P(k_1, k_2)$ 表示一条边的两个端点度数分别为 k_1 和 k_2 的概率.在无关联的网络中,显然 $P(k_1, k_2) = P(k_1) \cdot P(k_2)$.但节点联合度分布对于刻画网络的拓扑特性不具备很强的直观性,我们需要更加直观地度量参数来区分不同的网络,于是聚集系数、mixing 和 rich-club 等衡量参数应运而生.下面,分别介绍聚集特性、mixing 特征和 rich-club 现象的概念.

1.1 聚集特性

聚集特性是最早用于研究网络关联性的,它用于刻画一个节点邻居之间的亲疏程度.Dorogovtsev 指出,有 3 种不同的衡量聚集特性的参数,分别为:

(1) 局部聚集系数 $C(k)$

$$C(k) = \langle m_{nn}(k) \rangle / (k(k-1)/2).$$

其中, $\langle m_{nn}(k) \rangle$ 表示节点度为 k 的节点的邻居之间平均存在的连接数.

(2) 平均聚集系数(mean clustering)

$$\bar{C} = \sum_k P(k)C(k).$$

其中, $P(k)$ 表示图中任意一个节点的度数为 k 的概率,即节点的度分布.

(3) 聚集系数(clustering coefficient)

$$C = \frac{\sum_k P(k) \langle m_{nn}(k) \rangle}{\sum_k P(k)k(k-1)/2} = \frac{\sum_k k(k-1)P(k)C(k)}{\langle k^2 \rangle - \bar{k}}.$$

其中, $\langle k^2 \rangle$ 表示节点度的二阶矩,而 \bar{k} 表示平均节点度数.

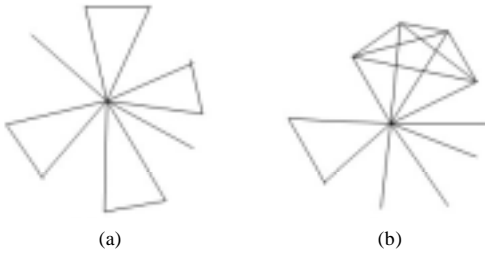


Fig.1 Inconsistency between \bar{C} and C
图 1 \bar{C} 和 C 的不一致性

目前,在研究领域通常使用平均聚集系数来衡量聚集特性,但是我们发现这一刻画参数并不总能正确地反映网络的真实情况.聚集系数(clustering coefficient)是最准确地从全局来刻画网络的聚集特性的参数.在某些情况下,平均聚集系数和聚集系数可能存在不一致性,即存在两个 a 和 b ,满足 $\bar{C}_a > \bar{C}_b$, 但 $C_a < C_b$. 图 1 给出了这样一个例子(其中, $\bar{C}_a = 364/459 = 0.793 > \bar{C}_b = 277/495 = 0.560$ 但 $C_a = 12/53 = 0.226 < C_b = 33/71 = 0.465$). 鉴于此,我们建议使用聚集系数而不是平均聚集系数来刻画网络的聚集特性.

1.2 网络的mixing特性

Newman^[12]指出,在某些实际网络中,高度节点倾向于与其他高度节点连接;而在另外一些网络中,高度节点则倾向于和低度节点连接,但很多网络模型并不能模拟出这一特性.研究表明:众多的社会关系网络,如论文合作网络和演员合作网络,呈现出 assortative mixing 特性;信息网络和神经网络则显现 dissortative mixing 特性;而常用的 E-R 模型和 BA^[5]模型则没有任何倾向性.

1.3 Rich-Club现象

Zhou^[18]在研究 Internet 的 AS 拓扑过程中发现,虽然 AS 拓扑具有 dissortative mixing 特性,但高度节点之间仍然具有很高的连接概率.他把这个现象称为 rich-club 现象,引入了 rich-club coefficient 来对 rich-club 现象进行量化描述.在此基础上提出了目前最准确的 Internet 模型——PFP 模型.显然,rich-club 现象的存在,极大地缩短了平均路径的长度.

2 局部聚集系数和节点度的关联

用平均聚集系数并不能完整地刻画网络的聚集特性,就像用平均度数并不能刻画网络的度分布特性一样.我们采用局部聚集系数 $C(k)$ 来对网络的聚集特性进行衡量.结果表明,对于 Internet 的网络拓扑, $C(k)$ 与节点度 k 高度相关,并随着 k 的增加而递减. BA 模型和改进的 ESF^[7]模型并不能体现这个特征;而 PFP 模型能够较好地刻画这一特征.图 2 给出了 AS 拓扑和 3 种不同的网络模型中 $C(k)$ 和 k 的关系(其中,每个图的纵坐标 $C(k)$ 表示节点度为 k 的节点的局部聚集系数; $\log(k)$ 表示节点度取 10 为底的对数; degree 表示节点度数).

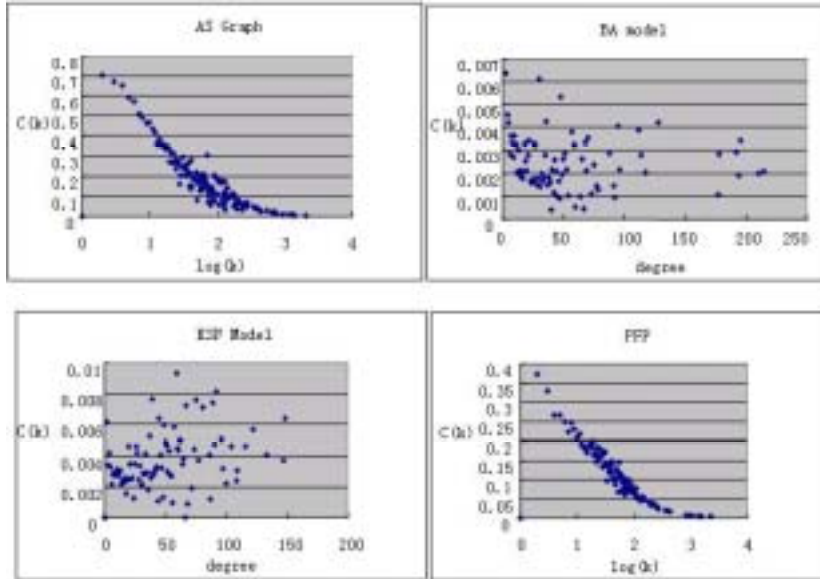


Fig.2 Correlation between $C(k)$ and k for AS graph and three other graphs generated by different network modes
图 2 AS 网络和 3 种不同的网络模型生成的拓扑图中 $C(k)$ 和 k 的关系

数据源描述.对于 AS 拓扑,我们使用的是由 CAIDA^[19]组织提供的实际数据,该数据是由 2004 年 3 月的数据综合生成的.其他 3 个模型均通过我们自己编写的程序来生成相应的网络拓扑图,每种模型都是生成 10 次取平均值所得.表 1 给出了这几个模型的主要度量参数以及与同等规模随机图的比较.

Table 1 Comparison of major metrics between AS graph and 4 network models
表 1 AS 图和用于测试的其他 4 种网络拓扑模型的主要度量对比

Networks	AS graph	BA model	ESF model	PFP model	Random graph
Number of nodes	9 204	10 005	10 099	10 005	10 000
Number of edges	28 959	30 005	34 317	25 699	30 000
Average degree	6.29	6.0	6.8	5.14	6.0
Clustering coefficient	0.025 8	0.002 6	0.003 3	0.011	0.000 6
Mean clustering	0.457	0.004 8	0.003 5	0.229	0.000 6

AS 拓扑中的局部聚集系数 $C(k)$ 与节点度 k 的关联性表明,实际网络中节点度和局部聚集系数是具有高度关联性的.而在无关联的网络中,对于任何 k ,都有 $C(k) = \bar{C} = C = \frac{\langle k^2 \rangle - \bar{k}^2}{N\bar{k}}$,与 k 无关.这个特征可以作为评价 Internet 网络模型好坏的标准之一.

3 Internet 的 mixing 特征

在上节中我们发现,实际的 AS 网络中低度节点具有很高的聚集系数($C(2) > 0.7$).为了解释这一现象,我们首先从另一个角度来考察网络的 mixing 特征.Alexei Vázquez^[20]以邻居的平均节点度来衡量一个网络的 mixing 特征,而我们将测量 k 度节点的邻居节点的度分布,尤其是低度节点的连接偏好特性,以此来解释为何 AS 网络中低度节点具有高聚集系数.特别地,对于 AS 网络,我们作了节点度为 2 的节点的邻居度分布图,如图 3 所示(其中,横坐标为节点度数取 10 为底的对数,纵坐标为相应节点度的概率).

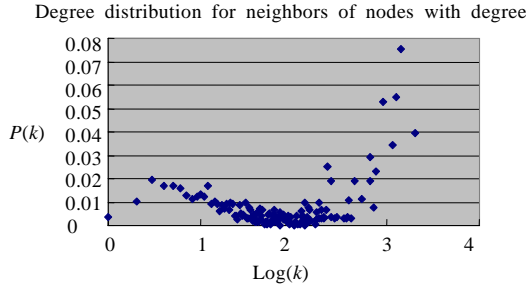


Fig.3 Neighbors' degree distribution of nodes with degree two in AS graph
图 3 AS 网络中节点度为 2 的邻居度分布图

同时,我们对该分布做了处理,将节点度按降序排列,取前 10% 的高度节点,测量了节点度为 2 的节点的邻居节点中高度节点的比例.由图 4 可以看出,节点度为 2 的邻居节点中有超过 50% 的节点为度数排名为前 2% 的高度节点.

无关联网络中沿着任意一条边到达的端点度为 k 的概率正比于 $kP(k)$.对于无关联的 power-law 网络来说, $P(k)=ck^{-r}$,通常 $2 < r < 3$.这样,沿着任意一条边到达的端点为 k 的概率正比于 ck^{1-r} .显然,对于任意 $k_1 < k_2$, $ck_1^{1-r} > ck_2^{1-r}$.因此,对于不呈现任何 mixing 特性的 power-law 网络来说,节点度为 k 的邻居节点中依然是低度节点占大部分,BA 模型就是这样的一个例子.图 5 给出了 BA 模型中节点度为 3 的邻居节点的度作为按度降序排列百分比的累积分布.

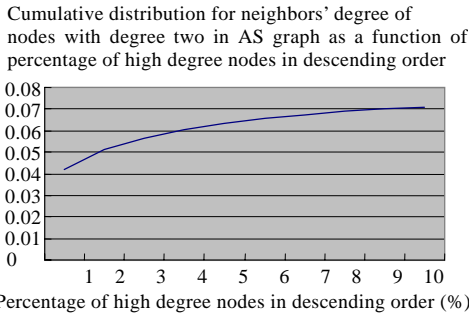


Fig.4 Cumulative distribution for neighbors' degree of nodes with degree two in AS graph as a function of percentage of high degree nodes in descending order

图 4 AS 网络拓扑中节点度为 2 的节点的邻居节点的度数作为高度节点降序排列百分比的累积分布图

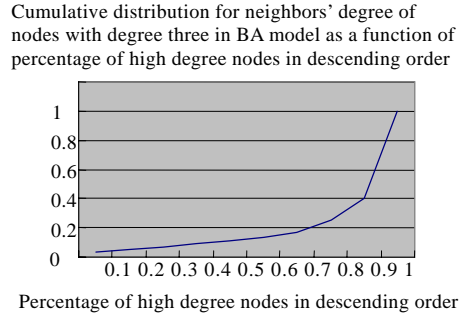


Fig.5 Cumulative distribution for neighbors' degree of nodes with degree three in BA model as a function of percentage of high degree nodes in descending order

图 5 BA 模型中节点度为 3 的节点的邻居节点的度数作为高度节点降序排列百分比的累积分布图

对比图 4 和图 5,我们可以很清楚地看出实际的 AS 网络和 BA 模型在 mixing 特性上的不同.这是由于上述概率没有考虑边的目的节点和源节点的关联性.而在很多实际网络中,随机选取边的另一端的度数与源节点的度数相关,如在 AS 拓扑中,节点度为 2 的端点引出的两条边的另一端的度数显然不正比于 ck^{1-r} ,而是高度倾向于高度节点.从这一点来看,网络也呈现出 assortative mixing 特性,即低度节点倾向于与高度节点连接.

4 Rich-Club 的形成机制

Zhou 在研究 AS 拓扑中发现了 AS 拓扑也存在 rich-club 现象,即高度节点之间存在很强的集团性,它的存在能够很好地解释 Internet 路由的高效性.他提出的 PFP 模型也由于能够模拟 rich-club 现象而成为目前最准确的 Internet 网络拓扑模型.PFP 模型的两个基石是:

(1) 正反馈的优先附着函数

$$P_k = \frac{d_k^{1+\delta \log_{10} d_k}}{\sum_i d_i^{1+\delta \log_{10} d_i}} \quad (1)$$

(2) Interactive growth 过程

网络的增长不仅仅是节点的增加,也是一种新节点和新的内部边互动增长的过程.具体的,增长步骤如下:

- 1) 以概率 p 增加一个新节点 N ,选择一个已有节点 H 与之连接,同时增加一条以 H 为端点的内部边;
- 2) 以概率 $q(q < 1-p)$ 增加一个新节点 N ,选择一个已有节点 H 与之连接,同时增加两条从 H 出发的内部边;
- 3) 以概率 $1-p-q$ 增加一个新的节点,选择两个已有节点 H_1, H_2 与之连接,同时增加一条以 H_1 或 H_2 为端点的内部边.

整个过程对于边的另一端的点的选择都遵从式(1)的正反馈优先附着函数.

在这两个基本元素中,Interactive growth 过程是导致 rich-club 和其他许多网络关联性的主要原因.但为何 PFP 模型生成的网络具有 rich-club 现象,而同样具备内部边增长机制的 ESF 模型或其他一些模型却不呈现 rich-club 现象?我们对 Interactive growth 过程作了仔细的分析,发现 Interactive growth 在增加内部边时的选择机制有别于其他内部边增长机制.正是这种带有关联性的新增内部边选择方式导致了 rich-club 的产生.在 ESF 模型中,增加内部边的选择过程如下:

- 1) 随机选择一个节点作为边的一个端点;
- 2) 使用优先附着函数选择另一个端点.

而 Interactive growth 过程却保证了新增内部边的两个端点都使用优先附着函数来选择,因此增加的内部边的两个端点具有关联性.这样,在当前网络状态下,增加边的两个端点是 k_1, k_2 的概率 $P(k_1, k_2)$ 为

$$P(k_1, k_2) = \frac{d_{k_1}^{1+\delta \log_{10} d_{k_1}} d_{k_2}^{1+\delta \log_{10} d_{k_2}}}{\left(\sum_i d_i^{1+\delta \log_{10} d_i}\right)^2} \quad (2)$$

上述机制使得新增加的内部边的两个端点有很高的概率都是高度节点,随着网络的演化,高度节点之间必然形成 rich-club.而 ESF 模型中新增加的内部边端点之间不具备关联性,因此不会产生 rich-club 现象.即新增内部边端点选择的关联性导致了 rich-club 的产生.为了验证此观点,我们提出了一种新的网络生成模型,该模型摒弃了 Interactive growth 过程,但在增加内部边时采用两个端点都使用优先附着函数进行选择.结果表明,通过该网络生成模型生成的网络同样也具有 rich-club 现象.

我们定义的网络生成模型每一步的增长方式如下:

- (1) 以概率 p 增加一个新的节点,并从该节点连一条边到已有网络,依然使用式(1)进行优先附着选择;
- (2) 以概率 q 增加一个新节点,并从该节点连两条边到已有网络;
- (3) 以概率 $1-p-q$ 增加一条内部边,内部边的两个端点都使用优先附着函数进行选择.

通过多次模拟,我们选择 $p=0.32, q=0.15$.在该参数设置下,我们生成了 10 个网络,并对其进行平均,每一个网络我们都执行 25 000 步.模拟所得的网络平均大小为 11 741,平均节点度数为 5.62,1%的最高度节点的 rich-club coefficient 为 0.314,assortative mixing coefficient 为 -0.210;而大小为 9 204 的 AS 网络拓扑,平均节点度数为 6.29,1%的最高节点的 rich-club coefficient 为 0.436,assortative mixing coefficient 为 -0.236.

由此可见,新定义模型同样能模拟 Internet 的 rich-club 现象和 dissortative mixing 特性.因此验证了我们的结论,即 rich-club 现象形成的内在机制是由于新增内部边的两个端点都采用优先附着函数进行选择,从而使新增内部边的端点之间存在了关联性.PFP 模型的 Interactive growth 过程由于自动保证了这一内在机制,而使得生成的网络呈现了 rich-club 现象.

5 为何低度节点具有很高的局部聚集系数

前面几节就 3 种网络关联特征分别进行了研究.在此基础上,我们将从一个整体的角度去研究这 3 种网络关联特征的关系.通过研究为何 AS 网络中低度节点具有很高的局部聚集系数,我们指出了聚集特性、mixing 特性和 rich-club 特性之间关联关系的存在性,为进一步研究网络的关联性指明了方向.

综合考虑上述 3 个网络关联特征,就能解释为何 AS 网络中低度节点具有很高的局部聚集系数.这主要是由于以下原因形成的:

- (1) 由于在第 3 节讨论的 AS 网络具有 assortative mixing 特征,使得低度节点更倾向于和高度节点连接.正如第 3 节所讨论的,在节点度为 2 的节点的邻居中,有超过 50% 的节点属于 2% 的高度节点集合;
- (2) 由于高度节点呈现 Rich-club 现象,而低度节点的邻居大部分都是高度节点,因此,低度节点的局部聚集系数必然比较大.

上述原因表明,assortative mixing 和 rich-club 特性共同导致了 AS 网络中低度节点具有高局部聚集系数的现象.进一步地,一个具有 assortative mixing 特征的网络如果存在 rich-club,那么其三角形数目相比于不存在 rich-club 的网络必然剧增,因此也必定呈现更高的聚集系数(clustering coefficient).

这个因果关系也告诉我们,不同的网络关联特征并不是独立存在的,而是内在地具有某种关联关系,它们相互作用而形成一有机体;将某个关联特征割裂开来去研究是不全面也不系统的.要完整地了解一下复杂系统的形成机理并解释某种现象,我们需要同时考虑不同关联特征的相互联系和相互作用.

6 结论和展望

网络的关联特性对于网络内在结构的研究具有重要意义,它有助于我们更深入地理解一个看似纷繁复杂的系统其内在的机理.本文首先指出,由于聚集系数(clustering coefficient)和平均聚集系数(mean clustering)之间可能存在的 inconsistency,我们应该使用前者来正确地刻画全局网络的聚集特征.在此基础上,本文采用了局部聚集系数来衡量一个网络的聚集特征,指出 Internet 的 AS 层网络拓扑的局部聚集系数和节点度紧密相关.BA 模型和 ESF 模型都无法模拟这一特性,而成功模拟了 rich-club 现象的 PFP 模型则能较好地反映该关联性.本文进而从另一个角度考察了 Internet 的 assortative mixing 特性,利用 k 度节点的邻居节点的度分布能够全面而直观地刻画 Internet 的 assortative mixing 特性.AS 网络的低度节点的邻居节点中高度节点占很高的比例,这与 BA 模型正好相反.本文还进一步深入研究了 PFP 模型,发现了 rich-club 现象形成的内在机制,即新增内部边两个端点带有关联性的选择方式,是导致网络是否具备 rich-club 现象的原因.如果新增内部边的两个端点都采用优先附着函数进行选择,则网络就会具备 rich-club 现象,否则就不会具备.我们通过提出一个新的简洁的网络生成模型验证了这一观点.

Assortative mixing 特性和 rich-club 现象,正是导致 AS 网络局部聚集系数和节点度具有高度相关性的原因,特别是低度节点具有很高的聚集系数,而这正是 3 个关联特征之间具有关联关系的写照.

对于网络关联性的研究尚未到达很深入的阶段,目前的研究依然处于独立研究某个网络关联特征的阶段.本文首次指出了这些关联特征之间存在关联性,应该系统地去研究这些关联特征及其相互联系和相互作用,对于这些特征的关联关系研究能使我们全面地了解网络的演化机理,而将某一个网络关联特征割裂开来的研究是不全面的.另外,目前的研究仍然集中于静态的网络结构的关联性,但对于 Internet 来说,网络的结构和流量、性能等动态量之间的关联关系同样具有重要的价值,这将是网络关联性研究的下一个热点之一.

References:

- [1] Krioukov D, Fall K, Yang X. Compact routing on Internet-Like graphs. In: Proc. of the IEEE INFOCOM 2004. Hong Kong: IEEE, 2004. 209-219
- [2] NSF 04-540: Toward mathematically rigorous next-generation routing protocols for realistic network topologies. 2005. <http://www.caida.org/projects/nets-nr/>

- [3] Albert R, Barabá AL. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 2002,74(1):47–97.
- [4] Faloutsos M, Faloutsos P, Faloutsos C. On power-law relationships of the Internet topology. *ACM SIGCOMM Computer Communication Review*, 1999,29(4):251–262.
- [5] Barabási AL, Albert R. Emergence of Scaling in random networks. *Science*, 1999,286(5439):509–512.
- [6] Medina A, Matta I, Byers J. On the origin of power laws in Internet topologies. *ACM SIGCOMM Computer Communication Review*, 2000,30(2):18–28.
- [7] Barabási AL, Albert R. Topology of evolving networks: Local events and universality. *Physical Review Letter*, 2000,85(24):5234–5237.
- [8] Bela Bollobas. *Random Graphs*, 2nd ed, Cambridge, 2003.
- [9] Doar MB. A better model for generating test networks. In: *Proc. of the GLOBECOM'96*. London: IEEE, 1996. 86–93.
- [10] Zegura EW, Calvert KL, Donahoo MJ. A quantitative comparison of graph-based models for Internet topology. *IEEE/ACM Trans. on Networking*, 1997,5(6):770–783.
- [11] Dorogovtsev SN. Clustering of correlated networks. *Physical Review E*, 2004,vol.69, no.027104.
- [12] Newman MEJ. Assortative mixing in networks. *Physical Review Letter*, 2002,vol.89, no.208701.
- [13] Zhou S, Mondragon RJ. Accurately modeling the Internet topology. *Physical Review E*, 2004,vol 70, no.066108.
- [14] Mahadevan P, Krioukov D, Huffaker B, Dimitropoulos X. Comparative analysis of the Internet AS-level topologies extracted from different data sources. 2004. http://www.caida.org/analysis/topology/as_topo_comparisons/
- [15] Tangmunarunkit H, Govindan R, Jamin S. Network topology generators: Degree-Based vs. structural. *ACM SIGCOMM Computer Communication Review*, 2002,31(4):147–159.
- [16] Tian B, Towsley D. On distinguishing between Internet power law topology generators. In: *Proc. of the IEEE INFOCOM 2002*. Vol 2. New York: IEEE, 2002. 638–647.
- [17] Zhang Y, Zhang HL, Fang BX. A survey on Internet topology modeling. *Journal of Software*. 2004,15(8):1220–1226 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/15/1220.htm>
- [18] Zhou S, Mondragon RJ. The rich-club phenomenon in the Internet topology. *IEEE Communications Letters*, 2004,8(3):180–182.
- [19] 2005. <http://www.caida.org>
- [20] Vázquez AV. Degree correlations and clustering hierarchy in networks: Measures, origin and consequences [Ph.D. Thesis]. *Scuola Internazionale Superiore di Studi Avanzati Int'l School for Advanced Studies*, 2002.

附中文参考文献:

- [17] 张宇,张宏莉,方滨兴. Internet 拓扑建模综述. *软件学报*. 2004,15(8):1220–1226. <http://www.jos.org.cn/1000-9825/15/1220.htm>



张国强(1980 -),男,江苏常州人,博士生,主要研究领域为计算机网络.



张国清(1965 -),男,副研究员,CCF 高级会员,主要研究领域为计算机网络.