

一种新型二分网络类局域世界演化模型*

田立新[†] 贺莹环 黄益

(江苏大学理学院, 镇江 212013)

(2011年9月21日收到; 2012年6月7日收到修改稿)

现实世界中复杂网络的演化存在很明显的局域选择现象, 然而目前关于二分网络中的局域世界演化模型研究较少. 因此, 本文建立了一个基于二分网络的类局域世界演化模型. 首先定义了网络节点度值的饱和度. 在此基础上提出了一种新型二分网络局域世界演化模型. 新节点加入系统不需要全局知识, 而是通过节点在网络演化的不同时刻度值饱和度为选择条件构造新节点的局域世界, 然后利用择优连接从局域世界中选择节点增加连边完成网络演化. 此类模型中新节点的局域世界是通过节点饱和度的限制被动生成, 因此又称为类局域世界模型. 通过模拟分析发现在节点度值饱和度的限制下择优连接并没有产生具有幂率特性的度分布, 而是生成了度分布相对均匀的二分网络, 即节点度值分布区间较小. 此外, 本文还给出了该网络的混合系数计算结果, 该结果显示网络同配性与网络参数的选择有关, 这一结果与网络邻点平均度的模拟结果一致.

关键词: 二分网络, 度分布, 节点度值饱和度, 局域世界

PACS: 89.75.-k, 89.75.Fb

1 引言

自然界和人类社会中存在各种各样的复杂系统. 我们被网络包围着, 几乎所有的复杂系统都可以抽象成网络模型, 这些网络往往有着大量的节点, 节点之间有着复杂的连接关系. 随着复杂网络的小世界效应^[1]及无标度性^[2]的发现, 复杂网络研究逐渐成为多个学科共同关注的前沿热点. 其中网络演化是当今网络科学研究的一个重要领域. 通过网络演化构建网络结构是对网络进行定量研究的基础, 这对网络的运行管理、大规模网络行为的理解等方面都是具有重要的意义. 同样, 研究和建立能够同时重现真实网络多种拓扑性质特征的网络局域演化模型, 有助于更好地理解 and 描述真实网络, 对于更进一步了解复杂网络的拓扑结构及其形成机理、研究复杂网络传播动力学和网络的鲁棒性都具有重要意义.

BA 模型是第一个无标度网络演化模型, 它捕

捉了无标度网络形成的两个必不可少的生成机理即增长性和择优选择, 说明了大规模复杂网络自组织成为无标度状态的原因. 在 BA 模型中, 每当有一个新节点进入网络时, 它会从所有的已存在节点中选择, 即它的“择优”是基于全局信息的. 这种假设条件在现实中是不常见的, 一般来说, 每一个新进入网络的节点只能获知一定范围内的信息, 它没有能力去完全了解整个网络的情况, 因而也只占有和使用整个网络的局部信息, 在局部范围内进行择优, 或者以其他原则在局域范围内选择连接对象. 例如研究表明世界贸易网, 每个节点代表一个国家, 两个国家之间有贸易关系, 则相应两个节点之间存在连接边. 许多国家都致力于加强与各自区域经济合作组织内部的国家之间的经济合作和贸易往来. 因此在世界贸易网中, 优先连接机理是存在于某些区域经济体内部的. 类似地, 在 Internet 中, 计算机网络是基于域-路由器的结构来组织管理的, 一台主机通常是只与同一域内的其他主机相连, 而路由器则代表它内部域的主机与其他路由器相连. 其中, 优

* 国家自然科学基金 (批准号: 71073072)、国家社会科学基金重大项目 (批准号: 12&ZD062)、江苏大学学生科立项 (批准号: 10A144) 和重庆师范大学青年基金 (批准号: 10XLQ001) 资助的课题.

[†] E-mail: tianlx@ujs.edu.cn

先连接机理不是对整个网络,而是在每个节点各自的局域世界中有效.所有这些实例说明在诸多实际的复杂网络中存在着局域世界.因此在不完全信息情况下,确定网络的连接机理,构造其演化模型,是十分有意义的. Li 与 Chen 于 2003 建立的局域世界演化网络模型(李陈模型)^[3]对 BA 模型做了改进,在整个网络中随机的选取 M 个点作为新加入的节点的局域世界,新加入的节点以择优的概率在 M 个节点中选择连接对象.这样形成的网络其度分布随 M 的变化在指数分布到幂率分布之间演化.但由于 M 个节点的随机性选择,使得选出的 M 各节点可能毫无关联,例如公交网络,如果要加入一个新的站点,按照李陈模型构造局域世界的方法构造,选出 M 个站点,如果选出的站点间没有合作运营关系并且与新站点相距甚远,那么新站点与其构成合作运营关系,会受到地理位置,连接成本的限制.因此采取这种构造局域世界的方式偏离了现实网络,为了更好的描述现实,王延庆^[4]对李陈模型加以改进,以某一节点的第一层节点和第二层节点作为新加入节点的局域世界,并给出了计算机模拟的模型度分布.文献[5]通过计算节点间的度相关性,在确定局域世界的大小后选择度相关性大的节点作为新加入节点的局域世界构造了局域世界网络模型分析和数值模拟的结果发现如此的模型度分布遵循幂率分布.文献[6]基于随机性选择确定了新加入节点的局域世界,通过改变局域世界内的连接概率和局域范围建立了 LWD 网络模型,并研究了模型度分布、平均路径长度、聚类系数关于参数的变化.文献[7]研究了局域世界网络的无标度性和社团结构,发现局域演化网络的度分布随着局域规模 M 的大小实现了从指数分布到幂率分布的一个过渡.文献[8]提出了一个基于随机行走和策略选择的复杂网络局域演化模型,该模型最终自组织演化具有幂率特点的复杂网络.文献[9]研究了具有老化节点的加权局域世界网络的演化,并给出了节点的强度分布服从幂率的结论.

然而以上关于局域世界模型的研究都是建立在单点网络上的,对于广泛存在的二分网络并未深入研究.事实上在二分网络的演化中也存在在局域范围内选择连接的现象.例如能源供需网络,无论供应节点还是需求节点,在选择合作的对象时,总是受地理因素、运输的便利性等因素的影响而不能在全局范围内选择连接对象,只能是在相对比

较有利的范围内选择.再例如在生产商和批发商所组成的二分网络中,以批发商的选择为例,在选择合作的生产商时,会以地域和产品质量为考虑因素进而把合作范围确定在一定的范围内.二分网络在现实中是普遍存在的,例如之前提到的能源供需网络,演员-电影网^[1]、生产商-供应商网^[10]等.此外文献[11, 12]的研究表明任何的复杂网络都可看成是具有某特定特点的二分网络.因此,对于二分网络局域世界演化模型的研究是很有必要的.

在局域世界大小确定后,范围内节点的选择大多是随机选择的.事实上,在网络节点的选择中是双向的选择,即新加入的节点以某种条件进行选择,同时在实际中有被选择资格的节点也根据自己的因素(比如节点增加新边的成本)决定是否要加入新节点的选择范围成为候选节点.例如国际航空网,出于对航线的经济效益的考虑,会尽量减少多条航线连接到同一个飞机场,如果考虑所有的航线布局,由于受到时间和空间的限制,网络中不可能出现拥有较多航线的飞机场.每个飞机场所拥有的航线数不会相差非常大.而新建的飞机场在考虑空间距离的基础上自然只能在与有意开辟新航线的机场中选择进而建立关系.随着航线网络的演化,一些机场为了扩大规模增加盈利,可能会在下一个时间步会追加投资开辟新航线进而增大机场的盈利,那么这些机场在这个时刻就成为了新节点局域世界内的候选节点.

本文引进节点饱和度的概念来作为已经加入网络的节点是否需要扩大连边的衡量标准,即通过计算网络演化的每一个时间步每一个旧节点的饱和度来确定节点是否还可以接受新的连边,即是否要加入新节点的局域世界成为候选节点以增加连边.通过节点在网络演化的不同时刻饱和度变化来确定新加入节点的动态局域世界.在该模型中,不同类节点集合中新加入的节点,只能在对应节点集中从度值小于其饱和度的节点子集中以择优的概率完成连边.在此需要说明的是节点饱和度的定义中涉及了网络中不同节点集的平均度这一全局信息,但是通过节点饱和度的限制却使新加入系统的节点仅能从度值小于其饱和度的子集中择优选择连接,该子集就是新节点的局域世界.这样的局域世界并不是由新节点来决定的,而是通过旧节点根据自身承载能力加以判断后被动生成的新节点的局域世界,因此我们又称此模型为二分网络类局域

世界的演化模型. 对于满足此类演化规则的二分网络我们直接基于原始的二分网络进行研究.

2 节点饱和度的定义

我们称

$$k_{i(\max)}(t) = \langle k(t) \rangle + \frac{\langle k(t) \rangle}{k_i(t)}, \quad (1)$$

为节点 i 在网络演化 t 时刻的饱和度, 其中 $\langle k(t) \rangle$ 表示在 t 时刻网络节点的平均度值, $k_i(t)$ 表示 t 时刻节点 i 的度值.

我们规定若节点 i 在 t 时刻的度值 $k_i(t)$ 大于或等于其饱和度 $k_{i(\max)}(t)$, 则该节点在该时刻不再连接新的边, 即节点在该时刻达到了度饱和, 不会加入局域世界成为新加入节点连接的候选节点. 反之若节点 i 在 t 时刻的度值小于其饱和度, 则该节点在该时刻可以接受新的连边, 即是局域世界范围内的候选节点. 通过 (1) 式的定义可知, 若 t 时刻节点 i 的度值比较小, 则 $\langle k(t) \rangle + \frac{\langle k(t) \rangle}{k_i(t)}$ 的值会比较大, 这样节点 i 在 t 时刻就不容易达到饱和, 即 i 点可能连接更多的边来增大本身的度值, 反之若 t 时刻节点 i 的度值比较大, 则 $\langle k(t) \rangle + \frac{\langle k(t) \rangle}{k_i(t)}$ 的值比较小, 即该节点在 t 时刻容易达到或已经达到饱和. 这样的限制会降低网络演化过程中节点度值的差异, 使得生成的网络度值分布比较均匀.

在实际网络中, 节点增加度值是受各种因素的影响和制约. 例如以企业为节点, 相互之间有合作贸易关系则建立连边, 以此所构成的网络中, 节点增加新的连边就有可能意味着企业业务的扩大. 每个企业由于受自身实力的限制在某一阶段内只能有一定的合作对象, 但是随着网络的演化更多更强大的竞争对手就会出现, 为了在市场中不被淘汰或者获取更大的利润, 企业会扩大自身实力以争取到更多的合作对象. 在企业决定是否需要追加投资扩大规模时, 会以市场的平均水平为标准进行考量, 因此本文所采取节点饱和度的定义方式作为判定节点是否要加入局域世界作为候选节点是合理的.

3 二分网络局域世界演化模型

我们的二分网络是由 $H = (M, N, E)$ 组成的, 其中 M, N 分别是两类节点集合, E 是存在于这两

类节点之间的边. 初始的二分网络是由 M 中的 m_0 个节点和 N 中的 n_0 个节点以及存在于 M 和 N 之间的 $a (\subset E)$ 条边组成. 其中 a 是在 M 中的 m_0 个节点和 N 中的 n_0 个节点之间随机连接的, 直至集合 M 和 N 中的任意节点的度值至少为 1, 则完成初始网络的演化. 我们记集合 M 中的节点 i 在 t 时刻的度值为 $k_i(t)$, 集合 N 中的节点 j 在 t 时刻的度值为 $J_j(t)$. 在之后的每一个时间步里循环执行以下过程,

1) 记集合 M 中度值满足 $k_i(t) < \left(\langle k(t) \rangle + \frac{\langle k(t) \rangle}{k_i(t)} \right)$ 的节点集为 $\Gamma_{M(t)} \subseteq M$, 集合 N 中度值满足 $J_j(t) < \left(\langle J(t) \rangle + \frac{\langle J(t) \rangle}{J_j(t)} \right)$ 的节点集为 $\Gamma_{N(t)} \subseteq N$ (其中 $\langle k(t) \rangle, \langle J(t) \rangle$ 分别为集合 M 和集合 N 的平均度值).

2) 集合 M 中加入一个新点, 加入的新点以择优的概率 $\frac{J_j}{\sum_{j \in \Gamma_{N(t)}} J_j}$ 与集合 $\Gamma_{N(t)}$ 中的 n 个该时刻以前加入的旧点相连接.

3) 集合 N 中加入一个新点, 加入的新点以择优的概率 $\frac{k_i}{\sum_{i \in \Gamma_{M(t)}} k_i}$ 与集合 $\Gamma_{M(t)}$ 中的 m 个该时刻以前加入的旧节点相连接.

给定网络的演化步数即可生成二分网络局域世界演化模型.

在该网络模型的演化中, 我们通过节点的饱和度和度分别给定了二分网络不同节点集合中新节点的局域世界, 集合中新加入的节点分别在各自的局域世界中以择优的概率选择连接对象. 此类局域世界不是由新节点决定的, 而是根据以前节点的饱和度限制生成的, 是一类新型局域世界. 由于网络演化使得节点的度值发生变化故而在不同时刻每个节点的饱和度也是不同的, 因此在不同时刻新加入节点的局域世界也是不一样的. 在现实世界中是存在满足此类演化规则的实际网络. 例如生产商与批发商所组成的合作网络, 新加入网络的批发商在选择合作对象时会选择规模大信誉好 (即以择优的概率选择) 的生产商合作, 同样生产商也会通过考虑市场的供需关系以及平均水平来决定是否要扩大生产开拓合作范围, 即是否要加入新节点的候选范围内. 如此新加入系统的节点只能在有需要增加连边扩大合作范围的节点中择优的选择连接.

4 二分网络模型的统计特性

对于二分网络统计特性的研究有两种思路: 第一种是把二分网络投影到单顶点网络, 然后以单顶点网络的研究方法进行网络分析, 第二种是直接基于原始二分网络来分析^[13]. 我们采用第二种方法对二分网络局域世界模型的度分布和同配性质进行模拟分析. 网络的模拟建立在网络演化 $t = 5000$ 步时形成的网络模型.

4.1 网络模型的度分布

网络中节点的度分布情况可用分布函数 $p(k)$ 来描述. $p(k)$ 表示的是网络中度数为 k 的节点个数占节点总数的比例. 在二分网络我们分别考虑了节点集合 M 和 N 的度分布函数 $p(k)$ 和 $q(J)$. 从图 1 中两个节点集合的度分布图可以体现出节点饱和度和所确定的局域世界抑制了择优连接所产生的节点度值分配不均匀现象. 网络两类节点的度值均匀, 没有出现 hub 点.

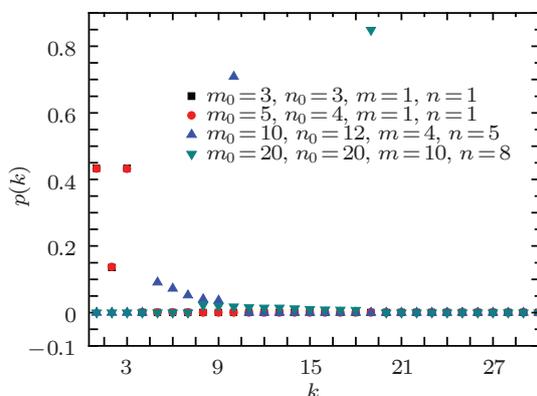


图 1 二分网络模型节点集 M 的度分布图

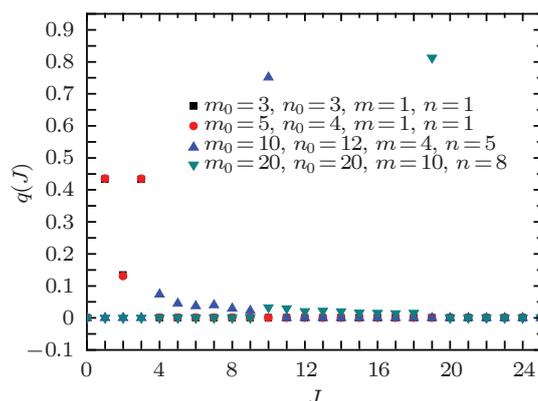


图 2 二分网络模型节点集 N 的度分布图

4.2 度相关性

度相关性描述的是网络中不同节点之间的连接关系. 如果高度值节点倾向于连接其他高度值的节点, 称为同配网络. 如果度高的节点倾向于连接其他低度值的节点, 称为异配网络. 网络的同配性 (异配性) 影响网络的结构和行为. 关于网络同配性的刻画有诸多不同的量, 本文利用网络混合系数^[14]的计算方法分别解析计算了该网络不同节点集合的混合系数来判断网络的同配性.

在网络完成演化以后我们先来计算节点集 M 的混合系数:

我们记 $k = \{k_1, k_2, \dots, k_{m_0+t}\}$ 为网络节点集 M 的度序列, 计算 M 的节点平均度

$$\bar{k} = \sum_{i=1}^{m_0+t} \frac{k_i}{m_0+t}.$$

将节点集 M 中的节点分为两个集合: $A_M = \{i | k_i > \bar{k}\}$, $B_M = \{i | k_i < \bar{k}\}$, 令 $O_i = \{j | (i, j) \in E, i \in M, j \in N\}$ 表示节点 i 的邻节点集合, 定义函数

$$a_j = \begin{cases} 1, & J_j > \bar{k}, \\ 0, & J_j = \bar{k}, \\ -1, & J_j < \bar{k}, \end{cases} \quad \forall i \in A_M, j \in O_i, I_i = \sum_{j \in O_i} a_j/k_i, \quad (2)$$

$$b_j = \begin{cases} -1, & J_j > \bar{k}, \\ 0, & J_j = \bar{k}, \\ 1, & J_j < \bar{k}. \end{cases} \quad \forall i \in B_M, j \in O_i, I_i = \sum_{j \in O_i} b_j/k_i. \quad (3)$$

计算 $I_M(g) = (\sum_{i \in A_M} I_i + \sum_{i \in B_M} I_i) / (m_0 + t)$, $I_M(g) \in [-1, 1]$ 即为网络节点集 M 的混合系数, 若 $I_M(g)$ 为负值则说明 M 是异配的, 即 M 中的高度值节点倾向于与 N 中的低度值节点相连接, 若 $I_M(g)$ 为正值即说明 M 是同配的, 若 $I_M(g)$ 为零即说明 M 是混配的.

利用上述的方法我们同样可以计算网络节点集 N 的混合系数 $I_N(g)$.

在表 1 中我们列出了当网络取不同的参数时所计算出的节点集 M 和 N 的混合系数. 从表 1 我们可以看出由参数 1 和参数 2 演化的网络的混合系数是负数, 即这样的网络是异配的. 由参数 3 和参数 4 演化生成的网络的混合系数是正数, 即这样的网络是同配的. 文献 [14] 中已提出混合系数与网络的规模无关, 而与网络的平均度有关. 从表 1 中我们也论证了这一点: 当节点集的平均度从参数 1 到参数 4 发生变化的时候网络的同配性也随之发生了变化. 以下我们采用邻节点平均度的模拟结果来论证上述关于网络同配性的解析结果.

对于给定度为 k 的节点 $i \in M$, 其邻节点的平均度为

$$k_{nn}(k) = \sum_J Jp(J|k), \quad (4)$$

其中

$$p(J|k) = \frac{\langle k \rangle p(k, J)}{kp(k)}$$

表示 M 中度为 k 的节点与 N 中度为 J 的节点邻接的条件概率, $p(k, J)$ 表示度为 k 和度为 J 的节点邻接的联合概率.

将所有度为 k 的邻节点平均度再取平均, 记为 $\langle k_{nn}(k) \rangle$, 若 $\langle k_{nn}(k) \rangle$ 随着度数 k 的增大而增大, 说明节点集 M 中的高度值节点偏向于与集合 N 中的高度值节点相连接. 反之, 若 $\langle k_{nn}(k) \rangle$ 随着度数 k 的增大而减小, 说明 M 中高度值的节点偏向于与 N 中低度值节点相连接.

同样, 对于给定度为 J 的节点 $j \in N$, 其邻点

平均度

$$J_{nn}(J) = \sum_k kp(k|J), \quad (5)$$

反应的集合 N 中节点的连接情况与集合 M 类似.

通过数值模拟, 图 3 和图 4 表明在网络取参数 1 和 2 时, 两类节点集 M 和 N 均是异配的, 参数 3 和参数 4 所生成的网络的节点集合均是同配的, 这一结果与我们的解析结果一致. 从图 3 和图 4 中我们也可以看出在节点饱和度的限制下所演化生成的类局域世界二分网络模型其节点度值的区间较小, 尤其是在参数 1 和参数 2 演化生成的网络节点的度值仅三个取值.

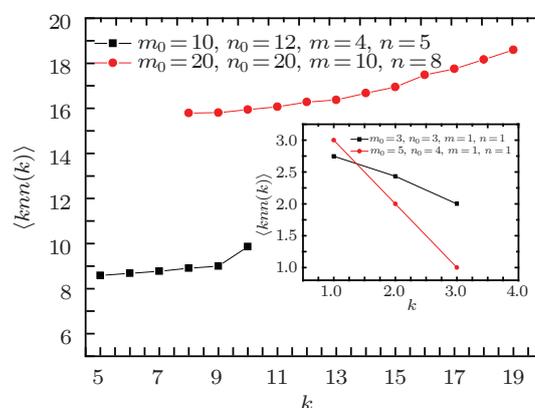


图 3 网络模型节点集 M 的邻点平均度

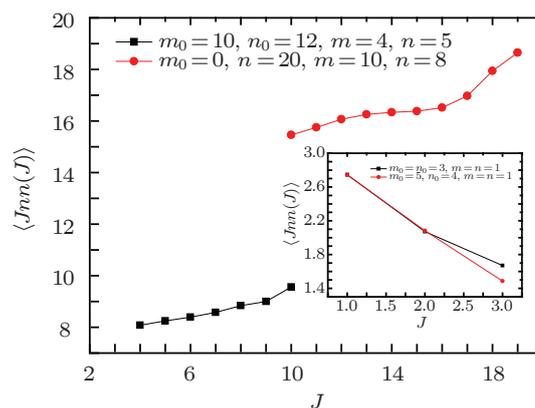


图 4 网络模型节点集 N 的邻点平均度

表 1 不同参数生成的网络混合系数对比表

参数	参数 1		参数 2		参数 3		参数 4	
	M	N	M	N	M	N	M	N
混合系数	-0.1909	-0.1912	-0.1911	-0.1929	0.5518	0.5645	0.7472	0.7337
平均度	1.996	1.996	1.9992	1.996	8.98503	8.98144	17.934	17.934

5 结论

在实际的网络中, 节点之间的连边是一种双向的选择, 不仅新加入系统的节点以某种原则选择与旧节点相连接, 同时已加入系统的节点也根据自身的条件决定是否增加新的连边. 本文引入了节点饱和度的概念, 以节点饱和度为标准判断节点是否需要加入新节点的局域世界. 在此基础上建立了以择优连接为原则的一种新型二分网络局域世界模型. 这样的局域世界不是由新节点了解的局部信息决定产生的, 而是由旧节点根据其自身的承载能

力判断以后被动生成的, 因此又称该模型为类局域世界模型. 对于生成的网络模型直接基于原始二分网络方法进行分析, 通过仿真模拟发现此类网络的节点度分布并没有因择优连接而呈现幂率特性, 相反显现出度值分布区间相对较小的性质. 此外, 我们还解析的计算了网络的混合系数, 发现在网络取不同的生成参数时其同配性是会发生变化的. 这一解析结果与网络不同集合的邻点平均度的解析结果是一致的. 该结果对于具有局域特性且度分布区间小的二分网络具有参考价值.

-
- [1] Watts D J, Strogatz S H 1998 *Nature* **393** 440
- [2] Barabasi A L, Albert R 1999 *Science* **286** 509
- [3] Li X, Chen G R 2003 *Physica A* **328** 274
- [4] Wang Y Q 2008 *Master Dissertation* (Nanjing: Nanjing University of Posts and Telecommunications) (in Chinese) [王延庆 2008 硕士学位论文 (南京: 南京邮电大学)]
- [5] Xuan Q, Li Y J, Wu T J 2007 *Physica A* **378** 561
- [6] Gu Y Y, Sun J T 2008 *Physics Letters A* **372** 4564
- [7] Shuhei Furuya, Kousuke Yakubo 2010 *Physica A* **389** 5878
- [8] Jiang Z H, Wang H, Gao C 2011 *Acta Phys. Sin.* **60** 05893 (in Chinese) [姜志宏, 王晖, 高超 2011 物理学报 **60** 058903]
- [9] Wei G H, Duan Z S, Chen G R, Geng X M 2011 *Physica A* **390** 4012
- [10] Wang D H, Zhou L, Di Z R 2005 *Physica A* **363** 359
- [11] Jean-Loup Guillaume, Matthieu Latapy 2006 *Physica A* **371** 795
- [12] Jean-Loup Guillaume, Matthieu Latapy 2004 *Information Processing Letters* **90** 215
- [13] Wu Y J, Zhang P, Di Z R, Fan Y 2010 *Complex System and Complexity Science* **7** (1) (in Chinese) [吴亚晶, 张鹏, 狄增如, 樊瑛 2010 复杂系统与复杂性科学 **7** (1)]
- [14] Shi D H 2011 *Theory of Network Degree Distributions* (Beijing: Higher Education Press) p160 (in Chinese) [史定华 2011 网络度分布理论 (北京: 高等教育出版社) 第 160 页]

A novel local-world-like evolving bipartite network model*

Tian Li-Xin[†] He Ying-Huan Huang Yi

(Faculty of Science, Jiangsu University, Zhenjiang 212013, China)

(Received 21 September 2011; revised manuscript received 7 June 2012)

Abstract

In complex networks, node degree values are limited by some practical factors. The saturation of node degree, which is a function of network evolution time, is defined first. We propose a novel evolving bipartite network model based on preferential attachment in local-world, which is generated by node saturation restrictions, not new node selection. So we also call it local-world-like model. However, the numerical simulation results display that the degree distribution does not obey the power-law distribution. We find that the degree value interval of this local-world-like bipartite network is small. There is no hub node. In addition to these, we analyze mixing coefficient of the network and find that the assortativities of the network are different when the network is generated by different initial parameters. Such a result accords with our simulated result.

Keywords: bipartite network, degree distribution, node degree saturation, local-world

PACS: 89.75.-k, 89.75.Fb

* Project supported by the National Nature Science Foundation of China (Grant No.71073072), the National Social Science Fund Project (Grant No. 12&ZD062), the Students Research Foundation of Jiangsu University (Grant No. 10A144), and the Youth Foundation of Chongqing Normal University (Grant No. 10XLQ001).

[†] E-mail: tianlx@ujs.edu.cn