

· 复杂性科学 ·

## 无权网络零模型的构造及应用

许小可<sup>1,2\*</sup>, 崔文阔<sup>1</sup>, 崔丽艳<sup>1</sup>, 肖婧<sup>1</sup>, 尚可可<sup>3</sup>

(1. 大连民族大学信息与通信工程学院 辽宁 大连 116600; 2. 贵州省公共大数据重点实验室(贵州大学) 贵阳 550025;  
3. 南京大学新闻传播学院计算机传播实验中心 南京 210093)

**【摘要】**静态无权网络是目前最常见的复杂网络形式,这种网络零模型也被研究得最广泛和最深入。该文将无权网络分成无权无向网络和无权有向网络两种形式,分别研究了这两类网络的零模型构造及应用,其中重点是无权无向网络。首先根据不同阶数随机图理论阐述了无权无向网络由低到高各阶零模型的定义,然后描述了使用ER随机图、配置模型和基于断边重连等方式构造各阶零模型的过程及相关应用。针对断边重连这种最重要的零模型构造方式,论述了无倾向性断边重连、有倾向性同配或异配断边重连,以及检测网络是否具有富人俱乐部性质的局部断边重连等构造方式,并且首次将高阶零模型扩展到社团检测等网络中尺度特性的分析中。最后,阐述了无权有向网络1阶零模型的构造以及如何基于该零模型检测网络中存在的出入度匹配特性。该文发现网络零模型能为实证无权网络提供一个准确的基准,结合网络的统计量指标定性和定量地描述出实际复杂网络的非平凡特性以及这种非平凡特性的程度及来源。

**关键词** 配置模型; 零模型; 随机断边重连; 无权网络

中图分类号 TP391 文献标志码 A doi:10.3969/j.issn.1001-0548.2019.01.020

## Construction and Applications of Null Models for Unweighted Networks

XU Xiao-ke<sup>1,2\*</sup>, CUI Wen-kuo<sup>1</sup>, CUI Li-yan<sup>1</sup>, XIAO Jing<sup>1</sup>, and SHANG Ke-ke<sup>3</sup>

(1. College of Information and Communication Engineering, Dalian Minzu University Dalian Liaoning 116600;  
2. Guizhou Provincial Key Laboratory of Public Big Data (Guizhou University) Guiyang 550025;  
3. Computational Communication Collaboratory School of Journalism and Communication, Nanjing University Nanjing 210093)

**Abstract** The static unweighted network, whose null models have been studied widely and thoroughly, is the most common type of complex networks at present. In this study, we classify an unweighted network into two types of networks: undirected network and directed network. We study the construction and applications of null models for the two kinds of networks, especially the null models of unweighted undirected network are emphasized. Firstly, we illustrate the definitions of null models from low to high orders for unweighted undirected network according to the theory of random graph series. And then we describe the constructing process and related applications for  $1 \leq k \leq 3$  null models by using ER random graph, configuration model, edge swapping, and so on. For the edge swapping algorithm, which is the most important mode for constructing null models, we introduce non-tendentious random edge swapping, tendentious assortative or dis-assortative edge swapping, and local edge swapping for detecting whether the rich-club properties exist in a network. Moreover, the high order null models are firstly extended to analyze meso-scale network features such as community detection. Finally, we analyze  $1 \leq k$  null models of directed network and tried to detect four types of in-out degree assortativities. In this study, we find that null models can not only provide an accurate baseline for real-life networks, but also qualitatively and quantitatively describe non-trivial properties of empirical complex networks.

**Key words** configuration model; null model; random disconnect-rewiring; unweighted network

复杂网络是指由数量巨大的节点在一定时间或空间尺度上相互关联共同构成的复杂系统,在研究

中重点关注复杂网络是否存在非平凡性质以及该性质的产生机理。在实证网络研究中,一般会使用多

收稿日期: 2017-08-12; 修回日期: 2017-12-08

基金项目: 国家自然科学基金(61773091, 61603073, 61374170); 辽宁省高等学校创新人才支持计划(DC201502060201); 辽宁省重点研发计划项目(2018104016)

作者简介: 许小可(1979-), 男, 教授, 主要从事复杂网络、社交网络大数据及计算传播学等方面的研究. E-mail: xuxiaoke@foxmail.com

种统计量来刻画复杂网络的性质, 如描述静态无权网络的统计量有平均度、度分布、匹配系数、聚类系数和平均路径长度等<sup>[1-2]</sup>。

鉴于各种网络统计量的绝对数值往往是无量纲的, 而且不同复杂网络的规模大小各异、结构千差万别<sup>[3]</sup>, 因此学者们不仅关注实证网络统计量的绝对值, 更加关心真实网络和网络随机化零模型比较后的相对值<sup>[4-7]</sup>, 通过统计量的相对结果来说明网络的性质, 此时如何构造出一个合适的零模型作为实证网络的参照物就变得非常重要。通常把与一个实证网络具有某些相同性质的随机网络称为该实证网络的随机化副本, 这类随机化网络在统计学上被称为零模型<sup>[8-12]</sup>。

基于网络零模型的研究价值主要有以下 3 点: 发现、度量和寻因。具体来说, 一是可发现网络的非平凡性质, 二是可量化这种特殊性质的程度, 三是可对这些性质产生的原因进行探究。若实证网络和最简单的零模型在性质上没有任何差别, 那么说明该网络不具有非平凡性质。若实证网络和其对应零模型在性质上具有统计意义上的差别, 则说明该网络具有非平凡性质, 并且对于不同网络可以通过比较同一个量化指标来说明它们拥有这种性质的程度。参照零模型的假设检验可遵循从简单到复杂、从低级到高级的顺序提高零模型和实证网络的吻合程度, 最终发现实证网络的性质是由哪些关键因素决定的。

文献[9]较系统性地介绍了无权网络、加权网络和时效网络 3 种类型网络中零模型的构造与应用, 但对无权网络这种最简单、最常用的网络, 零模型研究和阐述不够深入与完善, 因此本文主要围绕无权网络的零模型展开, 删减了加权网络和时效网络部分。在原文的基础上有如下的延伸和改进。文献[9]仅对 1 阶配置模型进行了阐述, 而本文不仅重新叙述了 1 阶配置模型的构造过程, 也论述了如何构造 2 阶和 3 阶配置模型。对于断边重连方式构造零模型, 文献[9]仅介绍了 0 到 2 阶零模型, 而本文不仅增加了 2.5 和 3 阶零模型的构造过程, 并且对 0~3 阶每一种零模型都增加了新的典型应用案例。最后, 本文针对网络零模型领域的最新研究热点, 提出了针对社团结构的多种零模型, 这些零模型对于分析网络微观和中观尺度之间的联系具有重要的意义与作用。

### 1 网络模型研究的局限性

目前很多复杂网络实证研究不使用网络零模

型, 其流程如图 1a 所示, 首先计算实证网络的某个统计量, 然后通过该统计特征的绝对数值对网络性质进行描述。这种方式的结果往往是不准确的, 不同网络之间结果也不具有可比性。

参照零模型的实证网络研究方法如图 1b 所示, 该研究框架不仅要计算实证网络的某个统计量, 也需要计算它多个零模型网络的对应统计量, 然后引入经典的假设检验方法, 通过统计分析来确定实证网络性质是否和零模型网络之间存在显著性差异, 以及差异的显著性水平有多高, 最终确定其性质是由网络结构的哪些因素导致的。

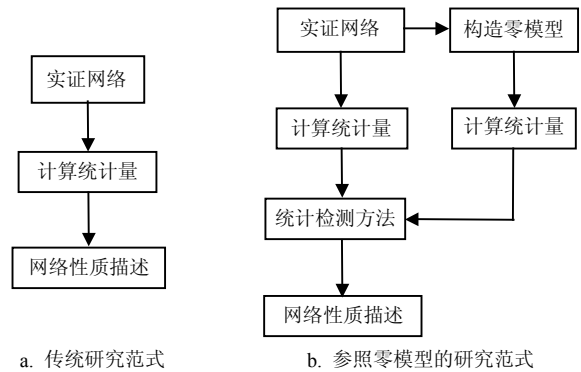


图1 传统网络研究与参照零模型网络研究范式的对比

无权网络包括无权无向网络和无权有向网络两种。无权无向网络是指网络中的连边既没有权重, 也没有方向, 如图 2a 所示。无权有向网络是指网络中的连边没有权重但有方向, 如图 2b 所示。无权有向网络的连边可以是单向的, 也可以是双向的, 通过连边方向来区分节点之间关系的发出方和接受方。

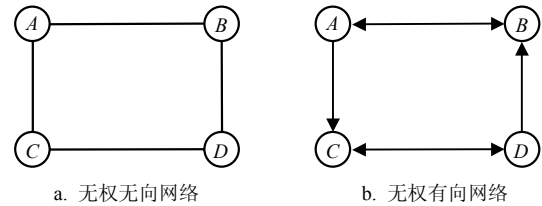


图2 无权网络的两种形式

无权无向网络是目前最常见的复杂网络形式, 这种网络也被研究得最广泛和最深入, 因此目前为其建立了各种各样的理论模型。网络模型的方法一般是首先计算实证网络的一些重要统计量, 然后根据这些统计量的数值来推断网络的形成机制, 最终通过模拟网络的形成机制来生成和原始网络某些统计特征类似的新网络。网络模型方法的优点是可以从无到有生成零模型网络, 而且生成网络的规模等参数可以人为控制, 具有很好的灵活性。这方面最有代表性的工作是: 20世纪60年代Erdős和Rényi提出的随机图(ER网络)、Watts和Strogatz在1998年提出

的“小世界”网络模型(SW网络)<sup>[13]</sup>以及Barabási和Albert在1999年提出的无标度网络模型(BA网络)<sup>[14]</sup>。

尽管网络模型在复杂网络研究领域获得了广泛应用,但是网络模型研究也存在以下两方面的局限性。一是建模方法对真实网络生成机制的推断不一定是正确和准确的,而且使用模型从无到有生成网络模型的时候往往只能保持少量特征和实证网络相似,而在更多的特征上两者还是有很大差异,既能保持形式简洁又能深刻描述复杂系统内在物理机制的网络模型很难构造。二是基于模型方法生成的网络往往对于真实网络某些统计特征的模拟是有偏的,即基于模型方法生成的网络中某个统计量往往都大于(或都小于)实证网络,在使用的时候无法判断出模型的这种有偏特性会对真实网络其他特性有哪些方面的影响。

相对于由统计量直接生成网络的模型方法,无权无向网络的零模型产生方法有两种,一种是基于配置模型的方法。配置模型是在ER随机图的基础上进行扩展产生的一种广义随机图,配置模型生成零模型的基本思想是:根据产生不同阶数零模型所需要的基本特性,将原始网络拆成相应的含有某种特性的基本子图,并通过不同的约束条件将基本子图重新随机组合到一起。配置模型可产生1阶、2阶或更高阶零模型,相对于BA等演化模型更接近于实证网络的各方面性质。

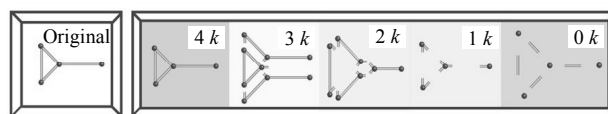
另一种生成零模型的方法是将实证网络进行置乱(断边重连)的方法,实证网络断边重连既可以随机断边重连,也可以使用有倾向性的断边重连。随机断边重连是很多应用中的常用方法,可以产生0~3阶网络零模型。有倾向性断边重连有两方面的应用:一是可检测网络是否具有富人俱乐部特性,二是可检测网络是否具有某种匹配特性<sup>[15]</sup>。

此外,本文构造了4种依赖于社团结构的零模型:社团内部断边重连零模型、社团间断边重连零模型、增强社团结构零模型和减弱社团结构零模型。社团结构是复杂网络重要的中尺度结构特征,可以揭示网络的结构与功能之间的关联关系<sup>[16]</sup>。无权有向网络的零模型只有一种构造方式,即将原始网络随机断边重连,一般使用该方法来产生对应的1阶零模型,适用于检测有向网络中节点之间的出入度匹配特性。

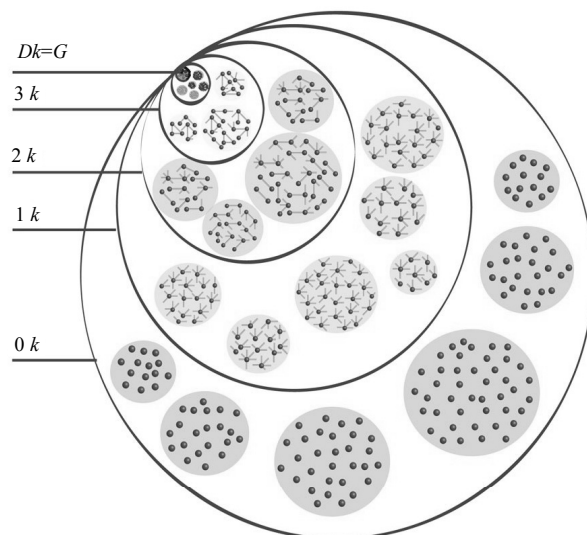
## 2 不同阶数网络零模型的定义

一般而言,使用网络零模型的主要目的是检测

实证网络的非凡特性,这就需要从粗糙到精确逐步逼近原始网络的性质。文献[1,4]为了逐步逼近实证网络,根据原始网络的基本特性,将网络中的基本组成单元分成不同的小模块。如图3a所示,最左侧的图为原始网络,右侧的图从右到左依次介绍了产生不同阶数的零模型所需要原始网络的性质,从而引入了不同阶数网络零模型的概念。同时,如图3b所示,不同阶数网络零模型之间并不是独立的,按照约束条件从少到多,存在一种包含关系,即  $0k \supseteq 1k \supseteq 2k \supseteq \dots \supseteq (n-1)k \supseteq nk$ , 任何一个  $n$  阶零模型都会包含  $n-1$  阶零模型的性质<sup>[3]</sup>。



a. 产生不同阶数的零模型



b. 不同阶数网络零模型关系

图3 不同阶数零模型网络之间的关系图(取自文献[4])

### 2.1 0阶零模型

0阶零模型是最简单也是最随机化的网络零模型,只需要与原始网络具有相同的节点数和平均度。平均度  $\langle k \rangle$  是指原始网络中所有节点的度的平均值,设  $m$  为原始网络连边的数量,  $n$  为原始网络节点的数量,则  $\langle k \rangle = 2m/n$ 。

### 2.2 1阶零模型

1阶零模型主要是保证与原始网络具有相同的节点度分布。1阶零模型与原始网络不但具有相同的节点数、相同的平均度,更重要的是具有相同的节点度分布  $p(k)$ , 度分布是指原始网络中节点度的概率或数目的分布<sup>[17]</sup>。若  $n(k)$  为原始网络中节点度为  $k$  的节点的数目,则度分布为  $p(k) = n(k)/n$ 。

2.3 2阶零模型

2阶零模型与原始网络具有相同的联合度分布  $p(k_1, k_2)$ <sup>[18]</sup>, 联合度分布是指每条边两端连接节点的度值数目(概率)。若  $m(k_1, k_2)$  为度为  $k_1$  和  $k_2$  的节点之间连边的总数, 则

$$p(k_1, k_2) = m(k_1, k_2) \mu(k_1, k_2) / 2m \quad (1)$$

式中, 若  $k_1 = k_2$ , 则  $\mu(k_1, k_2) = 2$ ; 否则,  $\mu(k_1, k_2) = 1$ 。

2.4 2.25阶零模型

2.25 阶零模型是根据原始网络中的联合度分布和平均聚类系数<sup>[19]</sup>这两个物理属性构造的, 要求与原始网络具有相同的联合度分布  $p(k_1, k_2)$  和平均聚类系数  $\bar{C}$ 。假设网络中的一个节点  $i$  有  $m_i$  条边将它和其他节点相连, 这  $m_i$  个节点就称为节点  $i$  的邻居节点。显然, 在这  $m_i$  个节点之间最多可能有  $m_i(m_i - 1)/2$  条边。 $m_i$  个节点之间实际存在的边数  $M_i$  和总的可能边数  $m_i(m_i - 1)/2$  之比定义为节点  $i$  的聚类系数  $C_i$ , 即:

$$C_i = \frac{2M_i}{m_i(m_i - 1)} \quad (2)$$

平均聚类系数  $\bar{C}$  定义为在原始网络中, 所有节点的聚类系数的平均值, 即:

$$\bar{C} = \frac{1}{n} \sum_i C_i \quad (3)$$

2.5 2.5阶零模型

2.5 阶零模型是在 2.25 阶模型的基础上, 对聚类系数增加更强的一点约束条件, 要求与原始网络具有相同的联合度分布  $p(k_1, k_2)$  和度相关的平均聚类系数  $\bar{C}(k)$ , 即:

$$\bar{C}(k) = \frac{1}{|N_k|} \sum_{i \in N_k} C_i \quad (4)$$

式中,  $N_k$  表示度为  $k$  的节点的集合。

2.6 3阶零模型

3 阶零模型与原始网络具有相同的联合边度分布  $p(k_1, k_2, k_3)$ , 联合边度分布考虑到 3 个节点之间的相互连接性, 主要有两种情况: 1) 开三角形如图 4a 所示, 即 3 个节点用 2 条边来连接称为  $p_{\wedge}(k_1, k_2, k_3)$ ; 2) 闭三角形如图 4b 所示, 3 个节点形成一个环称为  $p_{\Delta}(k_1, k_2, k_3)$ 。与原始网络具有相同的联合边度分布  $p(k_1, k_2, k_3)$ , 就要求零模型网络和原始网络具有相同的开三角形和闭三角形分布。

随着生成零模型约束条件的逐步增加, 构造出的高阶零模型比低阶模型更接近于原始网络, 因此高阶的3阶零模型比0~2.5阶零模型更接近于原始

网络。以此类推, 理论上还可以定义更高阶的零模型<sup>[3]</sup>。同时, 随着约束条件的增多, 可以被置乱的连边数也就越少, 生成的网络零模型越接近于原始网络, 但是这样进行连边置乱也就越困难。对于3阶及3阶以上网络零模型, 该类零模型受到的约束条件非常严格, 很多情况下不能在一个有限的计算时间内生成相应的零模型, 不具备实用性, 所以一般认为2.5阶零模型是现在能够实现的较为通用的最高阶零模型<sup>[1-2]</sup>。因此, 本文着重介绍可以将原始网络进行较程度随机化的0~2.5阶网络零模型。

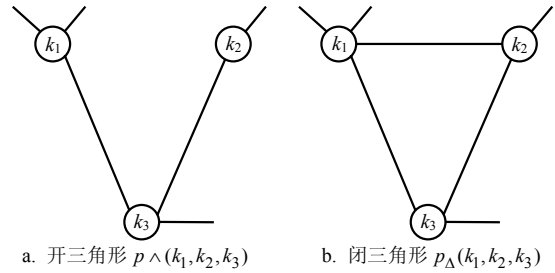


图4 3个节点相互连接的两种可能形式

3 ER随机图与配置模型

3.1 ER随机图

20世纪60年代匈牙利的Erdős和Rényi这两位数学家对随机图理论做出了重要贡献, 以他们名字简写的ER随机图是最为大家熟知的随机网络。ER随机图可以由两种方式构造出来。一种是给定  $n$  个节点和  $m$  条连边, 随机地从  $n$  个节点中选择两个没有相连的节点用连边连接起来, 重复上述过程直到连边数达到  $m$  条连边后停止选择, 构造出没有重边和自环的简单图。

另一种是给定  $n$  个节点, 以概率  $p$  任意连接两个不同的节点, 连通概率  $p$  介于0和1之间, 重复上述过程直到相连边数的概率达到  $p$ <sup>[5]</sup>。尽管ER随机图在很长一段时间内被大家广泛使用, 但该模型是一种完全随机的网络, 该模型中仅平均度与原始网络是相同的, 但是度分布是泊松的、聚类系数接近于零, 这些性质和实际复杂网络中幂律度分布和高聚类特性相差很大<sup>[6]</sup>。

通过将实际网络与它对应的ER随机图进行比较, 可以验证实际网络是否具有小世界特性。如表1所示, Barabási等构造了WWW、电影演员网、电力网这3个网络的ER随机图, 并将这3个随机图中的聚类系数和平均距离这两个统计量与实际网络中相应的统计量进行定量比较, 可以看到在实际网络和ER随机图具有相同的节点数和平均度的情况下, 它们

的平均距离几乎相同。同时,实际网络中的聚类系数远远高于相同规模的ER随机图,因此可以说聚类特性是原始网络所特有的属性而不能通过ER随机图产生,即实证网络具有小的平均路径长度和高的聚类系数,也就是我们常说的“小世界特性”。

表1 实际网络与ER随机图的统计量比较<sup>[7]</sup>

网络	$\langle k \rangle$	$L$	$L_{ER}$	$C$	$C_{ER}$
WWW	35.2	3.1	3.4	0.11	0.000 2
演员网	61.0	3.6	2.9	0.79	0.000 2
电力网	2.70	18.7	12.4	0.08	0.005

此外,也可以采用随机断边重连的方式产生ER随机图,也就是0阶零模型。但是无论采用哪种方式,0阶零模型只和原始网络具有相同的节点数和边数,生成的网络结构随机性很强,性质和原始网络相差较大。

### 3.2 广义随机图与配置模型

配置模型是在ER随机图的基础之上进行扩展产生的一种广义随机图,本节主要介绍基于配置模型构造网络零模型的方法。由配置模型生成零模型的具体思路是:首先根据产生不同阶数零模型所需要的基本特性,将原始网络拆成相应的含有某种特性的基本子图,通过不同约束条件将基本子图重新随机组合到一起。

本节主要介绍基于配置模型生成1阶、2阶、3阶零模型网络的过程。在1阶配置模型中通过给定原始网络的度序列来构造零模型网络,2阶配置模型通过给定原始网络中的联合度分布序列来构造零模型网络,3阶配置模型更具体地给出了每个节点中边角的分布情况,即原始网络中的“线”和“角”序列来构造零模型网络<sup>[3]</sup>。

#### 3.3 1阶配置模型

ER随机图是在保证与原始网络节点数和边数相同的情况下,采用随机连边方式生成的零模型。它仅仅保证了与原始网络具有相同的平均度分布,是随机性最强的一种网络。为了比ER随机图更精确地拟合原始网络,构造出与真实网络度分布相同的零模型,文献[8, 11]提出了广义随机图,即配置模型。

在1阶配置模型中,首先给定原始网络的度序列 $\{k_i\}$ ,在此前提下使用“配置”连边的方式生成了新的零模型网络,具体的构造过程如图5所示。首先将原始网络图5a拆分成A、B、C、D 4个独立的节点,在拆分的过程中保持每个节点的度值不变,从每个

节点中引出与度值数目相同的“线头”,如图5b所示,短线即为引出的线头。然后,每次随机的选择一对线头连接在一起形成一条边,如图5c所示。最终,将所有的线头都连接完成所形成的网络(如图5d所示)即为配置模型的1阶随机网络。1阶配置模型只要给定了真实网络的度序列,就可以生成与真实网络具有相同度序列的随机网络。

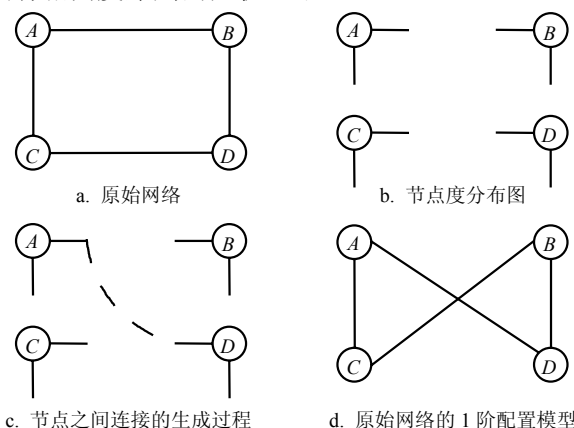


图5 原始网络和使用配置模型构造的1阶零模型网络

#### 3.4 2阶配置模型

使用配置模型也可以生成原始网络的2阶零模型,与原始网络具有相同的联合度分布。假定原始网络有 $m$ 条边,首先将该网络拆成 $m$ 个子图,每个子图为两个节点 $\{i, j\}$ 配对的形式。然后,随机选取一个子图,并将子图中度值为 $k_i$ 的节点 $i$ 视为起始节点。此时,随机在网络中挑选 $k_i - 1$ 个子图,并且每个子图中至少含有一个度值为 $k_i$ 的节点,最后将这 $k_i$ 个节点重新组合成一个新的度为 $k_i$ 的节点。持续随机选取一个分支节点作为起始节点重复上述过程,直到将所有的子图都重新组合就完成了2阶零模型的生成。

为了让读者更好地理解2阶配置模型的构造过程,以图6为例进行详细说明。原始网络图6a中有8条边,因此原始网络可根据边数拆成如图6b所示的8个子图,随机选取其中的一个2阶子图(如含C和G节点的子图),将子图中的节点C视为起始节点,由于节点C的度为4,因此需要从图6b中再随机选取3个子图,每个子图中至少有一个节点的度值为4,图6c显示了这4个度值为4的节点重新随机组合成一个新的子图的过程,进行组合后形成大子图的结果为图6c所示。接着,可以从图6d中随机选取一个有分支的新节点作为起始节点,重复聚合子图的过程,直到所有2阶子图的分支都得到链接最终生成了如图6e所示的2阶零模型。由于在子图的聚合过程中,每条

边两端节点的联合度都不会发生变化, 因此保证了与真实网络具有相同的联合度分布。

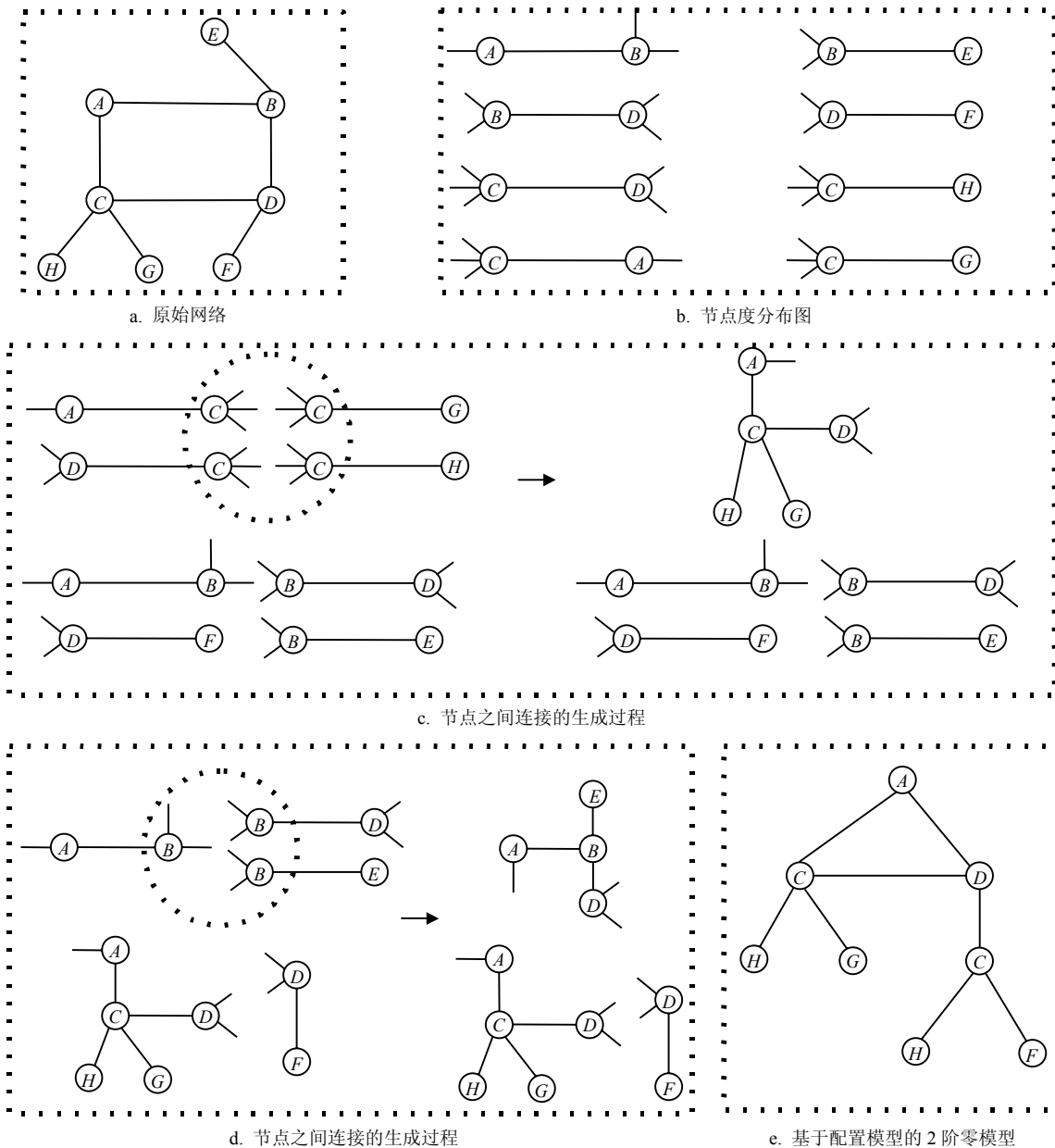


图6 原始网络和使用配置模型构造的2阶零模型网络

### 3.5 3阶配置模型

许多实际网络都具有明显的聚类特性, 而实际网络的高聚类性却没有在低阶配置模型, 如0~2阶模型中体现出来。为了解决这一缺陷, 文献[20]提出了能够保持实际网络聚类特性的3阶配置模型。

在3阶配置模型中将原始网络分成“线”和“角”两种基本元素, 这种形式的拆分保证了网络中各节点的聚类系数不变。定义 $s_i$ 为与节点 $i$ 相连的不构成三角形的单条边的数目,  $t_i$ 为节点 $i$ 所参与的三角形中角的数目。因此, 对于每个节点都可以得到它的联合“线”和“角”序列, 即 $\{s_i, t_i\}$ , 每一个线角序

列也可以称为一个线角子图。在给定网络联合线角序列的前提下, 可以随机选取一对线连接它们形成一条连边, 也可以随机选择不是同一节点的三个角连接它们形成完整的三角形。上述“线”和“角”聚合的过程可以同时进行, 也可以分别进行, 重复聚合过程用完所有的线角子图后就生成了3阶零模型。

为了让大家更好地理解3阶配置模型的构造过程, 下面以图7为例进行更详细的说明。根据前面的联合边角序列的定义, 可以得到图7a中所有节点的联合边角序列值:  $A(0, 2)$ 、 $B(0, 2)$ 、 $C(3, 0)$ 、 $D(3, 0)$ 、

$E(1, 1)$ 、 $F(1, 0)$ 、 $G(1, 0)$ 、 $H(1, 0)$ 、 $M(0, 1)$ 。根据该联合边角序列，先将原始网络拆成如图7b所示的9个边角子图，子图中的灰色阴影部分代表角，实线代表单独存在的边。在这些联合边角序列图中随机选取一对实线连接它们形成一条连边，也可以随机

选择不是同一节点的3个角连接它们形成完整的三角形，进行聚合后生成的大子图和剩余小子图如图7c、7d所示。不断重复角聚合过程或边聚合过程，用完所有的线角子图后就生成了3阶零模型，如图7e所示。

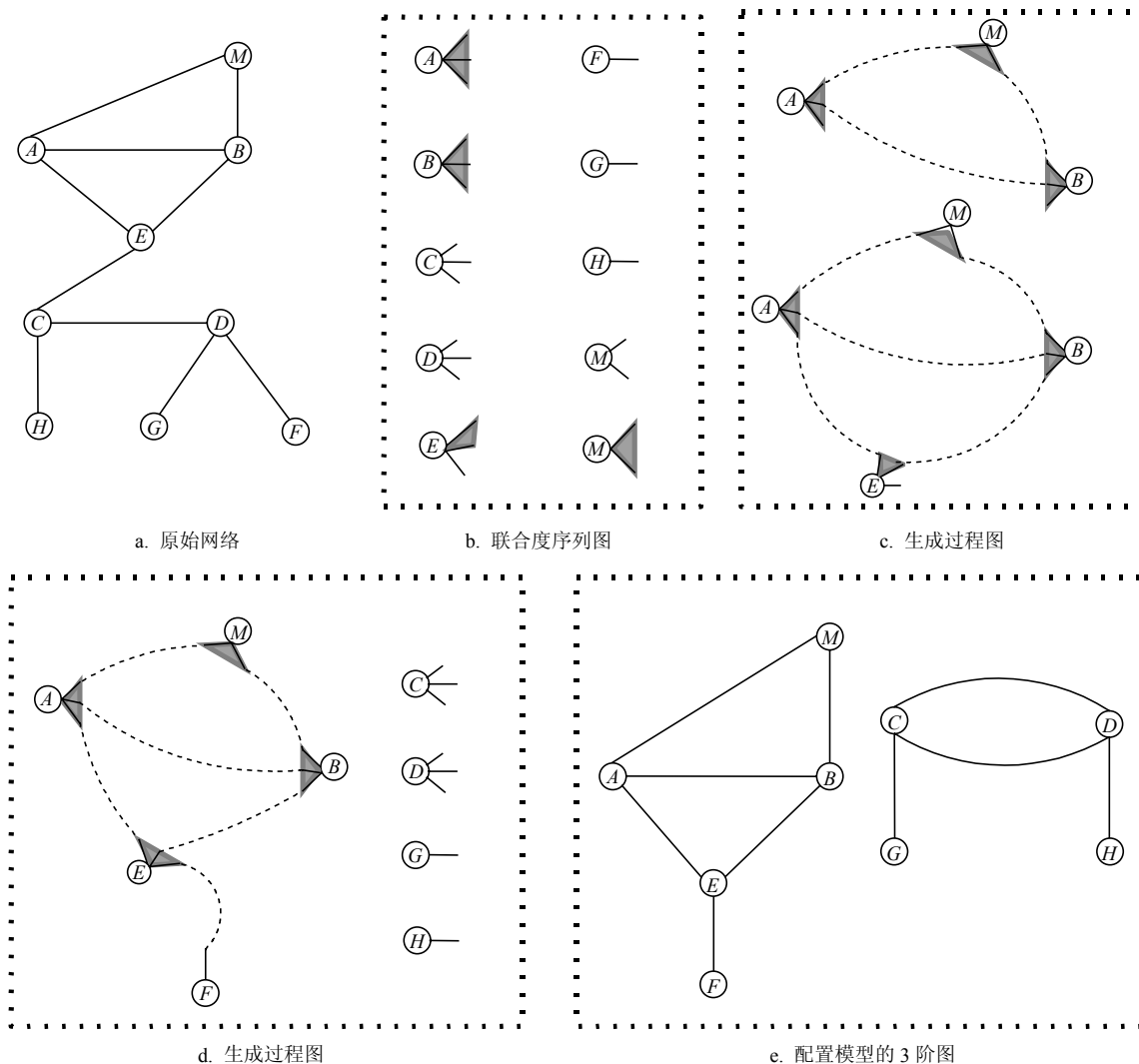


图7 原始网络和使用配置模型构造的3阶零模型网络

此外，依据3阶配置模型的生成原理，也可以尝试构造和原始网络具有相同高阶子图分布的配置模型，从而大大拓展了配置模型的使用范围<sup>[20-21]</sup>。但需要注意的是，尽管配置模型可以精确展现真实网络的度序列、联合度分布、甚至是网络的聚类特性和子图分布，但配置模型也有如下缺点：一是算法比较复杂，因此针对加权网络和时变网络很难使用；二是配置模型产生的网络为了保证连边配对的随机性，可能会存在自环和重边。针对第二个缺陷，可以采用移除自环和重边这种做法，但这种方式也有待商榷之处，因为这样生成的网络不是从所有可能

的匹配集中均匀选择节点或连边，这意味着网络的很多性质不再能够利用解析方法进行计算，至少无法利用当前已知的方法计算<sup>[22]</sup>。同时，由于存在自环和重边，无论是否删除这两种结构，配置模型的连通性等重要指标和原始网络具有较大差异，有些情况下不具有较强的实用价值。

#### 4 随机断边重连的零模型

用ER随机图或配置模型构造的零模型是从无到有生成新网络的过程，而用置乱算法构造的零模型则是将原始网络随机化的过程<sup>[23-24]</sup>。在静态无权

网络中常用的置乱算法就是随机断边重连, 该方法主要是在原始网络结构的基础上将原有连边随机地断开再重连, 使原始网络尽可能的随机化。在静态无权网络中, 相对于配置算法, 连边随机重连更简单、更容易操作, 不需要理解和运用复杂的数学公式、也不会产生自环和重边现象, 却能精确保持真实网络的一些物理属性, 因此被广泛应用到各种类型的实际网络分析中。

4.1 0阶零模型

ER随机图可以作为网络的0阶零模型, 本节采用随机断边重连的方式也可以生成网络的0阶零模型, 它们的共同点是都与原始网络具有相同的平均度分布。不同的是, ER随机图是根据原始网络的节点数和边数生成的新网络, 而随机断边重连的0阶零模型是在原始网络的上进行随机断边重连得到的。

基于随机断边重连的0阶零模型构造过程如图8所示: 首先随机从原始网络中断开一条连边AB, 然后随机选择网络中两个不相连的节点, 如图8a中A、D节点, 最后在A与D之间添加一条连边, 重连后的结果如图8b所示。为了使网络充分随机化, 一般应根据网络的规模将上述断边重连过程重复多次, 直到原始网络足够随机化后才认为生成了对应的零模型。根据复杂网络零模型的量化评估<sup>[25]</sup>中的研究, 为了保证生成的零模型网络足够随机化, 成功断边重连的次数应该为网络边数的两倍以上, 因此本文认为不同阶数零模型网络的随机化在有条件的情况下均应满足这一要求。

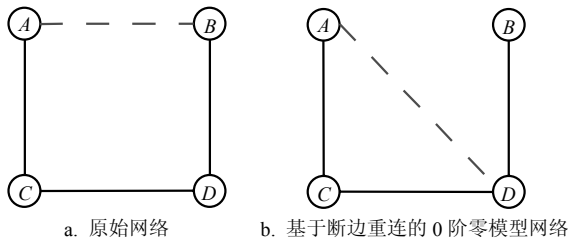


图8 原始网络和使用断边重连算法构造的0阶零模型网络

文献[26]在保证与大众生产合作网络具有相同节点数、边数的前提下, 对合作网络随机断边重连构造了0阶零模型, 并计算了原始网络与零模型网络的聚类系数和平均最短路径长度等多个静态统计量, 如表2所示。

通过分析表2的多个统计量, 可以看出实证网络与0阶零模型网络的平均最短距离长度相差无几, 这说明平均最短距离较小不一定是实际网络所特有的性质。实际网络中的聚类系数远远高于相同规模0阶零模型的系数, 因此可以说高聚类特性是原始网

络所特有的属性而非随机造成的。同样, 通过对比可以得出实际网络符合幂律分布而非ER图等随机模型的泊松分布。通过由零模型作参照物比较原始网络和0阶零模型的聚类系数、平均最短距离及度分布的参数指标证明了开源社区网络具有小世界特性和无标度特性。

表2 实证网络与0阶零模型网络统计量对比表<sup>[26]</sup>

静态参数	实证网络	0阶零模型网络
节点数	36 527	36 527
平均度	13.3	13.3
聚类系数	0.928	$3 \times 10^{-5}$ (平均值)
平均最短距离	4.438	4.351(平均值)
度分布	幂律分布	泊松分布

4.2 1阶零模型

1阶零模型的生成条件要高于0阶零模型, 1阶零模型需要在度分布不变的前提下进行随机断边重连。1阶零模型算法的连边置乱过程如图9所示。如果网络中有节点A与B相连, C与D相连, 且节点A与D不相连, B与C不相连, 具体情况如图9a所示。先将连边AB和CD切断, 再使节点A与D相连, B与C相连, 最终生成图9b。根据网络规模和实际需要不断重复上述过程, 就会使网络中的连边关系被随机置乱, 但是置乱前后每个节点的度均保持不变, 保证了置乱前后网络的度序列相同。从图9可以看出, 连边置乱算法破坏了网络拓扑结构, 因此可用来研究网络拓扑结构各种性质对于网络功能的影响。

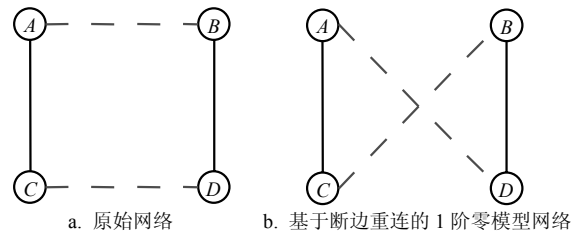


图9 原始网络和使用断边重连算法构造的1阶零模型网络

1阶零模型是原始网络非常重要的参照物。为了确定原始网络是否具有某种不平凡特性, 需要统计分析原始网络的这一特性是否显著地不同于零模型网络<sup>[27]</sup>。假设网络的一个统计量, 如某种子图(模体)在一个实际网络中出现的次数为  $N(j)$ , 而在零模型中出现的次数的平均值为  $\langle N_r(j) \rangle$ , 则两者的比值为:

$$R(j) = \frac{N(j)}{\langle N_r(j) \rangle} \tag{5}$$

在实际应用中, 如果  $R(j) > 1$ , 那么意味着实际网络的设计或者演化过程促进了该拓扑特征的出现。相反, 如果  $R(j) < 1$ , 则意味着实际网络的设计



或者演化过程抑制了该拓扑特征的出现。

在统计学中Z检验方法(平均值差异检验方法)是用标准正态分布的理论来推断差异发生的概率,从而比较两个平均数的差异是否明显。可以用Z检验方法进一步检验某种特性在实际网络中出现的统计量和在零模型中出现的统计量差异的显著性。具体而言拓扑特征  $j$  的统计重要性可用式(6)来刻画:

$$Z(j) = \frac{N(j) - \langle N_r(j) \rangle}{\sigma_r(j)} \quad (6)$$

式中,  $\sigma_r(j)$  是零模型中某个拓扑特征  $j$  出现的平均值  $\langle N_r(j) \rangle$  的标准差。Z值的绝对值越大就表示差异越显著,越能说明拓扑特征  $j$  产生于原始网络,而非随机造成的。

一般来说,Z值与网络的规模有关,网络的节点越多Z值就越高,因此不同的规模网络之间也就没有可比性<sup>[27]</sup>。为了将不同网络上的统计量指标统一到一个尺度上,对Z值做归一化处理使其标准化,得到了如式(7)所示的重要性剖面(significance profile, SP):

$$SP_j = \frac{Z_j}{(\sum Z_i^2)^{1/2}} \quad (7)$$

无向网络的断边重连1阶零模型网络最早被运用于蛋白质网络的研究中,在蛋白质网络中节点代表蛋白质,实验表明两个蛋白质之间如果能够形成化学键的就相互连接起来。在对网络子图结构的研究中,文献[28-29]利用连边置乱算法构造了1阶网络零模型,发现一些特定网络子图在原始网络中出现的频率远高于在零模型网络中出现的频率。文献[30]还对比了原始网络与连边置乱零模型中所有子图出现的频率,提出了重要性剖面算法将来自不同领域的19个网络分成了4个超家族。

在实际网络中也可采用相关性剖面来量化度相关特性。设  $m(k_0, k_1)$  设为度为  $k_0$  和  $k_1$  的节点之间相连边的总数,可以通过比较一个实际网络的  $m(k_0, k_1)$  和其1阶零模型所对应的均值  $\langle m_r(k_0, k_1) \rangle$  来分析实际网络的度相关特性。具体的做法是通过在平面上绘制出相关性剖面(correlation profile, CP)来刻画实际网络的度相关性,具体的R值和Z值相关性剖面计算方式如下:

$$R(k_0, k_1) = \frac{m(k_0, k_1)}{\langle m_r(k_0, k_1) \rangle} \quad (8)$$

$$Z(k_0, k_1) = \frac{m(k_0, k_1) - \langle m_r(k_0, k_1) \rangle}{\sigma_r(k_0, k_1)} \quad (9)$$

文献[31]将AS网络和利用连边置乱生成的零模型进行对比,通过量化网络相关性剖面的新方法检验了Internet网络中的度度相关特性。图10显示的是AS网络相关性剖面图,图10a绘制的是R值分布图,竖条表示的是  $R(k_0, k_1)$  值数值分布对照图。当  $R(k_0, k_1) > 1$  时,表明在度为  $(k_0, k_1)$  的节点之间的连接强度比相应的随机网络更强;当  $R(k_0, k_1) < 1$  时,表明在度为  $k_0, k_1$  的节点之间的连接强度比相应的随机网络更弱。图10b绘制的是Z值分布图,竖条是  $Z(k_0, k_1)$  值数值分布对照图,Z的绝对值越大就表示差异越显著。总的来说,R值是可检测原始网络度度相关程度是受到了促进还是抑制,侧重于定性;Z值可进一步检测这些连接强度(促进或抑制)的强烈程度,侧重于定量。

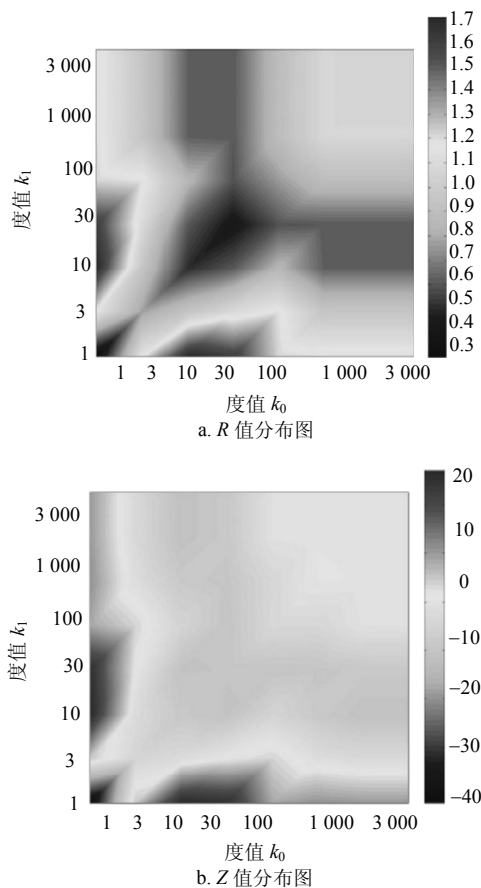


图10 AS网络的相关性剖面图<sup>[31]</sup>

通过图10可以看出:当节点的度值分布在  $3 \geq k_0, k_1 \geq 1$  时,通过式(8)计算得出的  $R(k_0, k_1)$  值从图10a中可以看出在0.3附近,表明了连接强度比相应的随机网络更弱。为了进一步刻画拓扑特征的显著性差异是否明显,用式(9)计算得出了  $Z(k_0, k_1)$  值的绝对值并通过图10b可以看出为-30左右,通过Z值的刻画可以说明原始网络的度度相关性与随机网络相

比差异非常显著, 因此可以说明当节点的度在  $3 \geq k_0, k_1 \geq 1$  范围时连边受到很强的抑制, 也可以说是度值小的节点倾向于和度值小的节点相连的这种度相关性受到了很强的抑制。

### 4.3 2阶零模型

静态无权网络中最常用的就是使用随机断边重连来产生1阶复杂网络零模型, 为了更精确地刻画网络的拓扑结构, 也可以用随机断边重连的方法构造2阶和更高阶次的零模型<sup>[2-3]</sup>。高阶零模型的随机重连算法可以看作1阶随机断边重连算法的拓展, 2阶零模型的随机断边重连算法是保持联合度分布不变的连边置乱, 其具体步骤与前述的1阶方法一样, 只是多了一个限制条件: 即要求交换前后两条边连接节点的度值是不变的。如图11所示, 只有节点B和D(或A和C)具有相同的度值时, 连边交换才可以进行, 以保证网络的联合度分布不变。随着零模型阶次的增加约束条件也越来越多, 可以被随机断开再重连的连边就越来越少, 因此网络拓扑被破坏得也越来越小, 所构造的零模型也越来越精确。

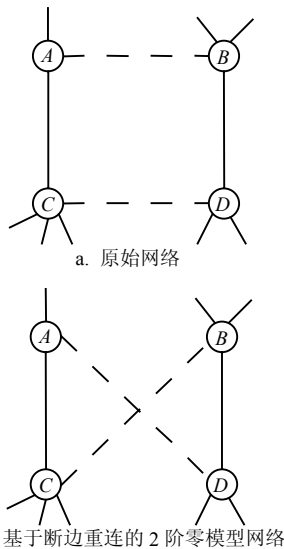


图11 原始网络和断边重连算法构造的2阶零模型网络

在网络的高阶零模型研究中, 文献[3]做了非常重要的系统性工作。一方面其系统性地研究了不同阶次复杂网络零模型的度相关特性; 另一方面, 在保持原始网络拓扑特性的前提下, 提出了构造原始网络不同规模零模型的算法并提供了软件Orbis供研究者参考和使用<sup>[2]</sup>。此外, 文献[1]讨论了如何结合不同网络抽样算法生成网络2.25阶和2.5阶零模型的理论算法与实现过程。

### 4.4 2.25阶零模型

2.25阶零模型不但需要在随机断边重连的过程

中保证2阶零模型的基本性质不变, 而且也需要保证在网络断边前后平均聚类系数的值不变。随着零模型阶数的增加, 约束条件也会变得愈加苛刻, 生成连边的成功率也越来越低。

2.25阶零模型构造的基本步骤是: 首先在原始网络中随机选取两条边, 要保证联合度分布不变即2阶零模型的基本性质不变。其次, 断边重连之前计算两条连边两端节点平均聚类系数的值, 经过断边重连之后计算新生成的连边两端节点平均聚类系数的值。若断边重连前后所选边上的节点中平均聚类系数的值不变, 则证明断边过程是成功的。若平均聚类系数的值不一致, 则证明所选的边不符合条件, 需要重新再选。

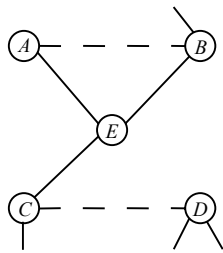
### 4.5 2.5阶零模型

随机断边重连生成的2.5阶零模型的约束条件与随机断边重连生成的2.25阶零模型的约束条件大致相同, 不同的是在生成2.5阶零模型过程中需要保证在网络断边前后度相关的聚类系数值不变。文献[1]结合网络抽样算法生成2.5阶零模型网络, 这里描述一下如何使用随机断边重连的方式来生成2.5阶零模型。

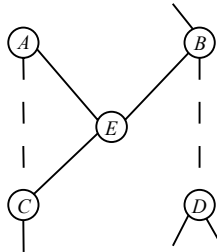
2.5阶零模型构造的基本步骤是: 首先在原始网络中随机选取两条边, 要保证联合度分布不变即2阶零模型的基本性质不变。其次, 断边重连之前计算两条连边两端节点度相关的聚类系数值, 经过断边重连之后计算新生成的连边两端节点聚类系数的值。若断边重连前后所选边上的节点中度相关的聚类系数值不变, 则证明断边过程是成功的。若度相关的聚类系数的值不一致, 则证明所选的边不符合条件, 需要重新再选。

以图12为例来具体阐述由原始网络生成2.5阶零模型网络的过程。在生成2.5阶零模型之前, 首先要计算原始网络中随机选取的两条边上4个节点的聚类系数, 然后得出这4个节点的度相关的聚类系数分布情况。在图12a中随机选取两条边AB和CD, 其中度值为2的A节点的聚类系数值为  $C_A = 1$ , 因此节点的度为2的平均聚类系数值为  $\bar{C}(2) = 1$ 。节点B的聚类系数  $C_B = 1/3$ , 节点C、D的聚类系数的值都为0, 且B、C、D节点的度都为3, 节点的度为3的平均聚类系数的值为  $\bar{C}(3) = 1/9$ 。此时, 就得到了在断边重连之前度值为2、3节点的平均聚类系数分布情况:  $\bar{C}(2) = 1, \bar{C}(3) = 1/9$ 。将选取的两条边AB和CD切断, 再使节点A与C相连, B与D相连。具体的生成过程如图12b所示, 在经过断边重连后的网络中A节点的度

为2, 经计算可得  $\bar{C}(2)=1$ ,  $B$ 、 $C$ 、 $D$ 节点的度都为3, 节点的度为3的平均聚类系数的值为  $\bar{C}(3)=1/9$ , 所以在经过断边重连之后度值为2、3的节点的平均聚类系数分布情况:  $\bar{C}(2)=1$ ,  $\bar{C}(3)=1/9$ 。通过对断边重连前后随机选取的两条边上4个节点的度相关的聚类系数的度值分布情况的对比, 可以看出其值是不变的。因此图12b是经过随机断边重连后生成的2.5阶零模型, 此模型保证了网络中的联合度分布及度相关的聚类系数值不变。

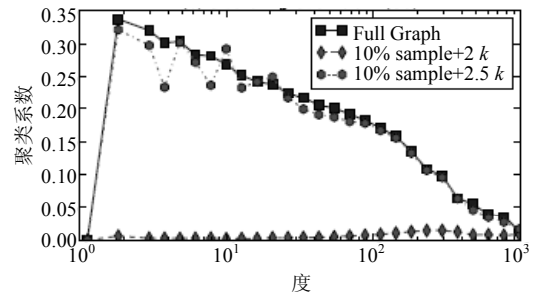


a. 原始网络

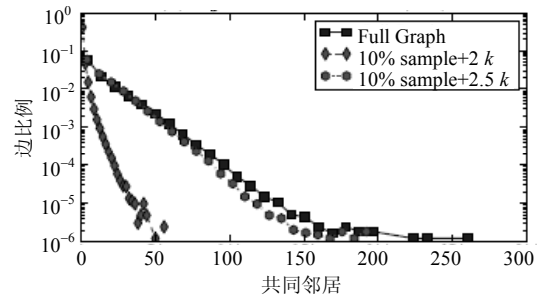


b. 基于断边重连的2.25阶(2.5阶)零模型网络

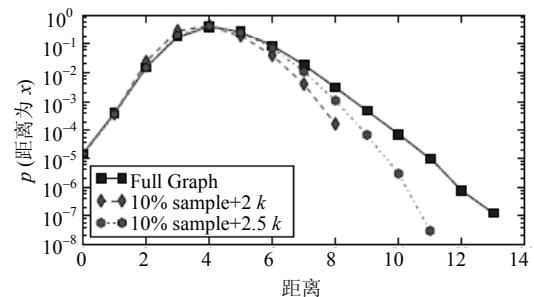
图12 原始网络和使用断边重连算法构造的2.25阶(或2.5阶)零模型网络



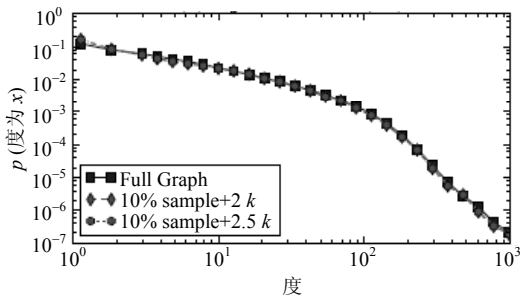
c. 聚类系数(CC)



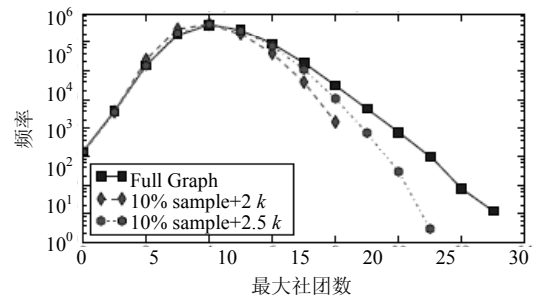
d. 边共享伙伴(ESP)



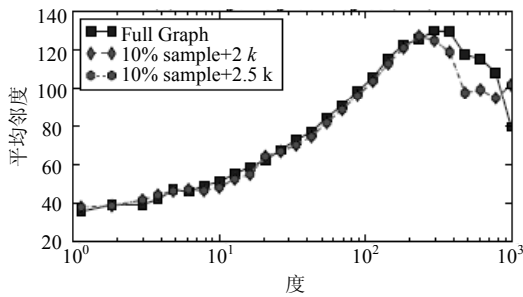
e. 最短路径(Sh.p.)



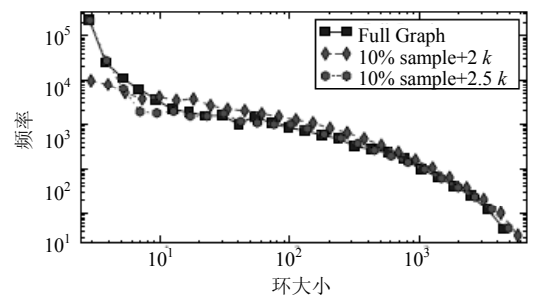
a. 度分布(DD)



f. 最大团



b. 平均邻度(knn)



g. 环基础

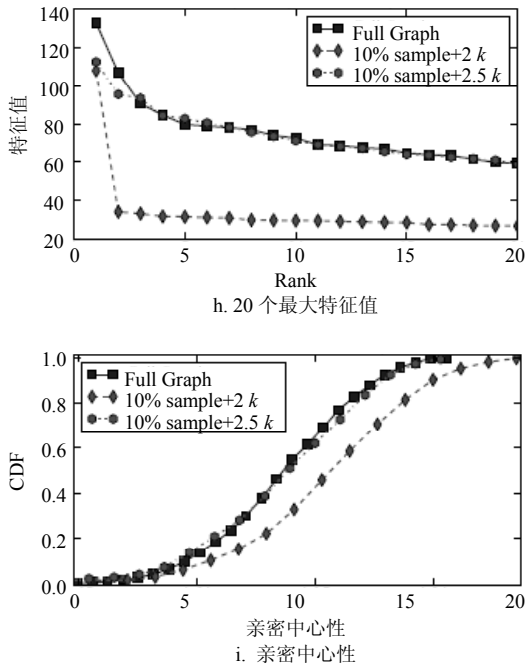


图13 Facebook局部网络的2阶零模型与2.5阶零模型比较<sup>[1]</sup>

文献[1]通过对节点的抽样生成了联合度分布和度相关的聚类系数的有效估计量并成功构建了2阶零模型网络和2.5阶零模型网络, 图13显示的是Facebook网站中新奥尔良城市子网络9种不同的性质分布图。从图13中可以明显看出9个性质图的分布情况, 在9个不同性质的分布图中2阶、2.5阶零模型与原始网络中相应的性质做比较, 可以清楚地看出2.5阶零模型相对于2阶零模型更接近原始网络。上述案例有效地说明了约束条件越多, 生成的网络零模型越能逼近于原始网络。

### 4.6 3阶零模型

对于3阶网络零模型, 该零模型受到的约束条件较多, 尤其对于一个大规模网络, 不能在一个较短的计算时间内生成相应的零模型, 因此有时不具备实用性。但是对于小规模网络, 尽管有时也不能保证遍历行, 但是3阶网络零模型在有限时间内还是可以构造出来并实际使用的, 因此下文中给出3阶零模型的构造过程。

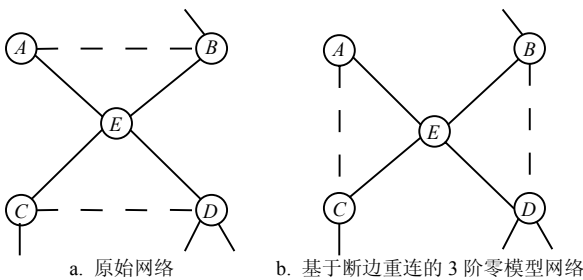


图14 原始网络和断边重连算法构造的3阶零模型网络

基于随机断边重连的3阶零模型构造的基本步骤是: 首先在原始网络中随机选取两条边, 要保证度相关的平均聚类系数不变, 即2.5阶零模型的基本性质不变。其次, 断边重连之前计算随机选取的两条连边连接节点以及它们邻居节点的开三角形和闭三角形的个数, 经过断边重连之后计算新生成的连边连接节点和其邻居节点的开三角形和闭三角形的个数。若断边重连前后所选边上节点和其邻居节点的开三角形和闭三角形的个数不变, 证明断边重连过程是成功的。若开三角形和闭三角形的个数不一致, 则证明所选的边不符合条件, 需要重新再选。

## 5 检验富人俱乐部效应的局部断边重连零模型

幂律分布也被称做无标度分布, 具有幂律分布的网络一般被称为无标度网络。在一个无标度网络中, 绝大部分节点的度相对很低, 但存在少量节点的度相对较高, 这样的节点被称为富节点或中枢(Hub)节点。很多复杂网络都具有幂律形式的度分布, 而择优连接是其中非常重要的一种幂律生成机制, 例如通常我们写一篇新文章时总是倾向于引用一些引用次数较多的文献。Hub节点在网络中占有比较重要的位置, 这样的节点由于拥有大量的社会连接会对各种扩散行为产生影响, 如在新产品的开发推广中具有Hub节点性质的人往往能起到扩大推广的作用。

### 5.1 网络富人俱乐部系数的定义

富节点之间产生连边的现象为富人俱乐部现象。富节点之间是否更倾向于产生连边在无权网络中可以用无权富人俱乐部系数来定量表示:

$$\varphi(n) = \frac{L_{>n}}{N_n} \tag{10}$$

具体计算过程如下: 首先定义节点度大于一定值的节点为富节点, 然后计算这些富节点之间的所有连边之和  $L_{>n}$ , 接着计算富节点个数之和为  $N_n$ , 最后再求出前面两项的比值。如果富节点之间两两相互连接, 那么上面计算出的比值应为  $(N_n - 1)/2$ 。一般情况下, 并不是所有富节点之间都存在连接, 计算出来的值越大, 则证明富节点之间连接的越紧密。当富节点之间没有连边时, 该值为0。这个比值是一个绝对的量, 这个量的大小是依赖于具体的实际网络, 因此为了有可比性, 很多时候都使用实际网络和其对应的随机化零模型网络进行对比, 判别公式为:

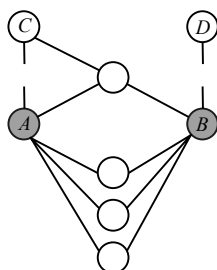
$$\rho(n) = \frac{\varphi(n)}{\varphi_{\text{null}}(n)} \quad (11)$$

式中,  $\varphi_{\text{null}}(n)$  代表的是实际网络随机化后的富人俱乐部系数。根据  $\rho(n)$  是否大于1作为加权网络中是否具有富人俱乐部现象的依据<sup>[28]</sup>。如果  $\rho(n) > 1$ , 表明原始网络具有富人俱乐部效应, 也就是说富节点之间更倾向于相互连接。 $\rho(n) < 1$ , 则表明原始网络不具有富人俱乐部效应, 即富节点之间不倾向于相互连接。

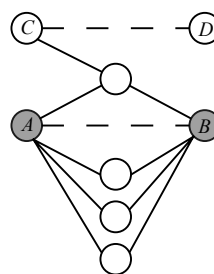
## 5.2 网络富人俱乐部零模型的构造过程

为了研究Hub节点之间的连接特性以及这种连接特性对于整个网络性质的影响, 文献[32]提出了复杂网络的富人俱乐部系数, 在Internet网络研究中没有使用该网络的零模型仅仅使用了几种常用网络做基准模型, 发现Internet网络具有富人俱乐部特性。文献[33]不仅仅计算了多种网络的富人俱乐部系数绝对值, 还计算了每个网络1阶零模型的富人俱乐部系数, 通过使用统计方法比较后得出了和文献[32]截然相反的结论: Internet网络不具有富人俱乐部特性<sup>[33]</sup>。

针对Internet网络是否具有富人俱乐部特性引起的诸多争论<sup>[33-34]</sup>, 文献[35]提出了改变网络富人俱乐部系数的局部连边随机重连算法, 再次有力证明了Internet网络确实存在富人俱乐部特性。具体的算法构造过程如图15所示, 其中网络中节点A和B是富节点, 重连以前这两个富节点之间无连接, 没有形成富人俱乐部。首先, 在网络中任意选取两个分别与AB相连的非富节点C和D(如图15a所示), 且C和D之间不存在连边。然后, 采用上文提到的连边重连算法将连边AC和BD切断, 再使节点A与B相连, C与D相连, 这样富节点A与富节点B相连后网络就有了富人俱乐部性质(如图15b所示)。上述算法描述的是如何使网络构造具有强的富人俱乐部特性, 如果将这一过程反过来运用, 也可构造出一个几乎没有富人俱乐部特性的零模型网络。



a. 原始网络



b. 基于断边重连的具有富人俱乐部特性网络

图15 原始网络和使用连边置乱算法构造的具有富人俱乐部特性网络

局部连边随机重连算法几乎可以任意改变中枢节点之间的连接特性, 通过将完全没有富人俱乐部特性的网络、原始网络以及具有极强富人俱乐部特性的几种网络进行比较, 就可以分析出富节点之间的连接特性对于整个网络性质的影响。文献[36]通过这一方法分析了多种网络中富人俱乐部特性对于网络聚类系数、匹配系数和平均路径长度的影响, 并讨论了在不同度分布网络中富人俱乐部特性的影响力大小。

## 6 检验匹配特性的有倾向性断边重连零模型

在各种实证网络中, 如果度大节点倾向于和度大节点相连, 度小节点倾向于和度小节点相连, 这种倾向性即为度匹配的正相关性; 或者度小节点倾向于和度大节点相连, 这种倾向性即为度匹配的负相关性。连边随机重连算法不仅破坏了网络拓扑, 也扰乱了网络本身的度相关性, 使连边不再具有同配或异配特性。文献[24]利用随机1阶断边重连算法构造的零模型与原始网络进行对比, 发现在蛋白质相互作用网络和基因调控网络中, 度大与度大节点之间的连边受到了系统抑制, 而度大与度小节点之间的连边出现的可能性则更大。此外, 文献[31]还将原始网络和利用连边置乱生成的零模型进行对比, 通过相关性剖面的新方法检验了Internet网络中的度度相关特性。以上研究都说明了度度相关特性对于网络的重要性, 那么有没有办法构造出满足某种度匹配模式的零模型呢?

如果想改变原始网络的度度匹配特性, 就需要有选择地进行断边重连来构造同配或异配网络。如图16所示, 原始网络图16a中有A、B、C、D 4个节点, 其中节点A与C相连, B与D相连, 相连节点之间的度有一定差距但相差不大, 属于中性网络。如果将原始网络的度大节点A和度大节点D相连、度小

点B和度小节点C相连(如图16b所示), 那么就增强了网络的同配特性。将这一过程反复进行, 就可以生成强同配性的网络。相反, 将原始网络的度最大的节点A和度最小的节点B相连, 其余节点相连(如图16c所示), 那么就增强了网络的异配特性, 反复进行就可以生成强异配性的网络。

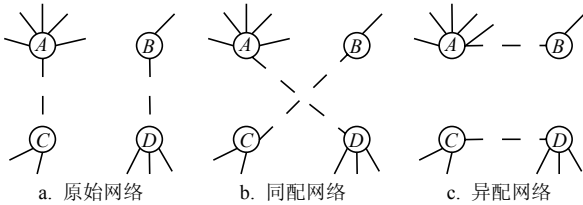


图16 复杂网络中3种不同的度匹配模式

使用上述有倾向性的连边重连算法生成强同配网络和异配网络, 可以研究网络拓扑结构对于度度相关特性的约束以及这种约束有何影响。文献[37-38]定义的匹配系数是复杂网络研究中最常用的衡量网络匹配特性的统计量, 具体公式如下所示:

$$r = \frac{M^{-1} \sum_i j_i k_i - \left[ M^{-1} \sum_i \frac{1}{2} (j_i + k_i) \right]^2}{M^{-1} \sum_i \frac{1}{2} (j_i^2 + k_i^2) - \left[ M^{-1} \sum_i \frac{1}{2} (j_i + k_i) \right]^2} \quad (12)$$

式中,  $M$  为网络中的边数;  $j_i$ 、 $k_i$  分别为第  $i$  条边

的两端节点的度数。如果  $r > 0$ , 表明网络是同配性质的, 如果  $r < 0$ , 表明网络是异配性质的。

理论上讲实际网络最大同配的时候该值应该接近于1, 最大异配的时候该值应该接近-1, 但实证研究表明并非如此。图17给出了3个网络——因特网、蛋白质相互作用网络、科学合作网络的匹配系数分布图。文献[35]在保证度分布不变、度分布与其度为  $k$  的富节点连通性不变的两种情况下, 对这3个原始网络分别进行随机重排、同配性重连和异配性重连, 每个原始网络都生成了相应的随机网络、强同配网络、强异配网络。

图17中的实线为在保持度分布不变的情况下按要对网络进行的断边重连, 虚线为在保持度分布与其度为  $k$  的富节点连通性不变的情况下按要对网络进行的重连, 并将重连后的网络用式(12)量化得到3个网络的同配系数  $r$  值的范围。由图可知, 在保持度序列不变的情况下, 在对蛋白质相互作用网络、科学合作网络进行随机重排的过程中发现重排后网络的匹配系数在0附近。在上述的两种网络分别进行强同配性重排和强异配性重排后所产生的1阶强同配零模型和1阶强异配零模型的同配系数  $r$  值不能完全覆盖-1到1这个范围, 但是蛋白质相互作用网络同配系数  $r$  值的范围比科学合作网络的  $r$  值范围更窄。

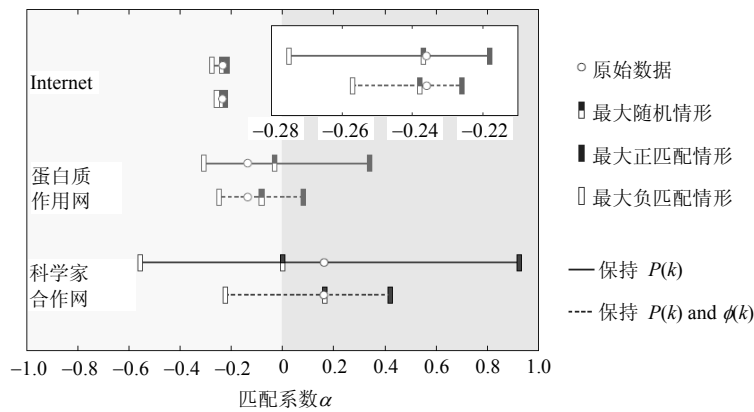


图17 不同网络的匹配系数r值分布情况图<sup>[35]</sup>

同时, 文献[38]发现实际的 Internet 网络, 在保证度序列相同的情况下无论怎么进行连边重连, 它的匹配系数总是在一个非常小的范围内且总是负值, 这说明在对实际网络进行最大同配重连的  $r$  值并非都会接近于 1, 最大异配重连的  $r$  值并非都会接近于-1。在保持度序列与富节点连通性不变的情况下, 分别对网络进行上述相同的随机重连、同配性重连和异配性重连, 可以看到相对应的匹配系数  $r$

值的范围变得更窄了, 这说明富节点之间的连通性会进一步影响  $r$  值的分布情况。

针对匹配系数这一缺陷, 文献[39]通过比较原始网络和它的零模型, 分析了 Internet 网络中超级节点引起网络匹配系数异常的原因并给出了一个简单的修正方式。如果一个网络中存在度远远大于其他节点的节点, 那么该节点可以视为网络中的“超级节点”, 超级节点和其他节点之间悬殊的度差别导致它

们和其他节点连接会导致负的匹配特性, 因此为了避免超级节点的干扰使整个网络的匹配特性度量不准确, 比较简单的方法就是先从网络中将超级节点过滤掉, 再计算网络的匹配系数。假定排除了超级节点后网络中的连边数为  $M_c$ , 则计算匹配系数的公式可简单修改为:

$$r_c = \frac{M_c^{-1} \sum_i j_i k_i - \left[ M_c^{-1} \sum_i \frac{1}{2} (j_i + k_i) \right]^2}{M_c^{-1} \sum_i \frac{1}{2} (j_i^2 + k_i^2) - \left[ M_c^{-1} \sum_i \frac{1}{2} (j_i + k_i) \right]^2} \quad (13)$$

式中,  $M_c$  为网络中排除超级节点后的边数;  $j_i$ 、 $k_i$  分别为第  $i$  条边的两端节点的度数。如果  $r_c > 0$ , 表明网络是同配性质的; 相反, 如果  $r_c < 0$ , 表明网络是异配性质的。

度相关性在网络中具有很强的影响, 如在网络的扩散现象中担当非常重要的角色, 它会对信息传播的规模和速度产生重要影响。对于无标度网络, 网络中的度相关性会影响谣言扩散的阈值和最终的扩散规模, 增加网络的度相关性可以显著地抑制扩散。无标度网络中存在大量度值小的节点, 增强度相关性意味着度值小的节点倾向于连在一起, 而非连向度值大的节点, 这使得度值大的节点在信息扩散过程中很可能变得孤立, 从而抑制了扩散。此外, 增强度相关性也会抑制 ER 随机网络中的谣言扩散, 但由于随机网络的度分布较为均匀, 这种影响比无标度网络中度相关性的作用要小得多。

## 7 依赖于社团结构的零模型

社团检测是近年来复杂网络领域非常重要的一个研究问题, 社团结构是复杂网络重要的中尺度结构特征, 社团可能对应于网络的基本功能单元, 可以揭示网络结构与功能之间的关联关系<sup>[40-44]</sup>。考虑到社团结构特征的重要性, 本文新构造了4种依赖于社团结构的零模型: 社团内部断边重连零模型、社团间断边重连零模型、增强社团结构零模型和减弱社团结构零模型<sup>[45-49]</sup>。

### 7.1 社团内部断边重连零模型

在前面介绍的随机断边重连0~3阶零模型中, 虽然每次仅改变网络的微观结构, 但是也消除了网络中原来存在的社团结构。而在本节中介绍的社团内部断边重连零模型, 它只改变社团内部连边的拓扑结构, 不改变社团之间的链接关系, 因此保持了原始网络的社团结构<sup>[50-53]</sup>。

社团内部连边重连零模型网络的构造过程如图18所示。首先采用某种社团检测算法将原始网络分为多个社团, 然后在保持其他连边结构不变的情况下, 每次只交换某一个社团内部连边。如首先选取图18a中所示的社团A中的两条虚线连边A1-A4和A2-A3, 采用断边重连1阶零模型的方式先将其断开后将A1与A3相连、A2与A4相连, 重连后的结果如图18b所示。

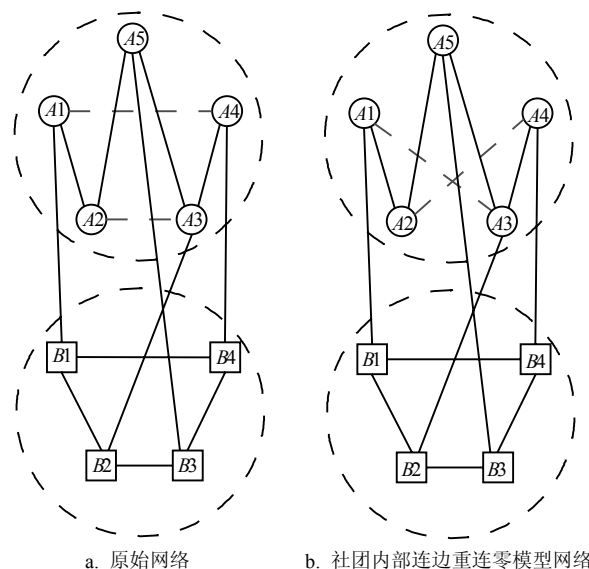


图18 社团内部断边重连零模型

在生成社团内部断边重连零模型的过程中, 除了使用1阶零模型的构造方式, 也可以使用2阶、2.25阶、2.5阶和3阶等阶零模型。此外, 该模型可以分别置乱不同社团内部的拓扑结构, 再计算网络的统计特征就可以知道每个社团内部结构在整个网络中的影响与作用。

### 7.2 社团间断边重连零模型

社团间断边连边重连零模型的构造过程如图19所示。首先将原始网络划分为多个社团, 然后在保持每个社团内部连边结构不变的情况下, 只随机置乱社团间的连边。如图19a中所示, 将社团A和社团B间的两条虚线连边A1-B1和A4-B4断开, 然后将A1与B4相连、A4与B1相连的结果如图19b所示。

比较图19a和19b两个网络的拓扑结构, 社团间断边重连零模型基本保持了原始网络的各种微观结构特征(0~3阶), 而仅仅是置乱了社团间的连边, 对社团结构造成一定的扰动。在构造这种零模型网络的过程中, 除了使用1阶零模型的构造方式, 也可以使用2阶、2.25阶、2.5阶和3阶等阶零模型。社团内部与社团间断边重连两种零模型构造方式可以有

效区分出社团内部拓扑结构和社团间连边关系对整个网络的影响。

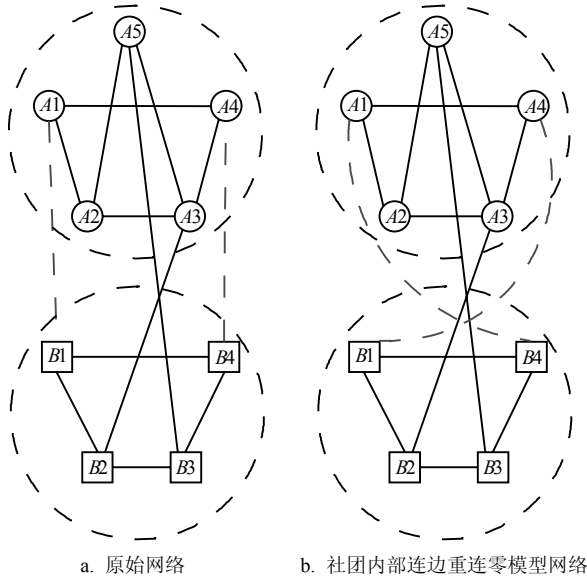


图19 社团间断边重连零模型

### 7.3 增强社团结构的零模型

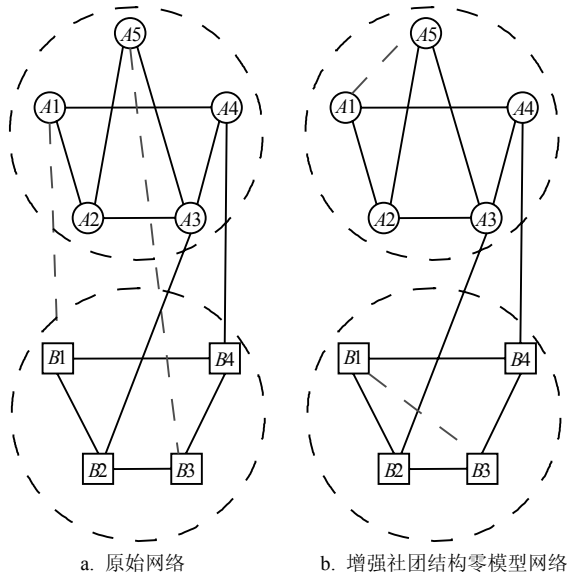


图20 增强社团结构的零模型

社团结构一般会呈现出社团内部的节点之间连接稠密、属于不同社团的节点之间连接稀疏的特点。如果要增强原始网络的社团结构, 那就需要减少社团之间的连边, 增加社团内部的连边。增强社团结构的零模型构造过程如图20所示。首先将原始网络划分为多个社团, 然后在保持其他连边结构不变的情况下, 将两个社团之间的连边交换为社团内部节点之间的连边。如图20a所示, 将社团A和社团B间的两条虚线连边A1-B1与A5-B3断开, 然后分别将社团

A中的两个节点A1与A5相连、将社团B中的两个节点B1与B3相连, 最后生成的网络拓扑结构如图20b所示。

### 7.4 减弱社团结构的零模型

基于断边重连的减弱社团结构零模型构造过程如图21所示。首先将原始网络划分为多个社团, 然后在保持其他连边结构不变的情况下, 将社团内部的连边交换为两个社团之间的连边。将图21a中所示的社团A内部的虚线连边A1-A4和社团B内部的虚线连边B1-B4断开, 然后将社团A和社团B间的节点A1与B4相连、A4与B1相连, 重新连接后的结果如图21b所示。

通过增强或减弱社团结构的零模型可以有效在保持真实网络拓扑结构基本不变的情况下, 增强或减弱社团结构特性。对于各种社团检测算法, 可有效识别出不同程度社团结构特性情况下, 每种社团检测算法的稳定性和鲁棒性。

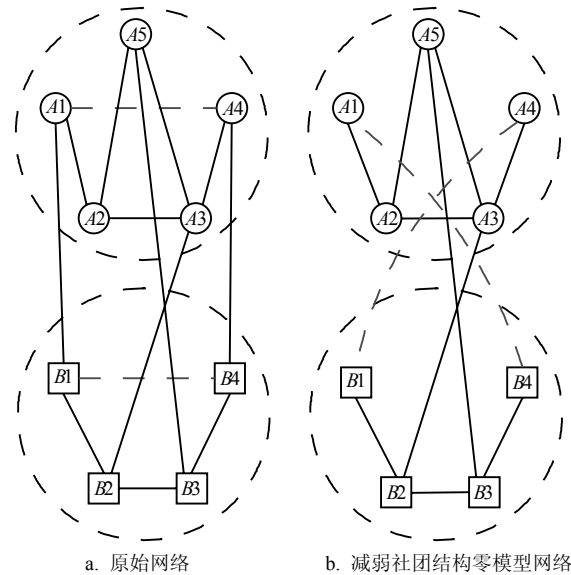


图21 减弱社团结构的零模型

## 8 无权有向网络的零模型

与无向网络不同, 有向网络中节点的度分为出度和入度两种。节点的出度是指以该节点为起点有方向性的指向另一个节点, 而节点的入度则表示其他节点有方向性的指向该节点。在有向网络中每个节点的出度和入度可能并不相同, 但是每条有方向性的边既是一个节点出度同时又是另一个节点的入度, 因此从网络的整体来看, 整个网络的出度总数总是等于入度总数。以微博用户关注为例, 普通人对名人的关注可以视为普通人的出度同时又可视为名人的入度, 在微博所构成的社交网络来说出度



的总数总是相等的。此外，部分网络中单个节点的出度与入度大小相似，可将此类有向网络转化为无向网络，从而利用无向网络的置乱算法对比研究网络的局部结构特征和链路预测算法的稳定性<sup>[54-55]</sup>。

对无向网络采用断边重连的方式生成零模型能够研究无向网络的度相关性<sup>[35]</sup>，同理，对有向网络进行断边重连可以研究网络的出入度相关性，即在有向网络中相连节点之间的出度和出度、入度和入度、入度和出度、以及出度和入度之间有何关系<sup>[56]</sup>。有向网络的断边重连与无向网络类似，不同的是需要分别考虑将节点的出边或者入边断边重连。

以1阶零模型为例，该算法的构造过程如图22所示，如果在有向网络中有节点A有出度指向B，C有出度指向D；且节点A与D不相连，B与C不相连(如图22a所示)。我们就将连边AB和CD切断，再使节点A的出度指向D，C的出度指向B(如图22b所示)。根据网络规模和实际需要不断重复上述过程就会使网络中的连边关系被随机置乱，但是通过置乱前后保证每个节点的出度和入度均保持不变，最终保证了置乱前后网络的出入度序列相同。

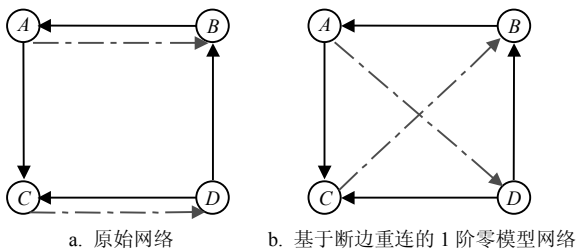


图22 原始有向网络和使用断边重连算法构造的1阶零模型网络

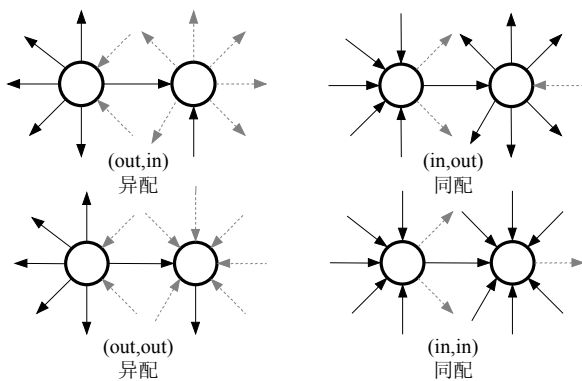


图23 有向网络的4种度相关性<sup>[56]</sup>

在无向网络中，可通过计算度同配系数是否大于0来判断网络中的同异配性质。而在有向网络中，每个节点有出度和入度两个不同的度，因此在有向网络中存在如图23所示的4种匹配特性的情况： $r(out,in)$ ,  $r(in,out)$ ,  $r(out,out)$ ,  $r(in,in)$ ，这4种情况

分别对应高出(入)度的节点有边指向高入(出)度的节点的倾向性程度，如  $r(in,out)$  表示高入度的节点有边指向高入度的节点的倾向性程度<sup>[56]</sup>。

在有向网络的研究中，文献[56]利用原始网络和零模型中度相关性的对比，提出了有向网络的同配系数可以用下式表示：

$$r(\alpha, \beta) = \frac{M^{-1} \sum_i (j_i^\alpha - \langle j^\alpha \rangle)(k_i^\beta - \langle k^\beta \rangle)}{\sqrt{M^{-1} \sum_i (j_i^\alpha - \langle j^\alpha \rangle)^2} \sqrt{M^{-1} \sum_i (k_i^\beta - \langle k^\beta \rangle)^2}} \quad (14)$$

式中， $\alpha$  表示出度类型； $\beta$  表示入度类型； $M$  网络的边数； $j_i^\alpha$  表示有向边  $i$  的出节点的度值； $k_i^\beta$  表示有向边  $i$  的入节点的度值； $\langle j^\alpha \rangle$  表示在零模型中出节点出现次数的均值； $\langle k^\beta \rangle$  表示在零模型中入节点出现次数的均值； $r(\alpha, \beta)$  的统计重要性可以用  $Z$  值来刻画：

$$Z(\alpha, \beta) = \frac{r(\alpha, \beta) - \langle r_r(\alpha, \beta) \rangle}{\sigma_r(\alpha, \beta)} \quad (15)$$

式中， $\langle r_r(\alpha, \beta) \rangle$  是1阶零模型中同配系数的均值； $\sigma_r(\alpha, \beta)$  是同配系数的标准差。通过对上面的  $Z(\alpha, \beta)$  进行归一化处理可以得到同配重要性剖面 (assortativity significance profile, ASP)：

$$ASP(\alpha, \beta) = \frac{Z(\alpha, \beta)}{(\sum_{\alpha, \beta} Z(\alpha, \beta)^2)^{1/2}} \quad (16)$$

ASP( $\alpha, \beta$ ) 可用来检测有向网络的匹配特性，当  $ASP > 0$  时，表明原始网络比零模型更具有同配性；当  $ASP < 0$  时，表明原始网络比零模型更具有异配性。ASP 绝对值的大小反映了原始网络同配或异配的强弱程度，并可用于比较不同的有向网络之间匹配特性的强弱程度。文献[56]在保证原始网络度序列不变的前提下，对WWW, Wikipedia, Political Blogs 3个有向网络进行随机重连生成了相应的1阶零模型，并根据式(16)计算了3个有向网络的同配重要性剖面的值，生成了图24所示的匹配性质网络对比图。图24中的横坐标表示有向网络匹配性质的4种度量，纵坐标表示 ASP 的值。以  $r(out,in)$  度量情况为例，通过对比3个网络的 ASP 值，对应于纵坐标可以看出3个网络的 ASP 值都小于0，因此可以确定这3个原始网络都具有异配性。通过对 ASP 的绝对值对比可以看出3个网络之间匹配特性的强弱程度为：Political Blogs > Wikipedia > WWW。

将同一个网络的  $r(out,in)$ ,  $r(in,out)$ ,  $r(out,out)$ ,  $r(in,in)$ ，进行横向比较，可以发现同一个有向网络

的不同出入度关系之间可能具有多种匹配特性, 在所给的3个有向网络中并不存在单一的同配或者异配性质。如它们的 $r(\text{out}, \text{in})$ 和ASP值都小于0, 因此可以确定这3个原始网络的出度-入度之间都具有异配性。相反, 在 $r(\text{in}, \text{out})$ 度量的情况下, 3个网络之间呈现出同配特性。然而在 $r(\text{out}, \text{out})$ ,  $r(\text{in}, \text{in})$ 两种度量的情况下, ASP值均接近于0, 可以认为这3个网络出度-出度和入度-入度既没有同配倾向性, 也没有异配倾向性。

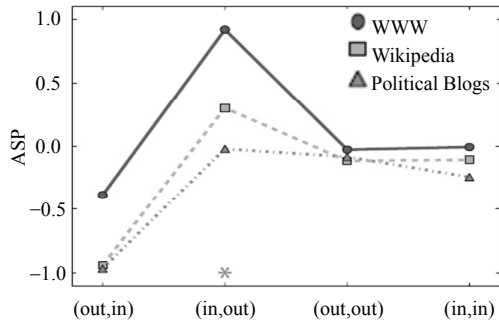


图24 3个有向网络的不同的度匹配性质<sup>[56]</sup>

## 9 结束语

综上所述, 本文分别研究了无权无向网络和无权有向网络的零模型构造及应用。对于无权无向网络, 重点总结了0阶ER随机图, 1~3阶配置模型, 以及基于0~3阶断边重连零模型的构造过程和它们的实际应用。其中, 断边重连零模型具有非常广泛的应用, 如可以帮助检测网络的随机性强弱, 判断网络是否具有富人俱乐部特性等。本文首次将高阶零模型扩展到社团检测等中尺度特性分析中, 构造了4种依赖于社团结构的零模型: 社团内部断边重连零模型、社团间断边重连零模型、增强社团结构零模型和减弱社团结构零模型。对于无权有向网络, 重点总结了基于随机断边重连的1阶零模型, 将原始网络保持度分布不变的情况下断边重连, 该模型适用于检测有向网络中节点与节点之间的出入度匹配特性。

本文主要研究网络零模型的构造方法及相关应用, 最大特点在于将理论和应用两方面进行了有机结合。在基于零模型来研究网络科学领域的各种实际应用问题时, 其基本流程和研究网络不平凡性质的过程是一致的, 区别在于不再单独依赖于网络统计量和网络模型。

传统研究中, 网络统计量虽然可以从一个侧面去刻画网络性质, 但有时无法和实际应用场景对应起来。基于零模型的方法为跨越网络统计量对网络

直接进行分析提供了可能性, 思路是将实证网络和它对应的一系列零模型网络当成黑箱, 然后绕过计算各种统计量这些步骤, 直接比较实证网络和零模型网络在某种实际场景(比如说链路预测)上的性能差异。如通过比较实证网络和它对应零模型网络使用同一链路预测方法的性能差异, 不但可以判断出该链路预测算法的准确性, 还能分析出该网络的哪些性质影响了链路预测的效果, 从而提出更有效的预测算法或检测模型。

相比于为网络建立各种模型的方法, 参照网络零模型的各种研究, 其最大价值是提供一套网络科学研究的新范式。这种严谨的科学研究方法有助于网络科学的研究者建立一套统一、完整而科学的研究方法, 让该领域研究者更加严格和科学的审视实证结果。

目前零模型中应用最广泛的是根据原始网络中度序列产生的1阶零模型, 约束条件的较为苛刻的2阶及2.5阶零模型的还未得到广泛的应用。导致这一情况有两方面原因: 一方面是高阶零模型的构造比1阶模型要复杂, 另一方面是很多研究者不清楚在不同阶数零模型的使用情境。对于实际网络的2阶和2.5阶零模型网络的具体实现, 文献[1-4]分别提供了几种开源实现的方法, 这几种方法可以供研究者使用和借鉴, 可以部分解决第一个问题。对于何种情况下使用何种零模型网络这一问题, 一般而言应该遵循由低阶到高阶的顺序来使用, 越是高阶的零模型, 它的性质越接近于原始网络。如通过与相应的2阶零模型对比可以清楚地发现, 2.5阶零模型的各种性质比2阶零模型更接近于原始网络。

网络性质可以分成微观、中观和宏观三个层面。微观性质主要一般是指平均度、度分布、匹配系数和聚类系数涉及到单个或少量几个节点(小于4个)之间的性质。中观性质则是指模体、模块度、派系、富人俱乐部等涉及到不少于4个以及一组节点(如社团)之间的性质。宏观性质是指平均最短路径长度、网络直径或半径、介数、接近中心性等涉及到网络所有节点的性质。目前在由低阶到高阶生成网络零模型的过程中, 这三种性质之间的相互影响我们还不得而知。因此, 今后应进一步探索在分别保证网络微观、中观和宏观特性的前提下, 对网络中的各种性质进行定性和定量检测, 研究不同尺度零模型之间的统一性和相互影响规律, 从而揭示出复杂网络多尺度的复杂性与随机性。

本文研究工作得到大连市青年科技之星项目支持计划(LR2016070)、贵州省公共大数据重点实验室开放课题基金(2017BDKFJJO)的资助,在此表示感谢。

### 参 考 文 献

- [1] GJOKA M, KURANT M, MARKOPOULOU A. 2.5K-graphs: from sampling to generation[C]//IEEE INFOCOM. New York: IEEE, 2013: 1968-1976.
- [2] MAHADEVAN P, HUBBLE C, KRIOUKOV D V, et al. Orbis: Rescaling degree correlations to generate annotated internet topologies[J]. ACM Special Interest Group on Data Communication, 2007, 37(4): 325-336.
- [3] MAHADEVAN P, KRIOUKOV D V, FALL K R, et al. Systematic topology analysis and generation using degree correlations[J]. ACM Special Interest Group on Data Communication, 2006, 36(4): 135-146.
- [4] ORSINI C, DANKULOV M M, et al. Quantifying randomness in real networks[J]. Nature Communications, 2015, 6: 8627.
- [5] BOLLOBAS B. Random graphs[M]. England: Cambridge University Press, 2001.
- [6] NEWMAN M E J. The structure and function of complex networks[J]. SIAM Review, 2003, 45(2): 167-256.
- [7] BARABÁSI A, ALBERT R. Statistical mechanics of complex networks[J]. Reviews of Modern Physics, 2002, 74(1): 47-97.
- [8] MOLLOY M, REED B. The size of the giant component of a random graph with a given degree sequence[J]. Combinatorics, Probability and Computing, 1998, 7(3): 295-305.
- [9] 尚可可, 许小可. 基于置乱算法的复杂网络零模型构造及其应用[J]. 电子科技大学学报, 2014, 43(1): 7-20.  
SHANG Ke-ke, XU Xiao-ke. Construction and application for null models of complex networks based on randomized algorithms[J]. Journal of University of Electronic Science and Technology of China, 2014, 43(1): 7-20.
- [10] 尚可可. 在线社交网络的零模型构造和行为预测研究[D]. 青岛: 青岛理工大学, 2013.  
SHANG Ke-ke, The study of null model construction and user behavior prediction for online social networks[D]. Qingdao: Qingdao Technological University, 2013.
- [11] MOLLOY M, REED B. A critical point for random graphs with a given degree sequence[J]. Random Structures and Algorithms, 1995, 6(2-3): 161-180.
- [12] NEWMAN M E J, SRROGATZ S H, WATTS D J. Random graphs with arbitrary degree distributions and their applications[J]. Physical Review E, 2001, 64(2): 26118.
- [13] WATTS D J, STROGATZ S H. Collective dynamics of 'small-world' networks[J]. Nature, 1998, 393(6684): 440-442.
- [14] BARABASI A L, ALBERT R. Emergence of scaling in random networks[J]. Science, 1999, 286(5439): 509-512.
- [15] 明均仁, 党永杰. 基于科研贡献度加权的作者合作网络对比研究[J]. 数字图书馆论坛, 2016(1): 34-40.  
MING Jun-ren, DANG Yong-jie. Comparative study of author collaboration network based on scientific contribution degree[J]. Digital Library Forum, 2016(1): 34-40.
- [16] 李慧嘉, 李慧颖, 李爱华. 多尺度的社团结构稳定性分析[J]. 计算机学报, 2015, 38(2): 301-312.  
LI Hui-jia, LI Hui-ying, LI Ai-hua. Analysis of multi-scale stability in community structure[J]. Chinese Journal of Computers, 2015, 38(2): 301-312.
- [17] 王林, 戴冠中. 复杂网络的度分布研究[J]. 西北工业大学学报, 2006, 24(4): 405-409.  
WANG Lin, DAI Guan-zhong. On degree distribution of complex network[J]. Journal of Northwestern Polytechnical University, 2006, 24(4): 405-409.
- [18] 陈庆华. 无标度网络的相关性和联合度分布[J]. 福建师大学报(自然科学版), 2006, 22(1): 1-6.  
CHEN Qing-hua. Correlations and joint degree distributions of scale-free networks[J]. Journal of Fujian Normal University (Natural Science Edition), 2006, 22(1): 1-6.
- [19] NEWMAN M E J. Clustering and preferential attachment in growing networks[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2001, 64(2): 025102.
- [20] NEWMAN M E J. Random graphs with clustering[J]. Physical Review Letters, 2009, 103(5): 58701.
- [21] NEWMAN M E J, KARRER B. Random graphs containing arbitrary distributions of subgraphs[J]. Physical Review E, 2010, 82(6): 66118.
- [22] NEWMAN M E J. Networks: an introduction[M]. 郭世泽, 陈哲, 译. 北京: 电子工业出版社, 2010.  
NEWMAN M E J. Networks: an introduction[M]. Translated by GUO Shi-ze, CHEN Zhe. Beijing: Publishing House of Electronics Industry, 2010.
- [23] HOLME P. Network reachability of real-world contact sequences[J]. Physical Review E, 2005, 71(4): 46119.
- [24] MASLOV S, SNEPPEN K. Specificity and stability in topology of protein networks[J]. Science, 2002, 296(5569): 910-913.
- [25] 李欢, 卢罡, 郭俊霞. 复杂网络零模型的量化评估[J]. 计算机应用, 2015, 35(6): 1560-1563.  
LI Huan, LU Gang, GUO Jun-xia. Quantitative evaluation for null models of complex networks[J]. Journal of Computer Applications, 2015, 35(6): 1560-1563
- [26] 曾进群, 杨建梅, 陈泉, 等. 基于零模型的开源社区大众生产合作网络结构研究[J]. 华南理工大学学报(社会科学版), 2013(2): 29-34.  
ZENG Jin-qun, YANG Jian-mei, CHEN Quan, et al. Research of the open source community cooperation network structure of the mass production based on null model[J]. Journal of South China University of Technology (Social Science Edition), 2013(2): 29-34.
- [27] 陈泉, 杨建梅, 曾进群. 零模型及其在复杂网络研究中的应用[J]. 复杂系统与复杂性科学, 2013(1): 8-17.  
CHEN Quan, YANG Jian-mei, ZENG Jin-qun. Null model and its application in the reserch of complex nexworks[J]. Complex Systems and Complexity Science, 2013(1): 8-17.
- [28] MILO R, SHEN-ORR S, ITZKOVITZ S, et al. Network motifs: Simple building blocks of complex networks[J]. Science, 2002, 298(5594): 824-827.
- [29] SHEN-ORR S, MILO R, MANGAN S, et al. Network

- motifs in the transcriptional regulation network of *Escherichia coli*[J]. *Nature Genetics*, 2002, 31(1): 64-68.
- [30] MILO R, ITZKOVITZ S, KASHTAN N, et al. Superfamilies of evolved and designed networks[J]. *Science*, 2004, 303(5663): 1538-1542.
- [31] MASLOV S, SNEPPEN K, ZALIZNVAK A. Detection of topological patterns in complex networks: correlation profile of the internet[J]. *Physica A: Statistical Mechanics and its Applications*, 2004, 333: 529-540.
- [32] ZHOU S, MONDRAGON R J. The rich-club phenomenon in the internet topology[J]. *IEEE Communications Letters*, 2004, 8(3): 180-182.
- [33] COLIZZA V, FLAMMINI A, SERRANO M A, et al. Detecting rich-club ordering in complex networks[J]. *Nature Physics*, 2006, 2(2): 110-115.
- [34] AMARAL N L A, GUIMERA R. Complex networks: Lies, damned lies and statistics[J]. *Nature Physics*, 2006, 2(2): 75-76.
- [35] ZHOU S, MONDRAGON R J. Structural constraints in complex networks[J]. *New Journal of Physics*, 2007, 9(6): 173.
- [36] XU Xiao-ke, ZHANG J, SMALL M. Rich-club connectivity dominates assortativity and transitivity of complex networks[J]. *Physical Review E*, 2010, 82(4): 46117.
- [37] NEWMAN M E J. Assortative mixing in networks[J]. *Physical Review Letters*, 2002, 89(20): 208701.
- [38] NEWMAN M E J. Mixing patterns in networks[J]. *Physical Review E*, 2003, 67(2): 26126.
- [39] XU Xiao-ke, ZHANG J, SUN J, et al. Revising the simple measures of assortativity in complex networks[J]. *Physical Review E*, 2009, 80(5): 56106.
- [40] NEWMAN M E J. Detecting community structure in networks[J]. *The European Physical Journal B-Condensed Matter and Complex Systems*, 2004, 38(2): 321-330.
- [41] DANON L, DIZA-GUILERA A, DUCH J, et al. Comparing community structure identification[J]. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(9): P09008.
- [42] FORTUNATO S. Community detection in graphs[J]. *Physics Reports*, 2010, 486(3-5): 75-174.
- [43] CAI Q, MA L J, GONG M G, et al. A survey on network community detection based on evolutionary computation. [J]. *Int J Bio-Inspired Computation*, 2016, 8(2): 84-98.
- [44] FLAKE G W, LAWRENCE S, GILES C L, et al. Self-organization and identification of web communities[J]. *Computer*, 2002, 35(3): 66-70.
- [45] TAYLOR D, SHAI S, STANLEY N, et al. Enhanced detectability of community structure in multilayer networks through layer aggregation[J]. *Physical review letters*, 2016, 116(22): 228301.
- [46] BAZZI M, PORTER M A, WILLIAMS S, et al. Community detection in temporal multilayer networks, with an application to correlation networks[J]. *Multiscale Modeling & Simulation*, 2016, 14(1): 1-41.
- [47] THAKUR S, DHIMAN M, TELL G, et al. A review on protein-protein interaction network of APE1/Ref-1 and its associated biological functions[J]. *Cell Biochemistry & Function*, 2015, 33(3): 101-112.
- [48] 李晓佳, 张鹏, 狄增如, 等. 复杂网络中的社团结构[J]. *复杂系统与复杂性科学*, 2008, 5(3): 19-42.  
LI Xiao-jia, ZHANG Peng, DI Zeng-ru, et al. Community structure in complex networks[J]. *Complex Systems And Complexity Science*, 2008, 5(3): 19-42.
- [49] CUI W K, SHANG K K, ZHANG Y J, et al. Constructing null networks for community detection in complex networks[J]. *European Physical Journal B*, 2018, 91(7): 145.
- [50] LENG Z F. Community detection in complex networks based on greedy optimization[J]. *Acta Electronica Sinica*, 2014, 42(4): 723-728.
- [51] 刘亚冰. 复杂网络中的社团结构特性研究[D]. 上海: 上海交通大学, 2010.  
LIU Ya-bing. Analysis of community structure in complex networks[D]. Shanghai: Shanghai Jiao Tong University, 2010.
- [52] 唐明, 崔爱香, 袁凯. 关注耦合网络及其传播动力学研究[J]. *复杂系统与复杂性科学*, 2011, 8(2): 87-91.  
TANG Ming, CUI Ai-xiang, XI Kai. On spreading dynamics on coupled networks[J]. *Complex Systems and Complexity Science*, 2011, 8(2): 87-91.
- [53] 汪小帆, 李翔, 陈关荣. 复杂网络引论——模型、结构与动力学[M]. 北京: 高等教育出版社, 2012: 217-218.  
WANG Xiao-fan, LI Xiang, CHEN Guan-rong. Introduction to complex networks: models, structures and dynamics[M]. Beijing: Higher Education Press, 2012: 217-218.
- [54] SHANG Ke-ke, YAN W S, SMALL M. Evolving networks-using past structure to predict the future[J]. *Physica A: Statistical Mechanics and Its Applications*, 2016, 455: 120-35.
- [55] SHANG Ke-ke, SMALL M, XU X K, et al. The role of direct links for link prediction in evolving networks[J]. *EPL (Europhysics Letters)*, 2017, 117(2): 28002.
- [56] FOETER J G, FOSTER D V, GRASSBERGER P. Edge direction and the structure of networks[J]. *Proceedings of the National Academy of Sciences*, 2010, 107(24): 10815-10820.