

# The Evolution of Repeated Prisoner's Dilemma Game on Scale-Free Network\*

Tao Wang<sup>1,2</sup>, Zhigang Chen<sup>2</sup>, Tingting Wang<sup>3</sup>

<sup>1</sup>College of Information Science and Engineering, Hunan University, Changsha

<sup>2</sup>School of Information Science and Engineering, Central South University, Changsha

<sup>3</sup>Archives of Hunan University, Changsha

Email: yunlinren@qq.com, czg@mail.csu.edu.cn

Received: Oct. 10<sup>th</sup>, 2012; revised: Nov. 2<sup>nd</sup>, 2012; accepted: Nov. 19<sup>th</sup>, 2012

**Abstract:** In P2P network and mobile P2P network, independent individuals need to share resources under limited bandwidths and powers. It is valuable to study how to increase the systems' cooperation while reduce free-riding. Repeated Prisoner's Dilemma game is widely studied in the fields of biology, sociology, economics and informatics. People's interests focus on the cooperation emerged in a system that individuals are selfish. We study the iterated games evolved on scale-free network with genetic algorithm, and reveal the cooperation mechanism of nodes in networks. Nodes can remember the game historical strategies and code them into genes. We show that memory length's effect to cooperation level, and that some characteristic genes and frequently used genes emerge after some generations. We also study cooperate strategy distribution on different degree-nodes. These results may give theoretic support to the design of a self-organized system which can support cooperation.

**Keywords:** Evolutionary Game; Genetic Algorithm; Complex Network

## 重复囚徒博弈在无标度网络中的演化\*

王涛<sup>1,2</sup>, 陈志刚<sup>2</sup>, 王婷婷<sup>3</sup>

<sup>1</sup>湖南大学信息科学与工程学院, 长沙

<sup>2</sup>中南大学信息科学与工程学院, 长沙

<sup>3</sup>湖南大学档案馆, 长沙

Email: yunlinren@qq.com, czg@mail.csu.edu.cn

收稿日期: 2012年10月10日; 修回日期: 2012年11月2日; 录用日期: 2012年11月19日

**摘要:** 在 P2P 和移动 P2P 网络中, 自主节点都需要在有限资源(带宽, 电源等)下进行通信和数据共享。如何提高系统的合作水平而减少背叛(搭便车), 是个值得深入研究的问题。重复囚徒博弈在生物学、社会学、经济学、信息学等领域正在被广泛的研究, 在个体自私的情况下整体涌现合作行为是人们感兴趣的焦点。本文利用遗传算法研究重复囚徒困境博弈在无标度网络中的演化, 揭示网络中节点产生合作的相关机制。网络中的节点记忆以前多次和相邻节点的博弈情况, 按照一定的编码方法转换成遗传算法中的基因, 本文研究了不同记忆长度对合作水平的影响, 特征基因的显现分布和基因使用频率, 合作节点的度分布情况等。这些研究结论对于设计一个自组织的具有合作机制的系统提供了理论上的支持。

**关键词:** 演化博弈; 遗传算法; 复杂网络

\*资助信息: 本文受到中国国家自然科学基金 61103202, 61073186, 60903058, 61272149 的部分资助。

## 1. 引言

近年来, 基于 Internet 的 P2P 技术发展迅猛, 通过网络中的普通节点共享数据和贡献资源, 其他用户可以从这些普通节点下载资源, 从而减少了网络中心节点的压力, 提高了整个网络效率。但是网络中的独立个体通常只有有限的带宽和其他资源, 出于对自身利益的保护, 这些个体倾向于使用资源而不提供资源。如何促进这种分布式系统中的节点的合作, 成为现阶段网络技术中的热点问题。

在人类社会和自然界中, 自私个体之间产生合作是一个普遍的现象, 得到了很多研究和重视<sup>[1-3]</sup>。应用博弈论来解释合作的涌现在这些研究中占据了很重要的地位<sup>[4-6]</sup>。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<sup>[4]</sup>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)<sup>[7]</sup>或无标度(scale-free)<sup>[8]</sup>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。另一方面, 文献[9,10]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<sup>[11]</sup>等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, 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, 而林海<sup>[12]</sup>将其应用到位于复杂网络中的群体。本文在前人研究的基础上, 进一步研究了囚徒博弈在 BA 网络的演化, 得到了一些有趣的结论。这些结论能够帮助人们设计出具有更好的合作性的自组织系统。

## 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  $T > R > P > S$ , 对于个体而言, 以背叛对合作得到的收益却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。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。我们构建了一个无标度网络, 利用遗传算法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。无标度网络则采用 Barabási 和 Albert 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  $L$  轮的博弈历史。假设  $L = 3$ , 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。

具体模拟步骤如下: 设记忆长度为  $L$ , 初始群体中个体的每一位等位基因随机赋值为 0 或 1, 前  $L$  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  $2 \times 2L$  轮重复囚徒困境博弈。之所以选择博弈次数为  $2 \times 2L$ , 是为了保证尽可能覆盖所有的等位基因, 例如, 当  $L = 3$  时, 可能的基因个数为 64, 每轮博弈的次数为 128。博弈后将所得的总收益对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, 用其中之一取代该节点的基因组。模

拟参数选取如下：所构建的网络包含 3000 个节点，节点平均度为 6。囚徒困境博弈的支付矩阵值取为  $R = 1$ ,  $T = r$ ,  $S = 0$ ,  $P = 0.1$ 。参数  $r = T/R$  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。 $r$  值表征背叛的诱惑力，当  $r$  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  $r$  值范围在 1~3 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  $2 \times 2$  轮博弈。在需要求平均值时，则对于每一个  $b$  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. 不同记忆长度对合作水平的影响

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。

从图 1 中可发现如下几个现象。1) 随着记忆长度的增加，对于不同的背叛值  $r$ ，整体合作水平由 50% 上升到稳定水平的演化代数并没有像想象中那样减少，反而有一定的增加。这是由于随着记忆长度增加，遗传基因的位数也增加，不同位对应的历史状态增加，而这些历史状态的取值要向高合作水平进化，其演化代数也就有所增加。2) 随着记忆长度的增加，对于不同的背叛值  $r$ ，稳定状态的合作水平并没有像想象中那样提高，反而有所下降。这是由于随着记忆长度增加，少量由背叛而获得高收益的行为也通过逐渐复杂的记忆历史记录下来，使得整体上合作水平有所下降。3) 随着记忆长度的增加，对于不同的背叛值  $r$ ，从初始合作水平到稳定合作水平的进化曲线变的平滑。

#### 3.2. 特征基因分布

特征基因是指系统中个体的基因位的统计平均

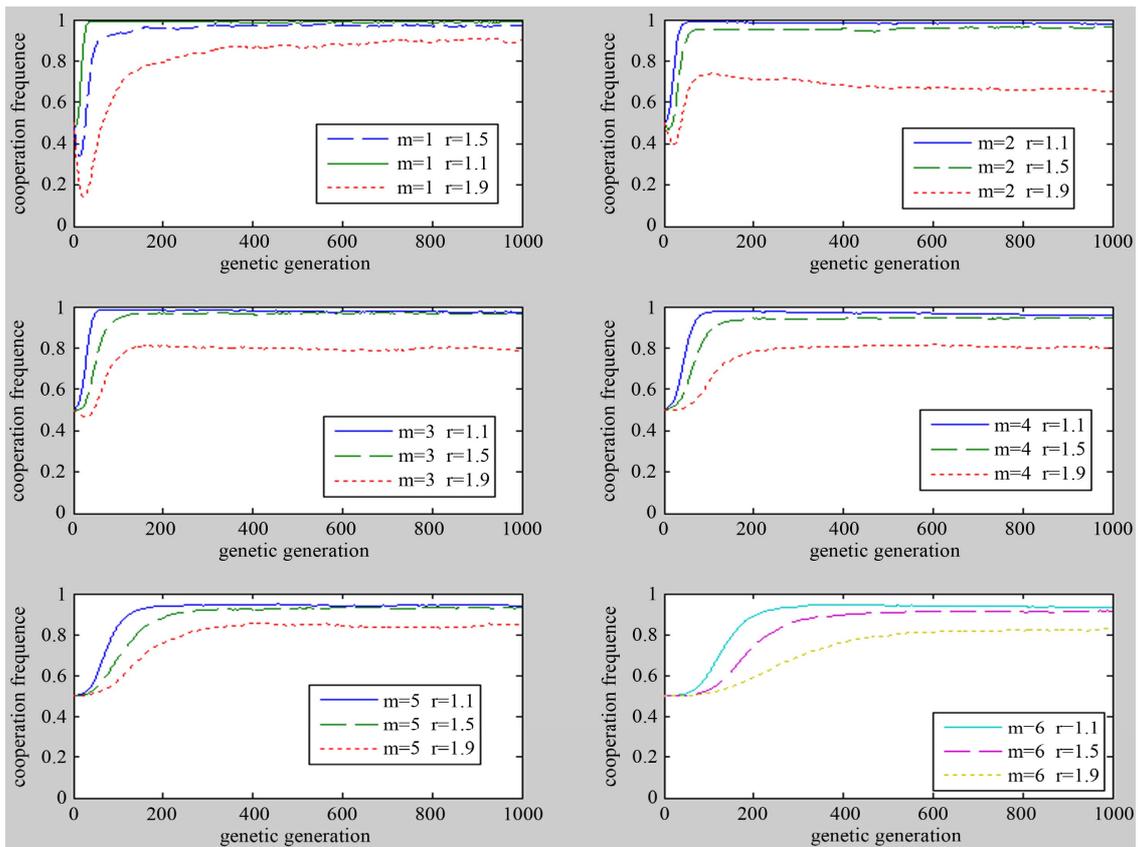


Figure 1. Cooperation ratio to different memory length and betray value r  
图 1. 对不同记忆长度和不同背叛值的合作水平

值，每个基因位对应于一个特定的博弈历史，当它为 0 时表示下次将背叛，为 1 时表示下次合作，根据统计平均值可以看出系统中的个体对特定博弈历史的策略倾向。

从图 2 中可以发现，不同记忆长度下，基因的特征点分布也不同，与文献[12]相似，我们也发现了以下一些规律。1) 记忆历史为全 0 或全 1 的都以较高概率在下一步选择合作；也就是说以前一直合作的节点

倾向于合作，以前一直不合作的节点由于没有收益也都会倾向于改变策略合作，如所有图中的 0，和最高位点。2) 一些基因演化出震荡的特性，特别是  $r$  值比较高时，如  $m = 3$  时的 25(011001)，38(100110)， $m = 4$  时的 102(01100110)，153(10011001)等等，这些基因使节点的收益增加。3) 欺负老实人与反欺负，如  $m = 3$  中的 21(010101)和 42(101010)， $m = 4$  中的 85(01010101)和 170(10101010)，这些基因最终有助于合作产生。

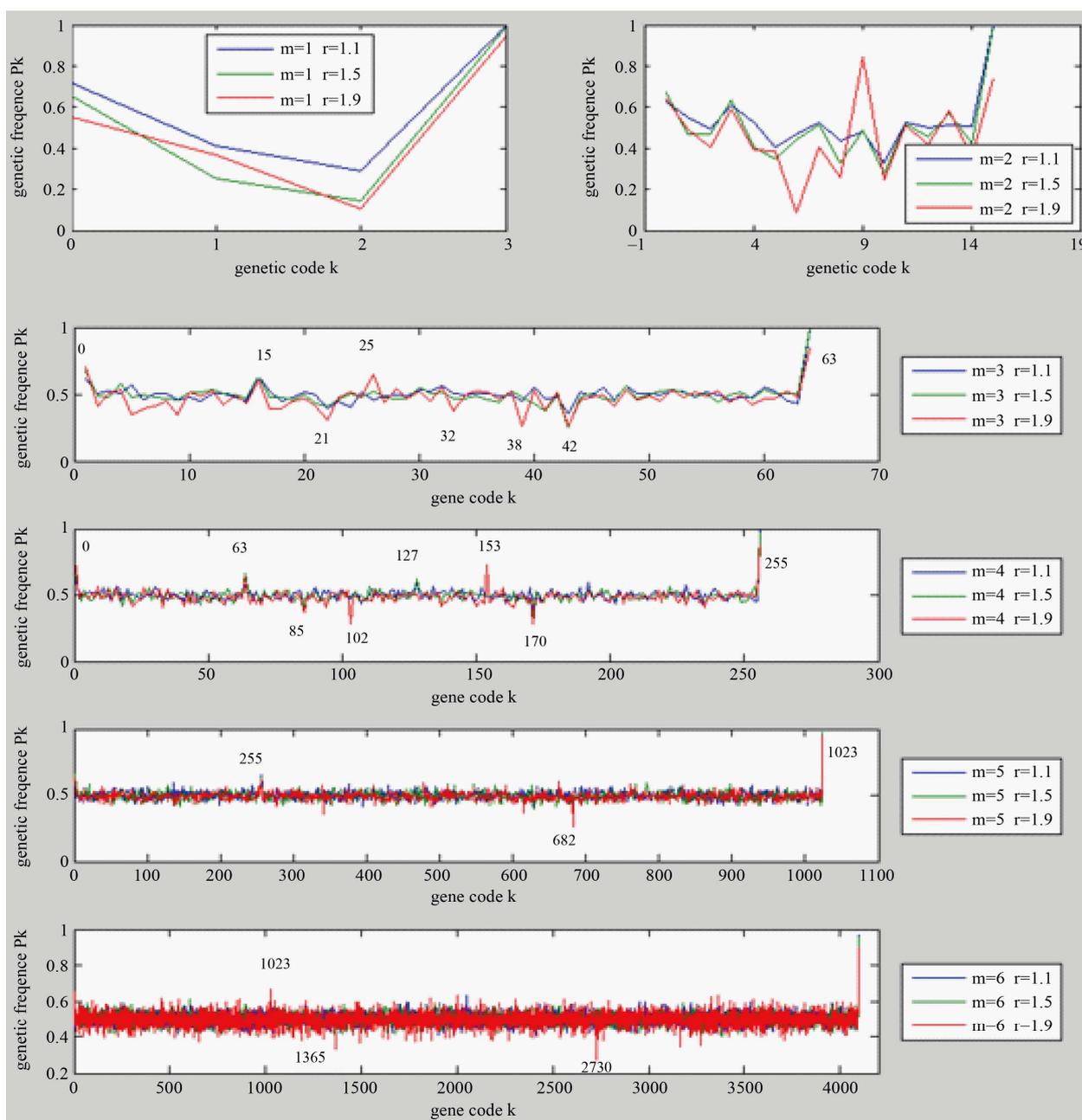


Figure 2. Distribution of characteristic genes  
图 2. 特征基因分布

### 3.3. 不同背叛收益下的基因分布与度分布情况

我们重点考察了记忆长度为 3 的时候, 基因被应用的分布情况。为了了解不同背叛值对基因应用分布的影响, 系统到达稳定以后(运行了 1000 代), 再运行 100 代, 统计每一代所有节点使用基因, 得到稳定后的基因使用分布图。

从图 3 中, 我们可以看出, 当  $r = 1.5$  时, 达到稳定合作水平后, 大部分节点对之间采用的策略都是合作, 基因 63 占了 90% 以上; 而当  $r = 1.9$  时, 使用 63 号基因的节点数下降明显只占 70% 左右, 而 25 号, 38 号基因各占到了 15% 左右;  $r = 2.1$  时, 25 号、38 号和 63 号基因各占了 30% 以上; 而当  $r = 2.5$  时, 25

号和 38 号基因占了接近 50%, 63 号基因比例很小。可见, 随着  $r$  的增加, 背叛者对合作者获得收益增加, 25 号(011001)和 38 号(100110)这种模式成对出现, 使得相互作用的两个节点收益最大化, 因此这两种基因占的比率逐步提高。这是因为当  $r$  增大, 背叛收益会增加, 节点选择背叛的概率加大。从宏观上来说,  $r$  值增加 2.1 后, 大部分节点都处于震荡中(01010101...如图 4(b))。也可以解释不管  $r$  值变得怎样大, 整体合作水平总可以稳定在 0.5 左右(如图 4(a))。

### 3.4. 合作策略的度分布情况

为了了解不同背叛值下, 合作策略与节点度之间

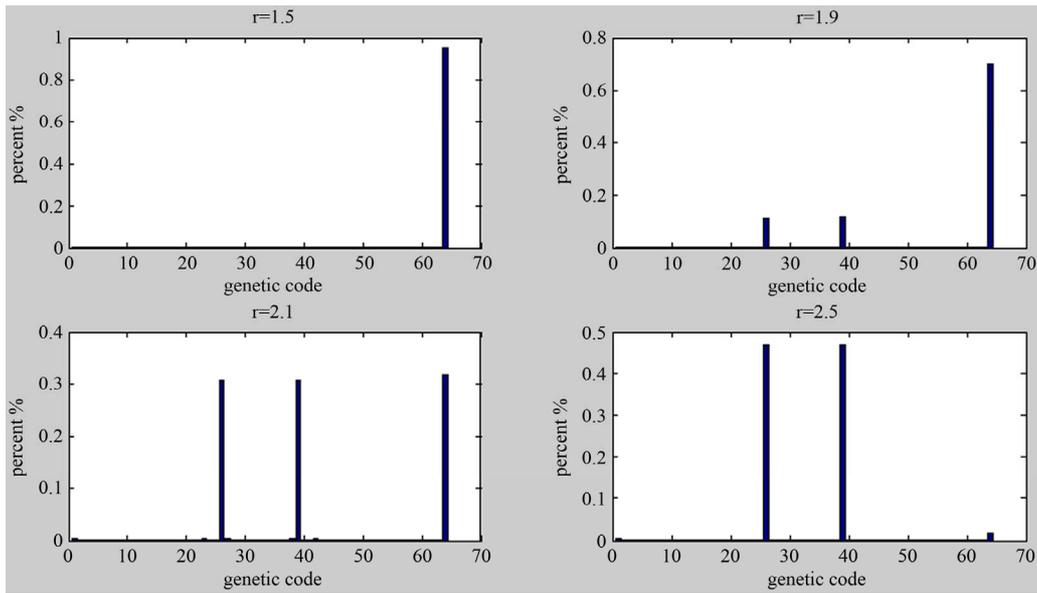


Figure 3. Gene's distribution to different betray value  $r$ . The figure gives the gene's distribution when memory length = 3  
图 3. 不同背叛值下的基因分布图。该图给出了记忆长度为 3 时, 基因被使用的分布情况

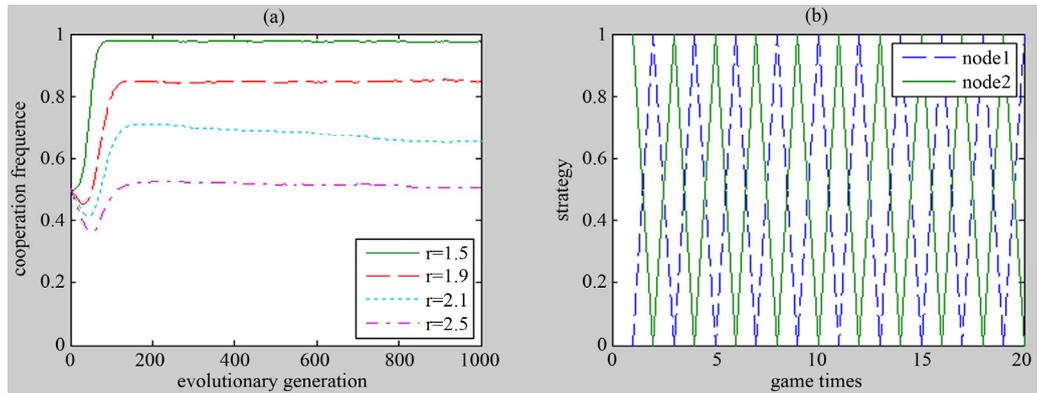


Figure 4. (a) The cooperation distribution when memory length = 3 under high  $r$  value; (b) The strategies two adjacent nodes adopt strategies when the system evolves to 1000 generations, where  $r = 2.5$   
图 4. (a) 记忆长度为 3 时, 高  $r$  值下的合作水平; (b) 演化第 1000 代时,  $r = 2.5$ , 两相邻节点一轮博弈前 20 次采取策略的情况

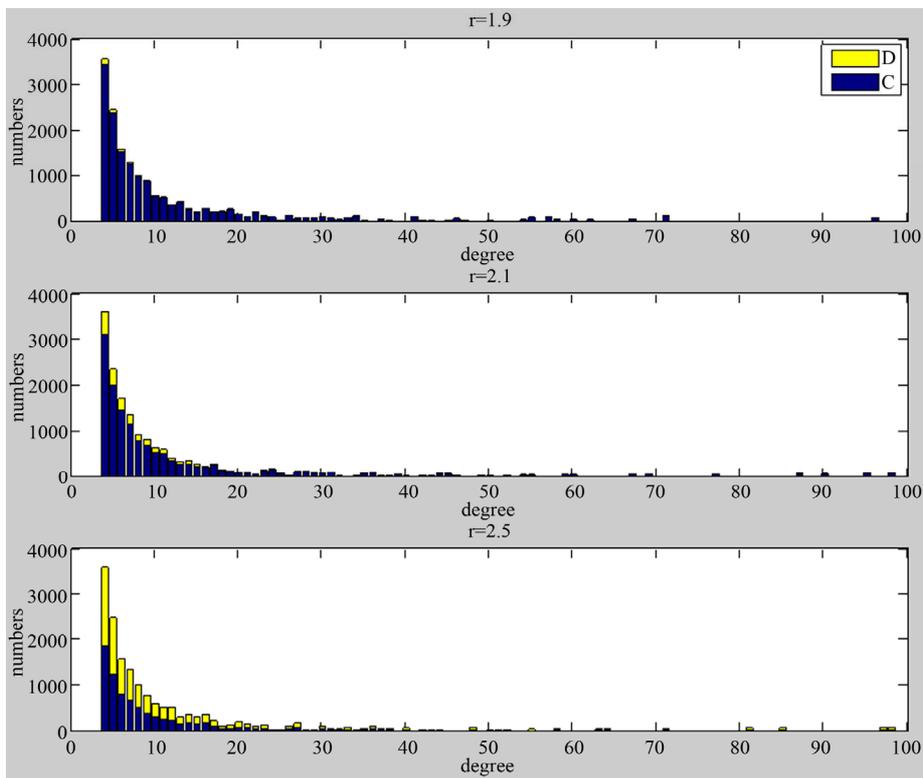


Figure 5. The cooperation-defection distribute on nodes' degree (There are total 3000 nodes, and the average degree is 6)  
 图 5. 合作 - 背叛策略的度分布(系统总节点数 3000, 平均度为 6)

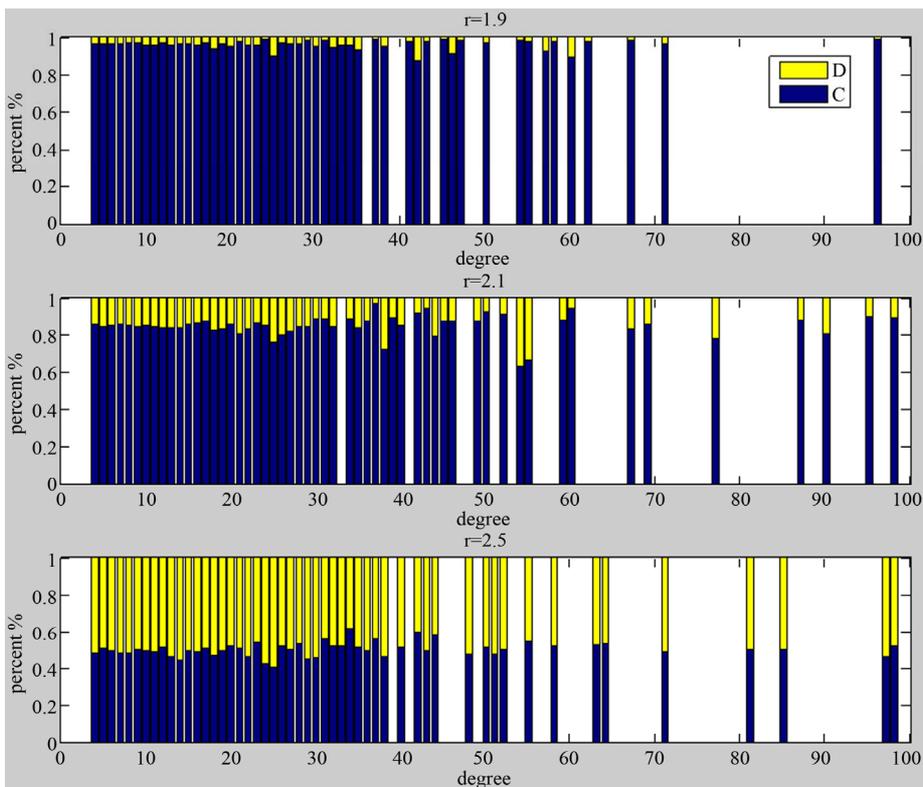


Figure 6. The cooperation-defection distribute on percentage of nodes' degree  
 图 6. 合作 - 背叛策略的百分比度分布

的关系，我们考察了记忆长度为 3 时，不同背叛值下合作策略的度分布。系统演化 1000 代以后，再运行 100 代，统计这 100 代不同的节点度上面采取合作策略和背叛策略的个数并求平均，得到图 5。

为了更清楚的看到策略的度分布情况，我们对每个度上面的合作背叛策略求了百分比，得到图 6。

从这个图我们可以清楚的看见，随着  $r$  值的增加，采用背叛策略的节点增加，但是背叛策略比较均衡的出现在不同的度的节点上，而不是出现在度小的节点或者度大的节点上。这可能与每个节点对于不同的邻居采用不同的策略有关。

#### 4. 结论

本文通过利用遗传算法研究 BA 网络上的囚徒博弈，得到一些结果。这些结果表明，在一定的情况下，个体的记忆长度增加并不能够有效的促进合作，特定的遗传基因在演化中逐渐显现出来，高背叛诱惑会导致节点策略处于震荡中，而与节点度没有必然联系。下一步的工作将考虑在更加多样的演化博弈中研究总体合作水平和其他参数的关系。

#### 参考文献 (References)

- [1] R. Axelrod. The evolution of cooperation. New York: Basic Books, 1984.
- [2] P. Hammerstein. Genetic and cultural evolution of cooperation. Cambridge: MIT, 2003.
- [3] K. Sigmund. Games of life. Oxford: Oxford University Press, 1993.
- [4] H. Gintis. Game theory evolving. Princeton: Princeton University, 2000.
- [5] M. Nowak, K. Sigmund. Tit for tat in heterogeneous populations. Nature (London), 1992, 355: 250-253.
- [6] M. Nowak, K. Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in Prisoner's Dilemma. Nature (London), 1993, 364: 56-58.
- [7] D. J. Watts, S. H. Strogatz. Network robustness and fragility: Percolation on random graphs. Nature, 1998, 393: 440-442.
- [8] A. I. Barabasi, R. Albert. Emergence of scaling in random networks. Science, 1999, 286(5439): 509-512.
- [9] F. C. Santos, J. M. Pacheco. Scale-free networks provide a unifying framework for the emergence of cooperation. Physical Review Letters, 2005, 95(9): Article ID: 098104.
- [10] H. Ohtsuki, C. Hauert, E. Lieberman and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. Nature, 2006, 441: 502-505.
- [11] R. Axelrod. Genetic algorithms and simulating annealing. London: Pitman, 1987: 32.
- [12] 林海等. 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[J]. 物理学报, 2007, 56(8): 4313-4318.