

基于BA增长网络的教师合作竞争网络模型

宫响¹ , 刘慧明² , 张淑华³

1. 青岛科技大学数理学院, 青岛 266061
E-mail: xiang_qd@163.com

2. 青岛科技大学网络管理中心, 青岛 266061
E-mail: huiming@qust.edu.cn

3. 青岛科技大学教务处, 青岛 266061
E-mail: shzhang@qust.edu.cn

摘要: 基于BA增长网络的形成机理, 采用二分图方法构建了教师合作竞争网络, 提出了一个平稳增长的教师合作竞争网络模型, 并对节点赋予了竞争力。用平均场方法计算了顶点项目度分布及度分布, 证明了该网络为无标度网络, 并对其度分布进行了数值模拟, 模拟结果与理论计算值很好的吻合。用动力学方法计算了聚集系数, 统计得出了网络的平均路径长度。分析了网络的拓扑结构对实际的指导作用, 为学校决策者提供了相应的决策信息。

关键词: 复杂网络, 合作竞争网络, 无标度, 小世界效应, BA网络

A Collaboration-competition Network Model for Teachers Based on BA Growing Networks

Xiang Gong¹, Huiming Liu², Shuhua Zhang³

1. Academic of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao 266061

E-mail: xiang_qd@163.com

2. Center of networks, Qingdao University of Science and Technology, Qingdao 266061

E-mail: huiming@qust.edu.cn

3. Office of teachers affairs, Qingdao University of Science and Technology, Qingdao 266061

E-mail: shzhang@qust.edu.cn

Abstract: Based on the development mechanism of BA growing networks, by a bipartite graph a model has been given which aims at a stationary growing social collaboration-competition network about teachers. In this model the nodes are endowed with competition. The degree distribution which obeys power-law form has been proved, and the simulation result is in agreement with the theoretic calculation. Both the clustering coefficient calculated by dynamic equation and the average path length obtained by statistic make this network have a small-world effect. The topological structure can provide some useful information to the policymakers of schools.

Key Words: Complex networks, Collaboration-competition networks, Scale-free, Small-world Effect, BA Networks

1 引言(Introduction)

近年来,复杂网络已成为揭示自然界和人类社会中各种复杂系统结构和行为的重要手段[1,2]。合作网络是一个复杂的社会网络,传统的合作研究,如科学计量学和信息计量学都是通过静态的数据采用统计方法分析合作网络的结构、规律,很少涉及网络的形成机制以及网络拓扑结构在网络研究上的作用。引用复杂网络的概念,可以以全局的、动态的视野,分析合作网络的动力模型、演化规律,对研究知识在网络中的传播途径、确定人才标准、制定科学技术政策有积极的作用。

合作网络与其他类型网络截然不同之处在于:它是由许多完全图组合而成的。这里的完全图是指任何不同两顶点之间都有边相连的简单图。合作竞争网络也是一种合作网络。而自然界中没有完全的合作也没有完全的竞争,是普遍的竞争与合作相结合的。对这样的普遍合作竞争网络已经有研究结果。例如,Heinous等定义了日本的各个公司(参与者)和他们申请的专利(项目)构成的合作-竞争网[3],并且重点研究了二分图向申请公司参与者的投影。数据包括1994-2003年间的4998464项专利。申请同一项专利的公司构成一个项目完全子图,他们当然是互相竞争的。A. M. Chmiel等报道了波兰公司的合作-竞争网络[4]。各个波兰公司定义为参与者节点,它们归属的产业类别(例如建筑施工、建筑材料等)定义为项目。凡是至少归属于同一种产业类别的公司之间连边,表示它们在销售市场上的既竞争、又合作的关系。

高考升学率是我国当今社会的一大热点,是人们关注的焦点,学校更以此作为衡量教师教学业绩的一项重要指标。影响高考升学率的因素有很多,而任课教师间的合作起着至关重要的作用,每位教师都想尽可能提高自己所教科目的成绩,但由于人的精力所限,学生在某一科目上有所偏重,必然会影响其他科目的成绩,所以同一个班级的任课教师之间又存在一定的竞争。由此本文选取某中学1996-2003年考取大学的学生的任课教师作为研究对象,讨论他们之间的合作竞争关系,提出了一个合作竞争网络模型,并对其度分布等统计特征进行了分析,得出了相应的区分不同教师角色的依据。研究结果可以为学校决策者提供一定的决策信息。

本文选取的时间段内高考模式为 $3+2x$,故每个班级的任课教师为包括班主任在内的5名教师。随着高考制度的改革,班级配备教师的人数也在变化,但每个班级的任课教师人数为常数是不变的。而只需在其合作竞争关系的讨论中对模型中选取旧顶点个数和新加入连线数变化,即可适应时代的需求。这也是本文网络的特殊之处,即每次参与合作竞争的人数为常数5。

山东省中青年科学家奖励基金项目(资助号:03BS125)

2 教师合作竞争网络模型(A Collaboration-competition Network Model for Teachers)

BA模型[5]有两个要点:第一,网络的顶点一个一个地增加;第二,新出现顶点与已经存在的顶点之间按照线性优选法则(即所谓“富者更富”法则)建立边的连接。由此模型得到的解析和模拟结果显示理想的幂律。Barabasi和Albert同时解析的说明了如果新出现顶点与已经存在的顶点之间完全随即的选择连接,那么顶点度的分布满足一个指数分布[5]。在此后的许多旨在改进或修正BA模型的建议中,我们特别注意以下两个模型。第一个模型是Bianconi和Barabasi提出的适应度(fitness)模型[6]。该模型是在BA无标度网络的演化过程中考虑竞争因素而提出的。在许多实际网络系统中,节点的度及其增长速度并非只与该节点的年龄有关,而是与节点的内在性质相关的,如个人的交友能力、WWW站点的内容和科研论文的质量等,Bianconi和Barabasi把这一性质成为节点的适应度(fitness)。该模型与BA模型的区别在于,在适应度模型中的优先连接概率与节点的度和适应度之积成正比,而不是仅与节点的度成正比。第二个模型是何大韧等建议的广义合作网络(Generalized Collaboration Networks)模型[7,8]。他们对中药方剂系统、中国旅游线路系统、淮扬菜、北京和扬州的公交线路系统、以及好莱坞演员的合作系统进行了实证统计研究,结果都证明了该模型的合理性。

2.1 教师合作竞争网络的构建(Formation of the Collaboration-competition Network for Teachers)

本文采用二分图(bipartite graph)[8,9,10]方法构建教师合作竞争网络。一类节点 U 可称为“项目”(act,本文中为考取大学的学生),另一类节点 L 可以称为“参与者”(actor,本文中为该生的任课教师)。一条连接两类不同节点的边代表一个参与者参加一个合作项目。如图1(a)所示。如果只考虑一类节点(常常是参与者节点),也称为向一类节点投影,就把两条共用一个另一类节点的边投影为这类节点之间的边,得到单模式图。如图1(b)所示。本文重点讨论任课教师间的合作竞争关系,这时2个参与者节点(教师)之间的边表示他们在教同一个学生中的合作竞争关系。每个考取大学的学生都可以看成5位任课教师互相作用构成的完全图项目,各个考取大学的学生中共同的任课教师可以看作是这些项目之间联系的桥梁,这样就构成了一个教师合作竞争网络,如图2所示。

顶点项目度 h_i 表示一个教师 L_i 曾经教过过的大学生人数,而顶点度 k_i 表示与教师 L_i 有合作竞争关系的教师人数。显然,在向参与者(任课教师)的单模式投影图中,一个考取大学的学生 U_j 的所有任课

教师两两之间都连接一条表示他们合作竞争关系的边,因此构成一个完全图,在这个完全子图中,每个顶点的顶点度都等于 k_i 都等于4。当然,两个(甚至更多个)这样的完全图很可能共用一条甚至好几条边,但如果计算多重边(multiple edges)(即把两个完全图共用的一条边认为是两条重叠的边,依此类推),对每个完全图其顶点度等于4的结论仍然成立。如果把包含多重边,每两个顶点都连边的图仍叫做完全图,则可以说向教师投影的单模式网络仍是完全图的集合。

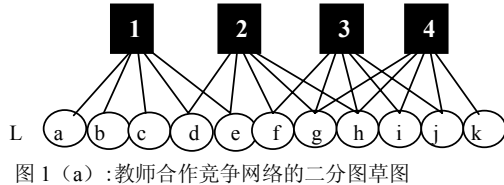


图1(a):教师合作竞争网络的二分图草图

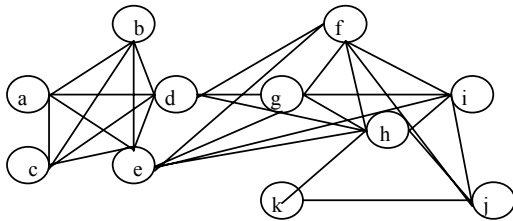


图1(b):图(a)的单模式投影

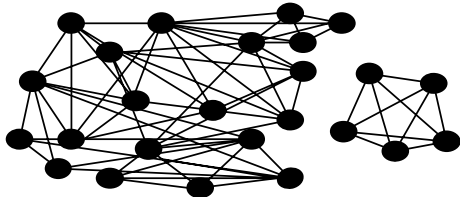


图2:教师合作网络的完全图项目示意图

2.2 教师合作竞争网络模型 (A Collaboration-competition Network Model for Teachers)

鉴于高考的重要性,学校领导选择高三教师会依据很多标准,但其中最重要的应该是教学能力。为此本文结合 BA 网络模型、广义合作网络模型及适应度模型和教师合作竞争的实际情况,提出教师合作竞争网络模型的算法描述如下:

- (1) 初始网络: $t=0$ 时有 m_0 个顶点,已经连接成若干个项目完全子图,它们的项目度 h_0 之和为 h_0 ;
- (2) 点的新增: 在第 t 时间步增加一个新顶点;
- (3) 边的新增: 新顶点按照择优概率

$$\prod(h_i) = \frac{\beta_i h_i}{\sum_j \beta_j h_j} \quad (1)$$

选取4个旧顶点,把这4个旧顶点和新顶点中两

两之间尚未连接的边都连上,构成一个新的项目完全子图。即在第 t 时间步增加4个4-完全子图和1个5-完全子图。

竞争力 β_i 定义为教师 i 的教学能力。而衡量一个教师教学能力的主要依据是以往的教学成绩,所以有无高三教学经验就成为选择高三教师的一条重要标准。但学校又不可能总是任用老教师教高三,总需要培养年轻教师,故对于没有高三经验的青年教师而言,高一、高二的教学成绩就成为能否继任高三教师的依据。但无论怎样只要能成为高三教师,就说明其具备了相当的教学能力,这样我们在本文中假设所有的高三教师的竞争力等同。

3 教师合作竞争网络拓扑结构分析 (Analysis of the Topological Structures for the Collaboration-competition Network)

3.1 度分布 (Degree Distribution)

人们为了计算网络度分布,已经创造了几种有效的方法。归纳起来可分为两类:一类是动力学方法,有平均场方法和率方程方法;另一类是概率论方法,有主方程方法和马氏链方法。本文采用平均场方法计算其项目度分布及度分布。

令 $h_i(t)$ 表示在 t 时间步顶点 i 的顶点项目度数, $k_i(t)$ 表示在 t 时间步顶点 i 的顶点度数。把 $h_i(t)$ 近似看作连续动力学函数,根据连续性理论,节点 i 的顶点项目度数 $h_i(t)$ 应该近似的满足下述动力学方程 (dynamic equation):

$$\frac{\partial h_i(t)}{\partial t} = \frac{4h_i(t)}{\sum_j h_j(t)} \quad (2)$$

由于在 t 足够大时, $\sum_j h_j(t) = 5(t-1) + h_0 \approx 5t$, 此方程可以简化为:

$$\frac{\partial h_i(t)}{\partial t} = \frac{4h_i(t)}{5t} \quad (3)$$

其初始条件为: 顶点 i 在 t_i 时刻进入系统, 此时其项目度为 $h_i(t_i) = 1$ 。由此解得: $h_i(t) = Ct^\alpha$, 其中 $\alpha = 4/5$, $C = 1/t_i^\alpha$ 。

由此可得连通度 $h_i(t)$ 小于 h 的概率表达式:

$$P[h_i(t) < h] = P\left[t_i > \frac{t}{h^{5/\alpha}}\right] \quad (4)$$

由于每一时间步都有且仅有一个顶点加入原网络, 因此 t_i 服从均匀分布, 其概率密度为:

$$P(t_i) = \frac{1}{m_0 + t} \quad (5)$$

故

$$P\left[t_i > \frac{t}{h^{1/\alpha}}\right] = 1 - \frac{t/h^{1/\alpha}}{m_0 + t} \quad (6)$$

由 $p(h) = \frac{\partial p[h_i(t) < h]}{\partial h}$, 可得顶点项目度分布为:

$$P(h) = \frac{5t}{4(m_0 + t)} h^{-9/5} \quad (7)$$

即项目度分布为幂律, 标度因子 $\gamma = 1/\alpha + 1 = 9/4$ 。

如 2.1 节所述, 当计算重复边时, 总有 $k_i = 4h_i$, 由此可得重复边的度分布为:

$$P(k) \propto k^{-\nu} \quad (8)$$

且 $\nu = \gamma$, 即度分布 $P(k)$ 和项目度分布 $P(h)$ 均为严格的幂函数, 而且标度因子相同。所以该网络属于无标度网络。无标度网络中极少量的度很大的顶点对于保持整个网络的同步稳定性具有至关重要的作用, 故顶点度很大的教师, 是学校的中坚力量, 对学校的公众形象和未来发展都会有相当大的影响。学校决策者可将此类教师作为班主任和中层领导干部的首选。

图 3 给出了网络计重复边的顶点度分布模拟结果, 图中横、纵两个坐标轴均为对数形式, 圆圈表示相应的统计数据, 实线表示理论结果。由于数值模拟中网络节点的数目有限, 而预测值则是 $t \rightarrow \infty$ 时的统计结果, 故出现幂律尾[11]现象。随着网络规模的增大, 模拟结果与理论计算值将越来越吻合。

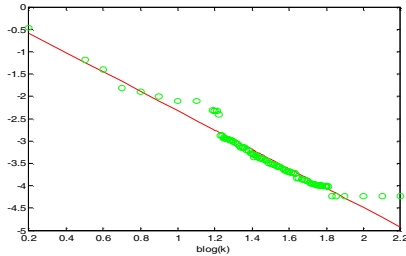


图 3: 网络计重复边的顶点度分布

3.2 其他统计特性(Another Statistical Characters)

本文选用了某中学 1996-2003 年考入大学的 1329 个学生(完全图项目), 245 个顶点(教师, 他们之间共有 1 513 条边)进行了统计。尽管数据量不大, 但是在采取了一些统计物理学上常用的减小统计误差的方法之后, 仍然可以得到不错的统计规律, 并且和在其他合作竞争网络中得到的结论以及模型的模拟结果一致, 说明对这样的数据的统计仍旧是合理可信的。

3.2.1 集聚系数 (Clustering Coefficient)

集群性起源于一个人的朋友中相互也是朋友的

人所占比例这一问题。节点的集聚系数描述网络中与其直接相连的节点之间的连接关系。网络的集聚系数为所有节点集聚系数的平均值。根据 Watts 和 Strogatz 对群集系数的定义[9], 计算网络群集系数的关键是计算 $E(l)$ 。目前文献中只有 BA 模型网络群集系数的经验结果[2]和解析结果。对于增长网络模型, 史定华等人提出了一种便于计算 $E(l)$ 的动力学方法[12]。本文采用此动力学方法计算教师合作竞争网络模型的群集系数。

假设该增长网络演化到第 N 时间步, 其中第 l 时间步加入的节点的群集系数 $C_l(N)$ 。由群集系数的定义, 关键是计算节点 l 邻域的连线期望数 $E(l)$ 。为此, 任意选择两个节点 i 和 j , 假定它们是第 i, j 时间步加入的, 且都与节点 l 有连线 (概率)。因为节点 l 邻域的任意两节点 i, j 的连线计算了两次, 故节点 l 邻域的连线期望数

$$E(l) \sim \left\{ \frac{N^{2\alpha-1} [l^{-2\alpha} (2\alpha-1)(\ln N - \ln l) - 2l^{-2\alpha} + l^{-1}]}{(2\alpha-1)^2} + \frac{l^{2\alpha-2} (2\alpha-1) \ln l + 3l^{-1} - 2l^{2\alpha-2}}{(2\alpha-1)^2} \right\} \quad (9)$$

由定义可得节点 l 的群集系数

$$C_l(N) \sim \left\{ \frac{(2\alpha-1)(\ln N - \ln l) - 2 + l^{2\alpha-1}}{(2\alpha-1)^2 N} + \frac{l^{4\alpha-2} (2\alpha-1) \ln l + 3l^{2\alpha-1} - 2l^{4\alpha-2}}{(2\alpha-1)^2 N^{2\alpha}} \right\} \quad (10)$$

用积分代替求和, 可计算网络群集系数为:

$$C(N) = \int_1^N C_l(N) dl / N \sim \frac{N^{-2\alpha-1}}{(2\alpha-1)^2} \times \left\{ \left[2\alpha + 3(2\alpha)^{-1} - 3 \right] N^{2\alpha} + \frac{2\alpha-1}{4\alpha-1} N^{4\alpha-1} \ln N + \left[\frac{1}{2\alpha} - \frac{10\alpha-3}{(4\alpha-1)^2} \right] N^{4\alpha-1} + \frac{2\alpha-1}{4\alpha-1} N^{4\alpha-1} \ln N + \left[(2\alpha-1)(\ln N + 1) + (2\alpha)^{-1} - 2 \right] N^{2\alpha-1} + \frac{10\alpha-3}{(4\alpha-1)^2} - \frac{3}{2\alpha} \right\} \quad (11)$$

由于上式过于复杂, 可按照 $\alpha (\alpha = 4/5)$ 值只保留 N 的最高项, 即第二项进行简化, 则

$$C(N) \sim N^{2\alpha-2} \ln N \quad (12)$$

求得整个网络的集聚系数 $C \approx 0.847$ 。这意味着在

教师合作网络中存在着这样一些团队,团队中人们彼此熟悉,而更容易形成合作关系。实际上,一般的合作网络的集聚系数都很大,如电影演员网的集聚系数为 0.79[9],淮阳菜系网的集聚系数为 0.60[13]。

3.2.2 平均路径长度 (Average Path Length, APL)

一个顶点的平均路径长度(也称为平均距离)是它到所有其他顶点的最短路径的平均值。这个平均值在网络分析中具有极高的重要性。这是一个顶点向心程度的指标,平均距离低的教师一般是第一个获得新信息的人,并且源于他们的信息将会比源于其他源头的信息传递的更快。这样平均路径长度就是一个获取信息能力和衡量教师向心力的指标。学校决策者可以将此类人作为班主任的首选。

网络的平均路径长度 L 定义为任意两个顶点之间距离的平均值,而顶点间的距离定义为连接这两个顶点的最短路径上的边数。统计求得 $L = 1.46$ 。由于本文网络数据量不大,其结果可能存在一定的偏差,但和其他合作网络中的得到的结论以及模型的模拟结果一致,说明对这样的数据的统计仍旧是合理可信的。

综合 3.2.1、3.2.2 考虑,该网络具有较大的集聚系数及较小的平均路径长度,因此,教师合作网络具有小世界效应[2] (Small-world Effect)。

4 总结与引申(Conclusion and Explicate)

本文我们提出了一个教师合作竞争网络模型,解析计算了节点的顶点度分布和项目度分布,证明该网络为无标度网络,并对其度分布进行了数值模拟,统计求得了节点的集聚系数和平均路径长度,得出这个网络具有小世界效应。由这些统计特性得出了相应的学校决策者需要的理论依据。由于高三教师的重要性,本章中选用线性择优方法是可行的,但实际生活中,常见的应该是择优和随机相结合的模式;另,本文模型中引入了竞争力 β_i , 但根据教师合作竞争网的实际情况,取相同值的 β_i , 然而对于其他网络肯定会存在不同的竞争力,怎样量化此竞争力,怎样计算其度分布、集聚系数等,都是我们后续研究的问题。

参考文献

- [1] M. E. J. Newman, The structure and function of complex networks, SIAM Rev., Vol. 45, 167-256, 2003.
- [2] R. Albert, A. L. Barabasi, Statistical mechanics of complex networks, Reviews of Modern physics, Vol.74, 47-94, 2002.
- [3] H. Inoue, W. Souma and S. Tamada, Spatial characteristics of joint application networks of Japanese patents. Physica A, Vol.383, 152-157, 2007.
- [4] M. Chmiel, J. Sienkiewica, K. Suchiecki and J.A. Holyst, Holyst, Networks of companies and branches in Poland physics. Physica A, Vol.383, 134-138, 2007.
- [5] A. L. Barabasi, R. Albert, Emergence of Scaling in Random Networks R. Science, Vol.286, 509-512, 1999.
- [6] G. Bianconi, Barabasi A. L., Bose-Einstein condensation in

- complex networks. Phys. Lett. A, Vol. 86, 5632-5635, 2001.
- [7] P. Zhang, K. Chen et al, Model and empirical study on some collaboration networks. Physica A, Vol.360, 599-616, 2006.
- [8] 何大韧.关于合作网络研究进展与前景的浅见.中国高等科学技术中心,CCAST-WL workshop series:Vol.170(1), Second National Forum on Complex Dynamical Networks, 205-212.
- [9] D. J. Watts, S. H. Strogatz, Collective dynamics of small-world networks. Nature, 393,440-442, 1998.
- [10] 何阅,张培培,许田,等.一个科研合作网的双粒子图自适应演化模型.物理学报, Vol.53, No.6:1710-1715, 2004.
- [11] Aherrere J., Sornette D., Stretched and exponential distributions in nature and economy: 'at tails' with characteristic scales. Eur. Phys. J. B., Vol, 2, 525-539, 1998.
- [12] 郭雷,许晓鸣.复杂网络,上海科技教育出版社,上海,2006.
- [13] 张培培,侯威,何阅等.淮阳菜系的网络描述,复杂系统与复杂性科学, Vol.2, 49-53, 2005.