

基于遗传算法的重复囚徒困境博弈策略 在复杂网络中的演化^{*}

林 海 吴晨旭[†]

(厦门大学物理系, 理论物理与天体物理研究所, 厦门 361005)

(2006 年 11 月 16 日收到, 2007 年 4 月 5 日收到修改稿)

利用遗传算法研究重复囚徒困境博弈策略在复杂网络中的演化. 研究表明, 处于复杂网络中有记忆的个体通过基因的复制、重组、变异和选择能够进化出一种自组织的合作机制. 这种合作机制既能够在群体中激发合作行为的产生, 加强和维护持续的合作行为, 同时又能对背叛的个体进行惩罚和报复, 因此能够促使复杂网络中进化出具有很高合作率的群体.

关键词: 复杂网络, 遗传算法, 进化博弈, 合作

PACC: 0175, 0565

1. 引 言

达尔文的进化论认为自然选择会促进有利于个体生存和繁殖的那些性状的发展. 那么, 生物界中普遍存在的合作行为是如何在自然选择的压力下进化出来的呢? 进化博弈论为理解合作行为的起源及演化提供了一个强有力的理论框架^[1-3]. 研究生物群体中进化博弈的传统方法通常假设个体是均匀混合的, 即群体中的任何一个个体都以同样的概率和其他个体相遇并进行博弈. 然而, 这种模型过于理想化, 因为现实中的生物个体活动范围总是有限的, 由此组成的群体具有一定的空间分布或者空间结构. 因此 Nawak 等^[4]又研究了位于二维规则格点上的进化博弈, 他们假设每个个体都占据一个格点并且只与自己的近邻相互作用. 研究发现, 对于囚徒困境博弈, 空间结构有利于合作行为的演化. 随着近年来复杂网络^[5-8]理论研究的兴起, 人们发现规则的格点并不能很好地描述现实中的各种关系网络. 现实中的各种网络包括社会网络、生物网络、技术网络等等, 其拓扑结构通常都具有小世界 (small-world)^[9]或无标度 (scale-free)^[10]的特征. 因此, 复杂网络理论为具有空间结构的生物群体的进化博弈论研究提供了

一个更为实际的框架^[11-13]. 另一方面, 文献 [11-13] 研究的个体可以认为是没有记忆的, 个体当前的行为仅取决于上一轮博弈的结果. Axelrod^[14]则假设个体能够记忆与其他个体前 3 轮的博弈情况. 利用遗传算法研究有记忆的个体在均匀混合群体中的演化, 得到了一些很有意义的结果, 如经过基因的复制、重组、变异和选择, 个体能够进化出一报还一报 (tit for tat, 也称为针锋相对) 的行为模式等等. 本文则把基于有记忆个体的遗传算法模型应用到位于复杂网络中的群体. 研究发现, 复杂网络中的群体能够演化出与均匀混合群体不同的一些行为模式. 这些行为模式能够促使复杂网络中进化出具有很高合作率的群体.

2. 模 型

囚徒困境是博弈论中的经典模型. 在囚徒困境博弈中, 两个个体同时决定是合作还是背叛. 如果两者互相合作, 则双方都得到 R 的收益. 如果互相背叛, 则得益 P . 如果甲合作而乙背叛, 则甲得益 S , 乙得益 T . 这里 $T > R > P > S$, 同时 $2R > S + T$, 即相互合作双方的总收益大于一方背叛另一方的总收益. 然而对于个体而言, 以背叛对合作得到的收益却

^{*} 国家杰出青年科学基金 (批准号: 10225420) 资助的课题.

[†] 通讯联系人. E-mail: lxwu@xmu.edu.cn

要大于以合作对合作.因此,囚徒困境模型展示了个体利益与群体利益的冲突.生物界群体中的个体经常面临囚徒困境的局面,例如有一种吸血蝙蝠能够把食物喂给未吃饱的同伴^[15],猫鼬(meerkat)为正在觅食的同伴警戒捕食者^[16]等等.在这类例子中,尽管单方面背叛行为的收益大于与对方合作时的收益,然而大自然还是演化出了不少相互合作的物种.

我们构建了4种不同拓扑结构的复杂网络,利用遗传算法^[17]研究位于复杂网络中的个体在囚徒困境模型的框架下如何演化出合作行为.规则网络、小世界网络、随机网络采用Watts和Strogatz^[9]提出的WS模型来构建,在该模型中分别令重连概率 p 为0,0.01,1即可得到上述3种网络.无标度网络则采用Barabási和Albert^[10]提出的BA模型构建.

网络中的每一个节点都代表一个参与博弈的个体,这些节点只和与自己有直接连接的邻居节点相互作用.个体可以记住与每个邻居最近3轮的博弈历史.用数字1表示合作,数字0表示背叛,可以用一个6位的比特串来表示最近3轮博弈的情况.例如 $H_{ij} = 100111$ 表示 i 节点与 j 节点最近3轮的博弈历史为: i 合作, j 背叛; i 背叛, j 合作; i 合作, j 合作.3轮博弈可能的历史记忆共有64种,即000000,000001,...,111111.用一个64位的比特串表示个体针对不同的博弈历史所采取的策略.这个比特串可以看成包含64个等位基因的基因组,这些等位基因按其所处的位置从0到63用十进制数字编号,代表64种可能的历史记忆.每个等位基因取值为1或0,代表相对于该历史记忆所采取的策略是合作还是背叛.例如,节点 i 的基因组为 $G_i = 10110011\dots$,其含义如下:0号基因取值为1,表示如果和某邻居的博弈历史为000000,则下一步对该邻居采取合作策略;1号基因取值为0,表示如果历史记忆为000001,则下一步采取背叛策略;.....

具体模拟步骤如下:初始群体中个体的每一位等位基因随机赋值为1或0,前3轮博弈随机采取合作或背叛的策略,然后每个个体根据自己的基因组和博弈历史与所有的邻居进行 n 轮重复囚徒困境博弈.博弈后将所得的总收益对邻居数目(即节点的度)求平均作为该个体的适应度.对每一个节点,以正比于适应度的概率在其邻居(包含该节点本身)中选取两个个体作为父代,父代基因组以一定的概率经过交叉重组和变异产生两个子代基因组(图1),用其中之一取代该节点的基因组.

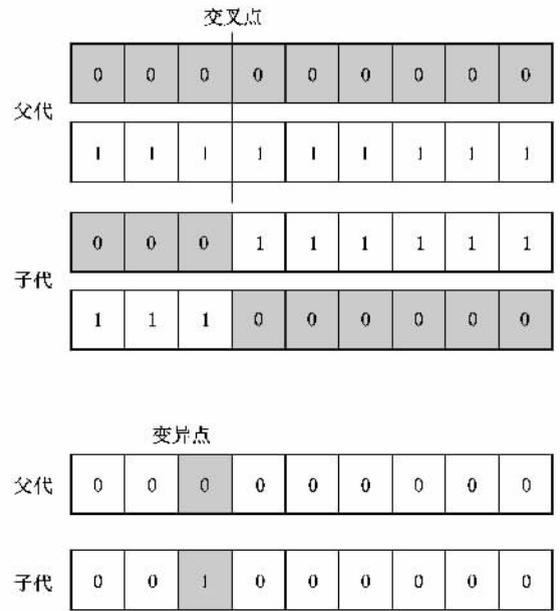


图1 基因重组和变异示意图

模拟参数选取如下:所构建的网络包含1000个节点,节点平均度为6.囚徒困境博弈的支付矩阵值取为 $R = 1, T = bR, b = 1-2, S = 0, P = 0.1$.参数 $b = T/R$ 表示当乙个体采取合作策略时,甲个体采取背叛策略所获得的收益与采取合作策略所获得的收益之比. b 值表征背叛的诱惑力,当 b 值增大时,背叛的个体将可能获得更大的收益,因此背叛行为对个体将具有更大的吸引力.我们取 b 值范围在1—2之间,用以考察系统中个体博弈策略模式的变化情况.模拟中,交叉重组概率取为0.7,变异概率为0.001.每个个体每一代都与邻居进行100轮博弈.在需要求平均值时,则对于4种不同拓扑结构网络的每一个 b 值,都取100个网络样本,在每一网络样本中模拟1100代进化,取后100代求各种平均值.然后,再将该平均值对样本数求平均得到最后的数据点.

3. 模拟结果及讨论

3.1. 合作频率

合作频率定义为在所有个体的总博弈次数中采取合作策略的比例.图2是4种网络中合作频率随遗传世代变化的关系.由图2可见,随着群体的遗传进化,不同网络中的个体合作频率都迅速上升,很快就达到很高的合作水平,即经过遗传进化之后系统

中的绝大部分个体都采取合作的策略。

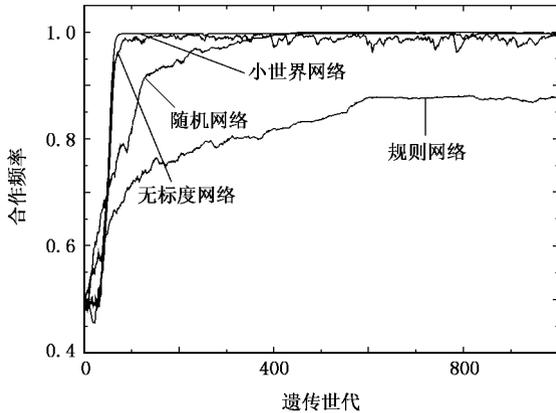


图2 $b = 1.5$ 时4种网络中合作频率随遗传世代的变化

图3是进化1000代后4种网络的合作频率与 b 值的关系。从图3可以看出,小世界网络中的合作频率高于随机网络和无标度网络,而规则网络的合作频率相对要低些。与文献[12]中采用的无记忆模型比较,对于无标度网络,两种模型都得到很高的合作频率,而对于规则网络,则由我们的模型得到的合作频率远高于无记忆模型。

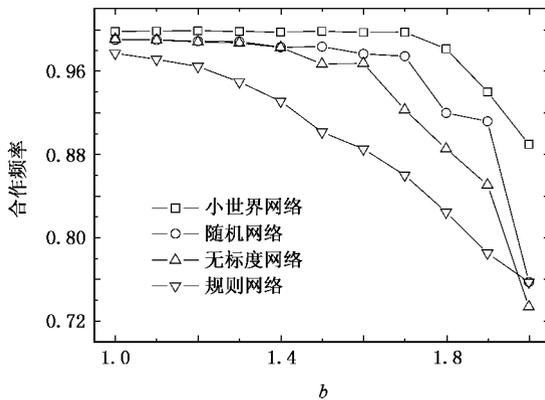


图3 4种网络中合作频率与 b 值的关系

3.2. 基因频率

我们考察演化1000代后网络中个体的基因分布情况,以研究系统通过遗传进化出何种机制促使复杂网络中产生高合作率的群体。为此,统计基因组中各个位置上值为1的等位基因在群体中出现的频率,即

$$p_k = \sum_i G_{ik} / N,$$

其中 k 是基因编号, G_{ik} 表示个体 i 编号为 k 的基因值, N 是网络节点数。图4(a)和(b)给出了 $b = 1.2$

和 $b = 1.9$ 时的 p_k 值。由图4(a)(b)可见,经过1000个世代的遗传进化之后,几种不同拓扑结构的网络中都演化出了类似的行为模式。为了更清楚地显示模式的共同特征,把 p_k 值对4种网络求平均,得到图4(c)(d)。

从图4(a)(c)可以看出,0,15,31,47,63号基因频率具有正峰值,说明群体对相应的记忆倾向于采取合作策略。从图4(a)(c)还可以看出,42号基因及21号基因具有负峰值,表示群体对于这两种记忆倾向于采取背叛的策略。我们发现,群体演化出的这种模式能够促进合作行为的产生和持续。0号基因对应的历史记忆为000000,即双方连续3次互相背叛。群体中大部分个体的0号基因值为1,说明当连续互相背叛使得双方都只能获取低收益的情况下,这些个体将会尝试与对方合作,从而跳出互相背叛的循环。因此该模式能够激发群体中合作行为的产生。15,31,47号基因对应的记忆分别为001111,011111,101111。这些位置上的基因频率具有正峰值意味着只要有连续两次成功的互相合作历史,个体就倾向于继续加强合作关系。对应于连续3次合作的63号基因111111,几乎所有的个体都采取合作策略,使得该基因频率呈现极高的正峰值。这说明当较稳定的互相合作关系已经建立起来之后,几乎所有的个体都倾向于继续维持这种关系。

此外,群体还演化出一种对背叛者进行惩罚和报复的模式,表现在42号基因频率具有负峰值。该基因对应于记忆101010,即如果个体连续3次合作而对方总是背叛情况下,则该个体下一步将不再合作,转变为背叛施以报复。这种模式使得网络中背叛的个体无法长期对合作者进行剥削,因此不容易侵入一个由合作者组成的团簇中。

上述几种基因模式的相互作用形成了一种自组织的合作机制。这种机制既能激发网络中合作行为的产生,又能不断加强和维持已经建立的合作关系,同时还能报复和惩罚不愿合作的个体。因此,能够在复杂网络中演化出如图2和图3所示的具有高度合作率的群体。

除此之外,图4还显示出一个有趣的模式,即21号基因呈现负峰值。该基因对应于记忆010101,即个体连续背叛而对方总是合作。网络中大约70%的个体对这种情况倾向于继续背叛以获取高收益,表现出一种“贪便宜”和“欺负老实人”的倾向。

图4(b)(d)是 $b = 1.9$ 时群体演化出的模式。

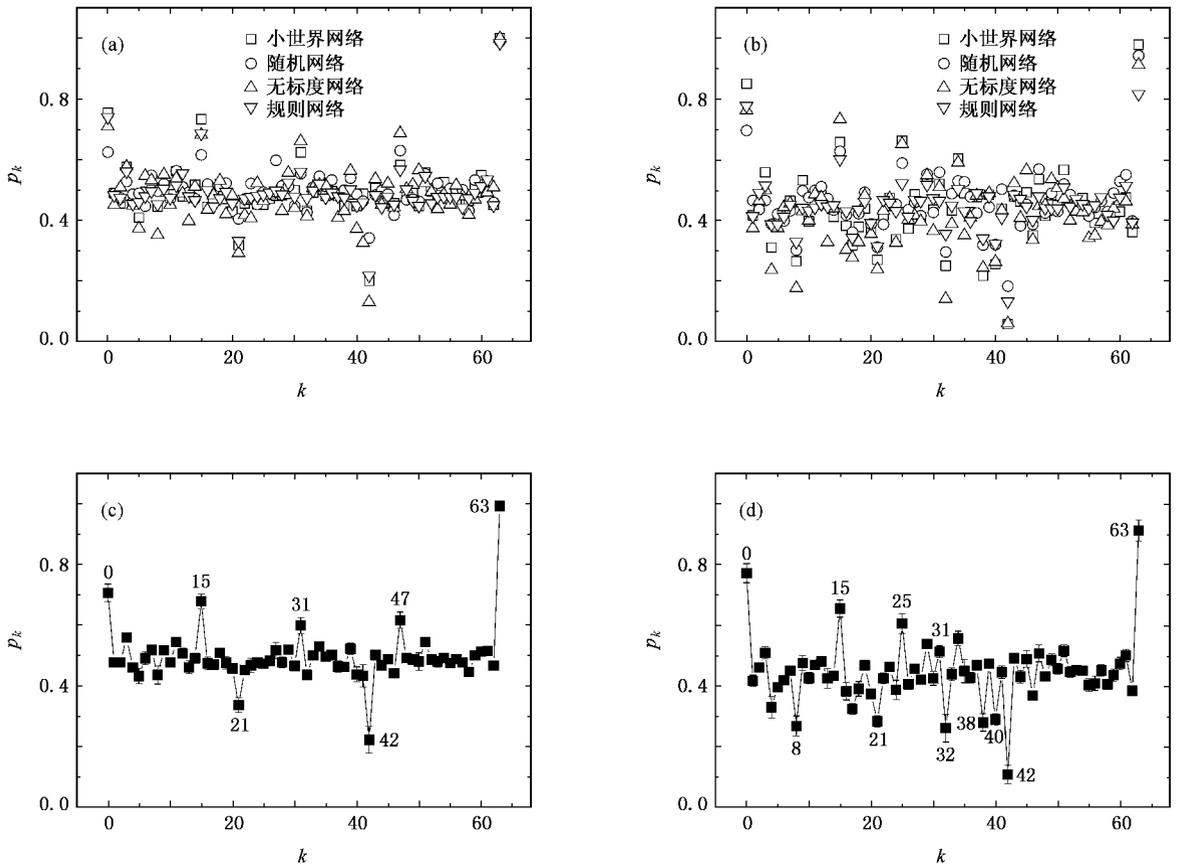


图4 基因频率模式 (a) $b=1.2$ 时4种网络中的基因频率 (b) $b=1.9$ 时4种网络中的基因频率 (c) $b=1.2$ 时基因频率对4种网络求平均 (d) $b=1.9$ 时基因频率对4种网络求平均

由图4(b)(d)可见,当 b 值即背叛与合作的收益比增大时,总体基因频率沿负方向略有移动,一些基因模式消失,另一些基因模式开始出现.但0,15,21,31,42,63号基因的基本模式仍然保持不变.图5给出了几种典型的基因频率随 b 值的变化曲线.

由图4(c)(d)及图5可以看出,当 b 值增大时,个体对合作者的要求开始变得“苛刻”.47号基因101111峰值消失,说明个体不再倾向于原谅对方的偶尔背叛.8号基因001000,32号基因100000开始出现负峰值,说明当个体尝试与相互背叛的对手合作却没有得到对方的回报时,个体将放弃合作的尝试而报之以背叛的策略.

图5还表明,当 b 值较大时,还产生了一种两个博弈个体交替作为背叛者和合作者的模式,即01,10,01,10,...的模式.具体表现在25号基因011001开始出现正峰值,而38号基因100110出现负峰值.原因是具有这种交替模式的两个个体每一轮博弈的平均收益为 $(b+0.1)/2$,而互相合作的两个个体每一轮的收益为1,因此当 b 增大时,交替模式的收益

开始接近甚至超过互相合作的收益,从而使得这种基因模式能够通过选择在群体中扩散.

比较图5不同网络中63号基因随 b 值的变化关系曲线,可以发现:对于小世界网络,63号基因频率基本不随 b 值增大而改变,因此这种网络中演化出的合作关系最为稳定.而规则网络中的个体则对 b 值较为敏感,一旦 b 值增大,规则网络中很快就出现一部分自私的背叛者对合作者进行剥削以获得高收益.随机网络和无标度网络则介于两者之间.这可以解释图3所示的结果:小世界网络中的合作率最高,其次是随机网络和无标度网络,规则网络相对于其他3种网络合作率最低.

3.3. 与均匀混合群体演化结果的比较

文献[14]研究了基于遗传算法的重复囚徒困境博弈策略在均匀混合群体中演化的结果,作者发现,均匀混合群体中可以演化出如下5种行为模式:个体对27号基因011011,47号基因101111和63号基因111111倾向于采取合作策略,而对0号基因

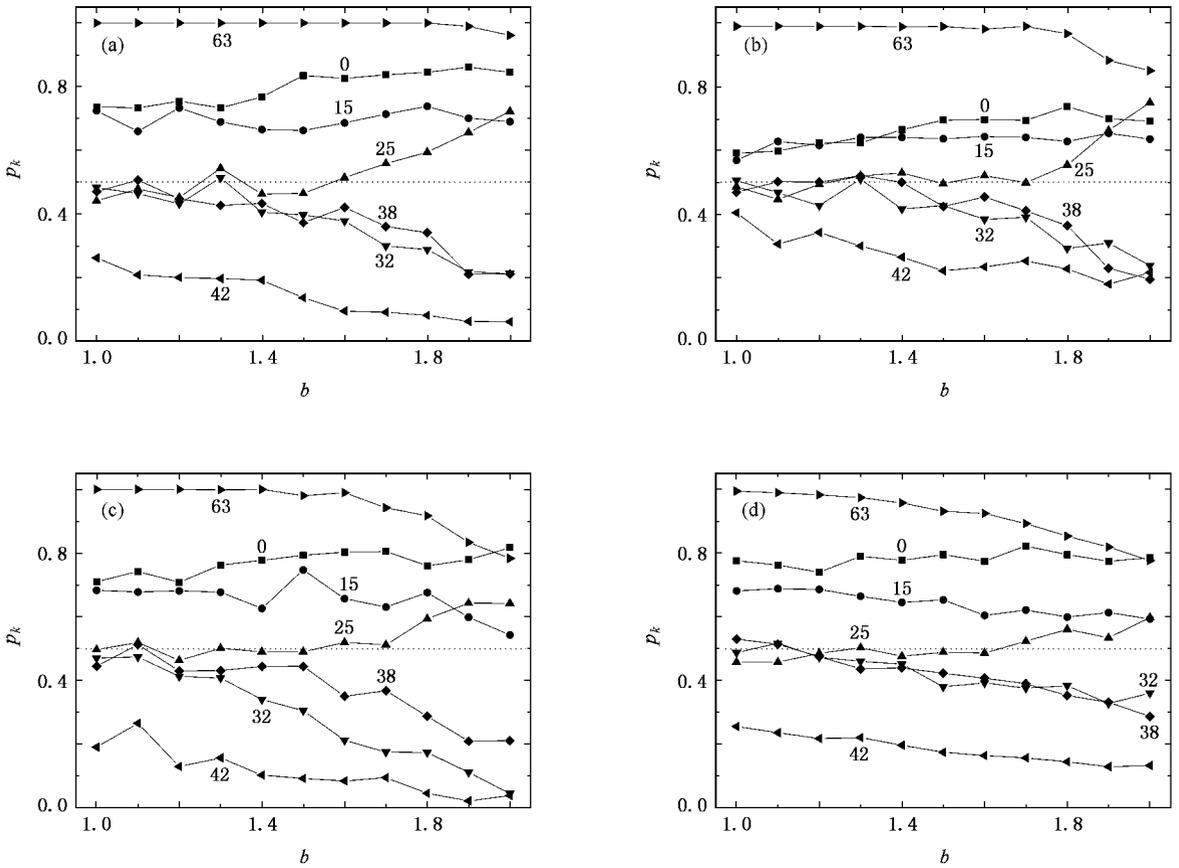


图5 4种网络中基因频率随 b 值的变化关系 (a)小世界网络 (b)随机网络 (c)无标度网络 (d)规则网络

000000 和 62 号基因 111110 采取背叛策略. 与本文在复杂网络中得到的结果比较, 我们发现两种模型对 111111 的记忆都采取合作的策略, 说明两种群体中的个体都倾向于维持连续的合作. 两者最大的区别在于对 0 号基因 000000 的反应: 复杂网络中的群体倾向于合作, 而均匀混合群体则倾向于背叛. 这说明复杂网络中的群体能够在连续互相背叛之后尝试与对方合作, 从而更容易激发合作行为的产生, 而均匀混合群体一旦落入互相背叛的恶性循环, 则很难从中跳出. 另外, 复杂网络中的群体还演化出 15 号基因 001111 31 号基因 011111 频率的正峰值模式, 使得个体之间的合作关系能够不断加强, 因此复杂网络中能够演化出比均匀混合群体具有更高合作率的群体. 此外, 复杂网络中合作的个体对于背叛者的报复显得更为“宽容”一些, 表现在均匀混合群体演化出的是一报还一报的策略, 即对 111110 的记忆采取背叛的策略, 而复杂网络则演化出“三报还一报”的策略, 即对 101010 的记忆采取背叛策略.

我们的模型综合了无记忆模型的网络结构特征及均匀混合群体模型的记忆效应, 这两方面的因素

共同决定了这样的群体能够演化出更高的合作率. 一方面, 处于网络中的个体能够形成由合作者组成的团簇, 团簇中的个体通过互相合作将获得比互相背叛的个体更高的收益^[4, 42]. 另一方面, 记忆效应使得个体能够识别合作和不合作的个体, 从而有针对性地与合作者合作, 对背叛者进行报复, 使得背叛的自私个体无法侵入由合作者组成的团簇中, 因此无法通过剥削合作者获取额外的高收益. 两方面因素的共同作用使得合作的个体比背叛的个体具有更高的适应度. 因此通过自然选择的优胜劣汰机制, 即使在 b 值较大的情况下, 合作行为仍然能够在网络中蔓延开来并最终占据主导地位.

4. 结 论

本文利用遗传算法研究复杂网络中重复囚徒困境博弈策略的演化. 研究表明: 在自然选择的压力下, 面临囚徒困境局面的生物个体追求自身利益最大化的结果, 并非必然导致个体的自私行为. 在我们的模型中, 网络结构和记忆效应两方面的因素决

定了合作行为能够通过自然选择从最初纯粹的自私机制中进化而来. 处于复杂网络中与其直接邻居进行重复囚徒困境博弈的具有记忆能力的个体, 在经过基因的复制、重组、变异和选择之后, 能够自然地进化出一种促进合作的行为模式. 0 号基因模式能够激发合作行为从无到有, 15 号和 31 号等基因模式促使新建立的合作关系得到进一步的加强, 而 63 号基因模式则能够维持持续的合作行为. 另外, 系统

不仅演化出促进合作的行为模式, 还演化出 42 号基因模式对背叛个体进行惩罚和报复, 使得背叛个体无法长期入侵合作者群体, 保证合作关系不致遭受破坏. 复杂网络中演化出的这几种基因模式的相互作用构成了一种自组织的合作机制, 能够促使网络中进化出比无记忆模型和均匀混合群体模型具有更高合作率的群体.

- [1] Maynard S J , Price G R 1973 *Nature* **246** 15
- [2] Gintis H 2000 *Game Theory: Evolving* (Princeton : Princeton University Press)
- [3] Axelrod R , Hamilton W D 1981 *Science* **221** 1390
- [4] Nawak M A , May R M 1992 *Nature* **359** 826
- [5] Albert R , Barabási A I 2002 *Rev. Mod. Phys.* **74** 47
- [6] Newman M E 2003 *SIAM Rev.* **45** 167
- [7] Li Y , Liu Y , Shan X M , Ren Y , Jiao J , Qiu B 2005 *Chin. Phys.* **14** 2153
- [8] 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]
- [9] Watts D J , Strogatz S H 1998 *Nature* **393** 440
- [10] Barabási A I , Albert R 1999 *Science* **286** 509
- [11] Lieberman E , Hauert C , Nowak M A 2005 *Nature* **433** 312
- [12] Santos F C , Pacheco J M 2005 *Phys. Rev. Lett.* **95** 098104
- [13] Ohtsuki H , Hauert C , Lieberman E , Nowak M A 2006 *Nature* **441** 502
- [14] Axelrod R 1987 *Genetic Algorithms and Simulating Annealing* (London : Pitman) p32
- [15] Wilkinson G S 1984 *Nature* **308** 184
- [16] Clutton-Brock T H , O ' Riain M J , Brotherton P N M , Gaynor D , Kansky R , Griffin A S , Manser M 1999 *Science* **284** 1640
- [17] Goldberg D E 1989 *Genetic Algorithms in Search , Optimization , and Machine Learning* (Massachusetts : Addison-Wesley)

Evolution of strategies based on genetic algorithm in the iterated prisoner ' s dilemma on complex networks *

Lin Hai Wu Chen-Xu[†]

(Institute of Theoretical Physics and Astrophysics , Department of Physics , Xiamen University , Xiamen 361005 , China)

(Received 16 November 2006 ; revised manuscript received 5 April 2007)

Abstract

Using genetic algorithm , we studied the evolution of strategies in the iterated prisoner ' s dilemma on complex networks . It is found that the agents located on complex networks can naturally develop some self-organization mechanics of cooperation by genome reproduction , recombination , mutation and selection , which can not only result in the emergence of cooperation , but also strengthen and sustain the persistent cooperation . At the same time , such mechanics punishes and takes revenge on defective agents , leading to a high cooperation rate on complex networks .

Keywords : complex network , genetic algorithm , evolutionary game , cooperation

PACC : 0175 , 0565

* Project supported by the National Natural Science Foundation for Distinguished Young Scholars of China (Grant No. 10225420).

[†] Corresponding author. E-mail : cxwu@xmu.edu.cn