

第5章

网络演化博弈

博弈论(Game theory)主要是研究个体在相互作用过程中如何获得最大利益的理论,是对合作与竞争关系的一种反映。一般而言,一个博弈通常有以下几个组成部分:①参与博弈的个体至少两个;②博弈个体可以从策略集中选取自己的博弈策略;③博弈结束后博弈个体可以得到的收益;④博弈个体进行策略更新的目的是为了达到最大收益。经典博弈论认为,博弈个体是非常理性的,博弈目的都是追求自己的最大收益,而且也知道其他博弈个体也是完全理性的;而演化博弈论以种群为研究对象,认为博弈个体是有限理性的,博弈个体的策略可能因变异而改变。

演化博弈理论源于对生态现象的解释。1960年,生态学家 Lewontin 开始运用演化博弈理论的思想研究生态问题。生态学家从动植物演化的研究中发现,动植物演化结果在多数情况下都可以用博弈论的纳什均衡概念解释。然而,博弈论是研究完全理性的人类互动行为时提出来的,为什么能够解释根本无理性可言的动植物的进化现象呢?我们知道,动植物的进化遵循达尔文的“优胜劣汰”生物进化理论,生态演化的结果却能够利用博弈理论合理地进行解释,这种巧合意味着我们可以去掉经典博弈理论中理性人假定的要求。另外,1960年生态学理论研究取得突破性的进展,非合作博弈理论研究成果也不断涌现并日趋成熟,演化博弈理论具备了产生的现实及理论基础。1973年,生态学家 Smith 和 Price 结合生物进化论与经典博弈理论在研究生态演化现象的基础上提出了演化博弈理论的基本均衡概念——演化稳定策略(evolutionary stable strategy, ESS)^[44],目前学术界普遍认为演化稳定策略概念的提出标志着演化博弈理论的诞生。此后,演化博弈理论逐渐被广泛用于生态学、社会学、经济学等领域。1978年,生态学家 Taylor 和 Jonke 在考察生态演化现象时首次提出了演化博弈理论的基本动态概念——模仿者动态(replicator dynamics)。至此,演化博弈理论有了明确的研究目标。演化博弈理论与经典博弈理论的不同主要体现在以下几个方面。

(1) 策略的内涵不同。在经典博弈论中,个体有特定的策略集,他们从策略集中选择策略进行博弈。演化博弈理论研究生物系统的演化,而在生物系统中是没有任何明确的策略集概念的。生物系统中的策略集实际上是由不同类型的物种组成的。策略由物种的不同表现型体现。因此,对生物系统而言,个体不需要从策略集中选择特定的策略进行博弈,他们只继承或遗传其父辈的表型特征即可。在继承或遗传其父辈的表型特征时,会产生一定的变异,这些变异决定了与其他个体的竞争优势或者适应度的大小。从这个意义上讲,对所有的物种,都可以应用博弈论的概念和思想加以研究。对生物系统而言,经典

博弈论中的收益对应演化博弈理论中的适应度的概念。另外,个体的适应度受很多因素的影响,个体间的相互作用只是其中一个方面。对于这种情况,可以通过引入一个背景适应度衡量其他因素的影响。通常,个体的适应度可以通过该个体的后代的数量加以量化。这样,经典博弈论中的超理性的观点就被演化博弈论中的适者生存所替代。

(2) 均衡的意义不同。经典博弈理论中的核心概念是纳什均衡,而演化博弈论中的核心概念是演化稳定策略。Smith 和 Price 是在研究动物之间为争夺食物、领域或配偶等有限资源而发生冲突时提出演化稳定策略概念的,它与经典博弈论中的纳什均衡息息相关但不相同。如果一个种群中的所有个体都采用某种策略,而其他任何小的突变策略都不能入侵这个种群,则称该策略为演化稳定策略。如果一个群体的行为模式能够消除任何小的突变群体,那么这种行为模式一定能够获得比突变群体更高的期望收益。随着时间的演化,突变者群体最终会从原群体中消失,原群体选择的策略就是演化稳定策略。就生态现象而言,由于我们把每一个种群的行为都程式化为一个策略,因此演化的结果将会是突变种群的消失。ESS 可以是纯策略,也可以是混合策略。混合策略可以解释为以某种概率采取一定的纯策略。如果用 x^* 表示 ESS,用 x 表示任意突变策略,则 $E(x^*, x^*) > E(x, x^*)$ ① 或 $E(x^*, x^*) = E(x, x^*), E(x^*, x) > E(x, x)$ ②,这里 $E(a, b)$ 表示策略 a 遇到策略 b 时的收益。设想现在有一个均匀混合的无限大种群,其中突变策略 x 初始时很少,式①表示 x 对 x^* 的收益小于 x^* 对于自身的响应,因此,突变策略 x 会慢慢地从该种群中消失,即策略 x^* 可以抵御突变策略 x 的入侵;式②表示突变策略 x 对 x^* 与 x^* 对自身的收益相等,但 x^* 对 x 的收益要高于 x 对自身的收益。这样,即使初始时 x 的数量很少,少到可以忽略,但由于 $E(x^*, x^*) = E(x, x^*)$,突变策略 x 仍有可能在种群中发展起来,但是,当发展到一定程度,由于 $E(x^*, x) = E(x, x^*)$, x 降低了自身的适应度,从而限制了它走向繁荣,即 x^* 最终抵制了 x 的入侵。ESS 强调了突变策略以很少的数量进入 ESS 的种群是不能成功入侵的,但并不排除突变策略以较大的数量进入到 ESS 种群时可以成功入侵。另外,演化稳定策略的原始定义只对无限大的种群成立,2004 年 Nowak 等人^[142]发展了适用于有限种群的相关定义。

(3) 相互作用的方式不同。从经典博弈理论到演化博弈理论,个体相互作用的内涵发生了转变。在经典博弈论中,参与博弈的个体要么作用一次,要么与相同的对手作用多次;而在演化博弈论中,所采取的策略由其表型决定,虽然也是进行多次博弈,但每个参与博弈的个体都是随机地从群体中抽取并进行重复、匿名博弈,他们没有特定的博弈对手,这就避免了经典博弈论中个体记忆的概念。在 1992 年的研究中,这种随机相互作用的要求被放宽了,规则格子和复杂网络上的演化博弈被大量研究^[143]。在这种情况下,个体既可以通过自己的经验直接获得决策信息,也可以通过观察在相似环境中其他个体的决策并模仿而间接地获得决策信息,还可以通过观察博弈的历史而从群体分布中获得决策信息。参与人常常会模仿好的策略,不好的策略会在进化过程中淘汰,模仿是学习过程中的一个重要组成部分,成功的行为不仅以说教的形式传递下来,而且也容易被模仿。参与人由于受到理性的约束而其行为是幼稚的,其决策不是通过迅速的最优化计算得到,而是需要经历一个适应性的调整过程,在此过程中,参与人会受到其所处环境中各种确定性或随机性因素的影响。因此,系统均衡是达到均衡过程的函数,要更准确地描述参与人行为,

就必须考察系统的动态调整过程。动态均衡概念及动态模型在演化博弈理论中占有相当重要的地位。

在自然界和人类社会中,自私个体之间能够产生大量的合作是一个“惊人”的现象,得到许多学者的重视与研究。在这些研究中,采用博弈论解释合作涌现的现象占据了重要的地位^[144]。既然要讨论合作的涌现,必然涉及相当数量的局中人,通常认为这些局中人以及他们之间的关系构成一个复杂网络,随着时间的演化,每个局中人都和他的邻居进行博弈,这就称为网络演化博弈。它的定义可以表述为

- ① 数量 $N \rightarrow \infty$ (或一个足够大的数量) 的局中人位于一个复杂网络上。
- ② 在每一个时间演化步,选取的一部分局中人按照一定法则以一定频率匹配进行博弈。
- ③ 所有局中人的策略更新法则相同,局中人采取的对策可以按照一定的法则进行更新,这种法则称为“策略的策略”。但是,法则更新比博弈频率慢得多,这就使得局中人可以根据上一次的更新对策成功与否选择或调整下一次更新。
- ④ 局中人可以通过感知环境吸取信息,然后根据自己的经验和信念在策略更新法则下更新策略。
- ⑤ 策略更新法也有可能受局中人所在网络拓扑结构的影响。

5.1 复杂网络演化博弈基本框架

复杂网络上的演化博弈动力学研究主要涉及以下几方面的问题^[145]: 第一,策略种类的选取。不同类型的博弈具有不同的策略集;第二,相互作用邻居和策略学习邻居的定义。在这方面,通常假设个体只与其直接邻居交互并且学习他们的策略,但是也有研究者认为这两者可以不同;第三,策略更新动力学的选取。策略的更新可以是同步的,也可以是异步的(即随机序列更新)。不同的更新方式可能导致不同的演化结果。此外,如果考虑个体记忆以及引入不同的偏好学习方式,那么模型还可以进一步复杂化。由于以上因素的限制,复杂网络上的演化博弈研究通常很难得到普遍的结果。对于一些简单的空间结构,尽管可以进行解析分析,但依然会存在很多限制条件。因此,系统的空间演化通常通过计算机模拟进行研究。

目前,复杂网络上的演化博弈研究主要包括如下 3 个框架。

(1) 研究网络拓扑结构对博弈演化动力学的影响。相互作用的网络结构和博弈规则具有同等的重要性。学者们主要以囚徒困境博弈和铲雪博弈为博弈模型,研究了规则网络、小世界网络、无标度网络、层次网络、关联网络上的演化博弈特性,同时深入研究度分布、平均度、集聚系数、度相关性、社团结构等拓扑特性对演化博弈特性的影响。

(2) 探索一些可能的支持合作行为涌现的动力学机制。现实世界系统中的个体是具有能动性的,他们拥有记忆和自我反省能力;他们是自适应的,在博弈过程中可以采用各种方式选择邻居学习;他们又是现实的,有自己的期望,可以根据自己的期望水平选择适合自己的策略。博弈个体之间相互影响,这种相互影响可能是非对称的,并且是动态调整的。这方面的研究主要是根据现实复杂系统的动力学特征,对博弈模型进行补充和改进。

(3) 研究博弈动力学和网络拓扑结构的共演化,即个体策略和网络拓扑结构协同演化的情形。策略与拓扑结构的共同演化对合作行为的促进作用是显而易见的。复杂系统最本质的特点是反馈,并利用反馈信息实现自适应和自组织:一方面,网络的拓扑结构对其上的动力学过程会产生影响;另一方面,网络上的动力学过程也会反过来塑造网络的结构。共演化机制能够促进合作就是因为引入了反馈机制,从而更好地抓住了现实复杂系统最本质的特征。

总之,在研究中,需要考虑现实复杂系统的基本特征,改进网络模型和博弈模型,使之更加接近复杂系统的结构特征和微观动力学机制,从更普遍层面上研究网络的拓扑结构对演化博弈动力学的影响。同时需要进一步研究动力学演化如何改变网络的拓扑结构,也就是探索网络结构与其上动力学过程的相互作用、协同演化行为。这对于在更加接近客观实际、更加普遍的层面上研究复杂网络的演化博弈动力学,揭示相关的物理现象和规律,理解复杂系统局部规则对宏观性质和功能的影响具有重要的理论意义。

5.2 网络博弈动力学

在传统的演化博弈理论中通常假设个体间以均匀混合的方式交互,即所有个体全部相互接触,然而,现实情况中个体间的接触总是有限的,个体仅与周围的少数其他个体接触。这样我们就可以在博弈理论中引入网络拓扑的概念。

复杂网络理论为描述博弈个体之间的博弈关系提供了方便的系统框架。网络上的节点表示博弈个体,边代表与其邻居的博弈关系。在每一时间步长,节点与其所有邻居进行博弈,累积博弈获得的收益,然后根据更新规则进行策略更新,如此重复迭代下去。网络上的演化博弈研究主要集中于3个基本的方向:

- (1) 研究网络拓扑结构对博弈动力学演化结果的影响。
- (2) 在一定的网络结构下探讨各种演化规则对演化结果的影响。
- (3) 网络拓扑和博弈动力学的共演化主要是自适应网络上的博弈动力学,即网络拓扑调整受博弈动力学影响。

每个模型都可以分成几个模块,如使用的博弈模型、更新规则、网络结构等。虽然使用的博弈模型和具体的模拟细节各不相同,但基本的模拟过程是类似的,这个模拟过程是分回合进行的,每个回合包含两步:

- (1) 网络中所有的参与者与其网络上的邻居进行博弈,并获得收益。每个参与者的收益为与其所有邻居发生博弈得到收益的总和。
- (2) 然后参与者将他的收益与他在网络上邻居的收益进行比较,按照一定规则改变自己的策略。

5.2.1 规则网络演化博弈

1. 规则网络——囚徒困境模型

Nowak 和 May 扩展了囚徒困境博弈模型,将参与博弈的个体置于二维格子上,每个

个体与直接相邻的4个邻居进行博弈,并累计收益,然后在更新策略时,一个个体与它的邻居比较本轮的收益,取收益最高者的策略作为下一轮博弈的策略,直到网络进入稳定状态为止。为了便于理论分析,Nowak采用了弱囚徒困境博弈,即令 $T=b>1, R=1, P=S=0$, 其中 b 表示合作收益。Nowak指出这种弱囚徒困境所得的演化结果与 $-1 \ll S < 0$ 时的结果相同。

Nowak发现引入空间结构后,通过演化,当 b 在一定范围内 ($1 \leq b \ll 2$) 时,合作者可以通过结成紧密的簇抵御背叛策略的入侵,如图5-1所示。

虽然这种合作簇并不固定,其形状也会随时间的改变而改变,但它并不会消亡,并且最终系统中合作者的比例(被称为合作频率,是衡量系统合作涌现程度的重要指标)会趋于稳定,如图5-2和图5-3所示。



图5-1 在方格上进行囚徒困境博弈得到的斑图

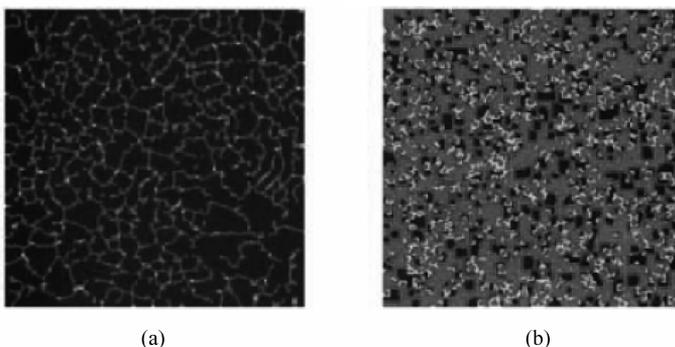


图5-2 200×200二维网格上的演化囚徒困境博弈形成的斑图

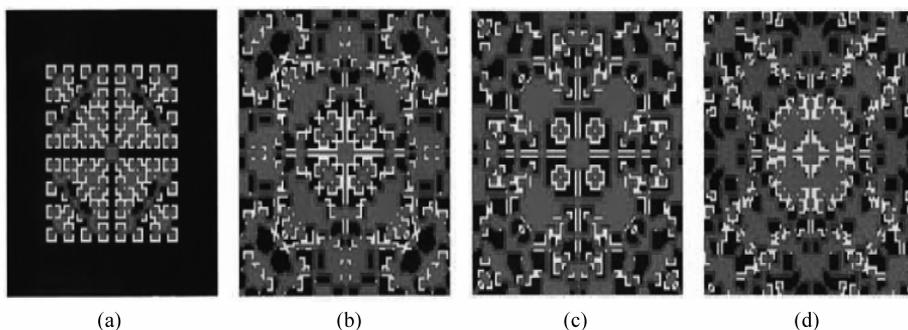


图5-3 99×99二维网格上的演化囚徒困境博弈形成的空间混沌

例如:对于上面提到的基于囚徒困境模型的规则网络博弈,基于费米函数的策略更新规则,利用平均场近似理论分析采取合作策略的个体的密度 ρ 随时间的演化。

$$\begin{cases} U_C = z\rho R + z(1-\rho)S = z\rho + z(1-\rho)c \\ U_D = z\rho T + z(1-\rho)P = z\rho b \end{cases} \quad (5-1)$$

$$P_{i \leftarrow j} = \frac{1}{1 + \exp[(U_i - U_j)/k]} \quad (5-2)$$

$$\frac{\partial \rho}{\partial t} = \rho(1-\rho)[W_{D \rightarrow C} - W_{C \rightarrow D}] = -\rho(1-\rho)\tanh\left(\frac{U_D - U_C}{2\kappa}\right) \quad (5-3)$$

他们发现当个体间的接触网络具有空间结构时,如方格网络,在囚徒困境博弈中合作行为能够出现并且稳定维持。其原因是在显著的空间结构效应下,合作者可以通过相互结成紧密的簇抵御背叛者的入侵。这个发现首次指出了网络结构对博弈演化起着重要的作用。

2. 规则网络——雪堆模型

Hauert 和 Doebeli 将博弈个体置于格子上,分别针对度为 3、4、6、8 的四种拓扑结构情况,根据雪堆博弈模型展开演化,如图 5-4 得出不一样的结论。

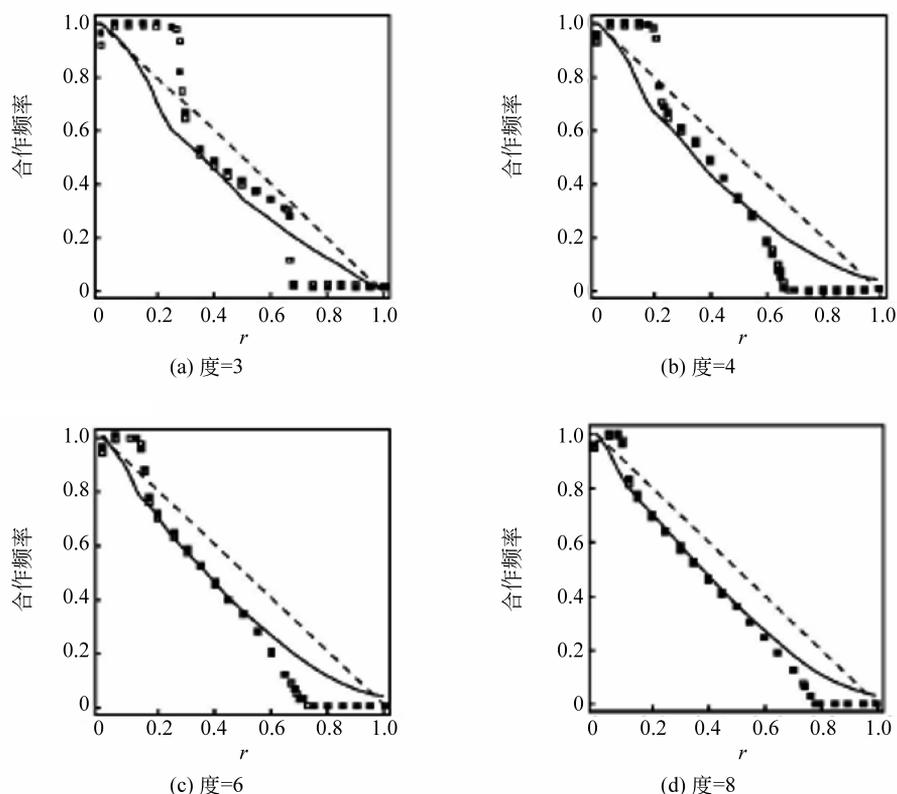


图 5-4 雪堆博弈模型演化结论

其中虚线表示模仿者动态下的演化稳定策略的合作频率,实线表示采用配对近似策略下的仿真结果,实心点和空心点分别代表同步演化和异步演化策略下的仿真结果。

规则格子上雪堆博弈的合作频率低于模仿者动态下的演化稳定策略,说明空间结构抑制了合作的产生。这是因为与囚徒困境的斑图不同,在雪堆博弈中合作者更容易聚成丝状簇(图 5-5)。

这就导致当损益比 r 较高时,背叛者容易入侵,使系统合作频率下降,这是雪堆博弈与囚徒困境在合作演化上的本质区别。

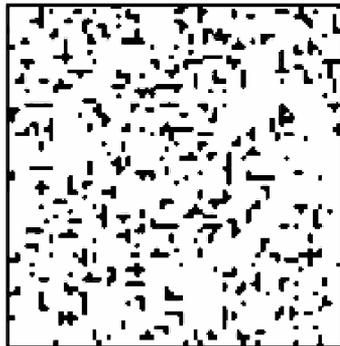


图 5-5 在方格上进行雪堆博弈得到的斑图

5.2.2 非规则网络演化博弈

1. 小世界网络——囚徒困境

Hauert 和 Szabo 基于规则方格,在保持度分布的前提下,对生成的均匀小世界网络和随机均匀网络做了研究。他们应用一种被广泛采用的随机演化策略:一个节点 x 更新策略的时候,随机在它的 k 个邻居中选择一个 y ,在下一轮中, x 以概率

$$p = 1 / \{1 + \exp[(P_x - P_y) / \kappa]\} \quad (5-4)$$

选择 y 本轮的策略作为自己下一轮的策略。上述公式来源于统计力学中的费米函数, κ 为环境中的噪声等不确定因素,设为 0.1, P_x 为 x 本轮的累积收益。研究表明:由于长程边的作用,均匀小世界网络和随机均匀网络比规则方格更利于合作的涌现:①小世界网络的异质性使其比规则格子更利于合作的涌现;②具有度异质特征的 WS 小世界网络比度均匀分布的小世界网络合作频率更高。

2. 小世界网络——雪堆博弈

Tomassini 等应用不同的演化规则作用在不同的重连概率的小世界网络上,细致地分析了小世界网络上的鹰鸽博弈。发现小世界网络的合作行为与博弈采用演化规则,收益比与小世界网络的重连概率息息相关。三者的交互作用使得空间结构时而促进合作的涌现,时而抑制合作的产生。

尚丽辉等针对现实生活中朋友关系网络的距离相关特性,研究了基于距离的空间小世界网络上的雪堆博弈(图 5-6),发现与规则网络相比,距离无关的小世界网络促进了合作的涌现;而距离相关的小世界网络中,幂指数增加导致了长程连接的减少和短程连接的增加,这使得网络在损益比较大时抑制合作的产生。

3. 无标度网络——囚徒困境

实际生活中很多网络(如因特网、航空网等)都具有无标度的特性,其节点的度分布满足某种幂律的特性。Santos 对比了规则格子、随机图、随机无标度网络和 BA 无标度网络对合作涌现的作用(图 5-7),认为由于无标度网络中节点之间的度存在极大的差异,合作行为容易在度大的节点之间传播,进而带动大量小度节点在无标度网络中传播。也就是说,无标度网络是目前最有利于合作涌现的网络结构。

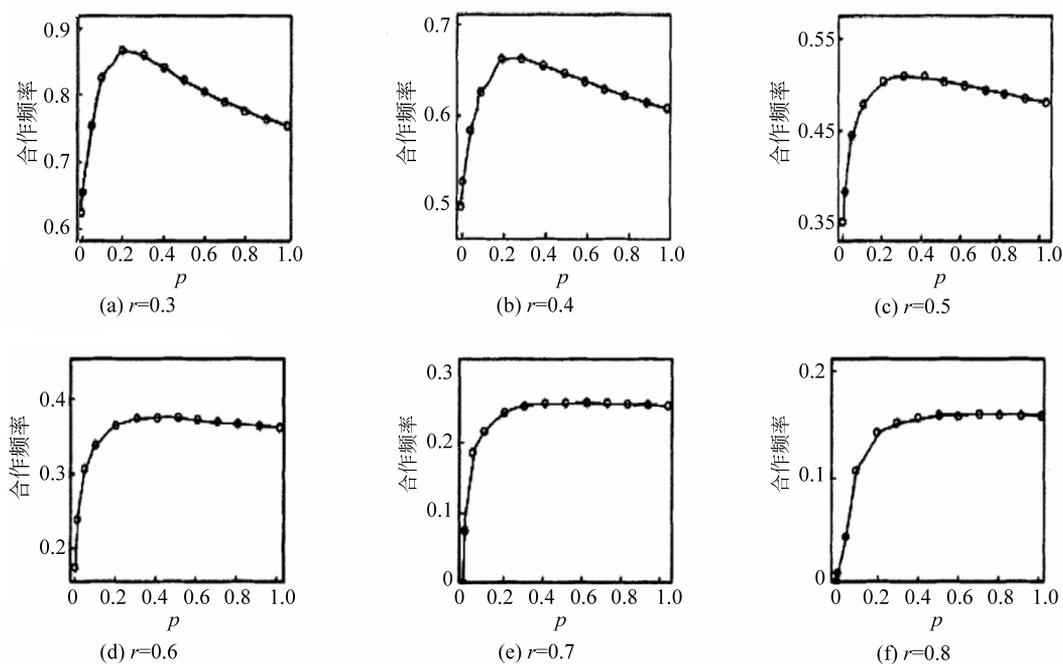


图 5-6 不同幂律指数下距离相关的小世界网络上的雪堆博弈合作曲线

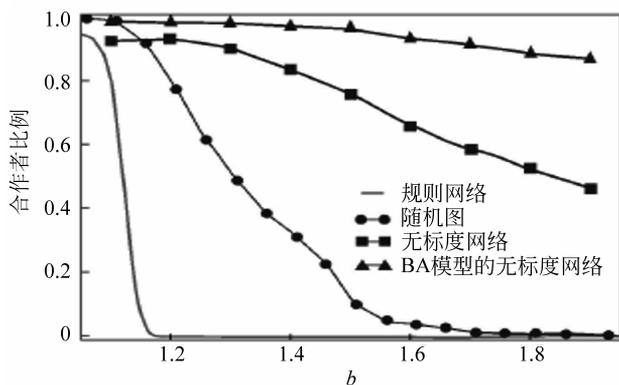


图 5-7 网络平均度为 4

Gomez-Gardenes 根据个体的稳定时的状态,将其划分为 3 类:纯策略者、纯背叛者和策略摇摆者。

4. 无标度网络——雪堆博弈

Santos 将研究无标度网络上囚徒困境的方法移植到雪堆博弈上,观察到类似于图 5-7 的现象,这说明无标度特性同样有利于雪堆博弈中合作的涌现。

通过对小规模网络(128 个节点)进行仿真,弱化了影响合作涌现的无标度网络其他统计学特性,着重突出了节点度的异质性的因素,再次验证了关于异质因素促进合作涌现

的一般性结论,指出无标度网络为研究演化博弈理论提供了统一的理论框架。

荣智海等研究了无标度网络上的扩展雪堆博弈(即一种可从雪堆博弈连续变化到囚徒困境的博弈),发现无标度网络异质性的增加使得合作的稳定性增强。而且对于相同的纯合作比例,纯背叛者比例增加,策略摇摆者比例减少。这说明越异质的网络,个体越倾向于选择稳定策略。

5. 度相关性对两类博弈的影响

Rong 等首先研究了无标度网络的度相关性对合作行为的影响。研究表明:在囚徒困境中,中性网络(即呈现度不相关特性的网络,如 BA 网络)的中心节点对于大度邻居与小度邻居的选择是最合理的,既与少量中心节点相连,又与他们共享很少量的邻居。所以其较之同配或异配网络的合作频率更高,最利于合作的涌现。

当无标度的网络结构呈现同配性质,即连接度大的节点倾向于和连接度大的节点建立连接时,由于中心节点和边缘节点(连接度一般较小)的“通信渠道”的减少,使得中心节点的合作策略难以传播出去,网络总体的合作频率呈现下降趋势。反之,如果无标度网络呈现度异配性时,中心节点之间的联系被切断,一方面不利于合作策略在中心节点之间扩散,抑制合作频率的上升;另一方面被孤立的中心节点可以和周围小度节点凝结成坚固的簇,即使背叛的诱惑非常大,也能有效抵御背叛策略的入侵。

对于雪堆博弈,越同配的网络其背叛者拥有越小的平均度,这说明与囚徒困境博弈类似,由于网络变得同配后中心节点对于小度节点的控制能力减弱,进行雪堆博弈的背叛者也主要集中在小度节点。异配网络当 r 较小时,雪堆博弈的合作频率会低于均匀混合状态的均衡频率。可见度相关性对于囚徒困境博弈的结论完全适用于雪堆博弈。如图 5-8 所示,两图中横坐标为背叛相对于合作的收益 b ,纵坐标为合作频率 ρ_c , r_k 为度相关性系数。

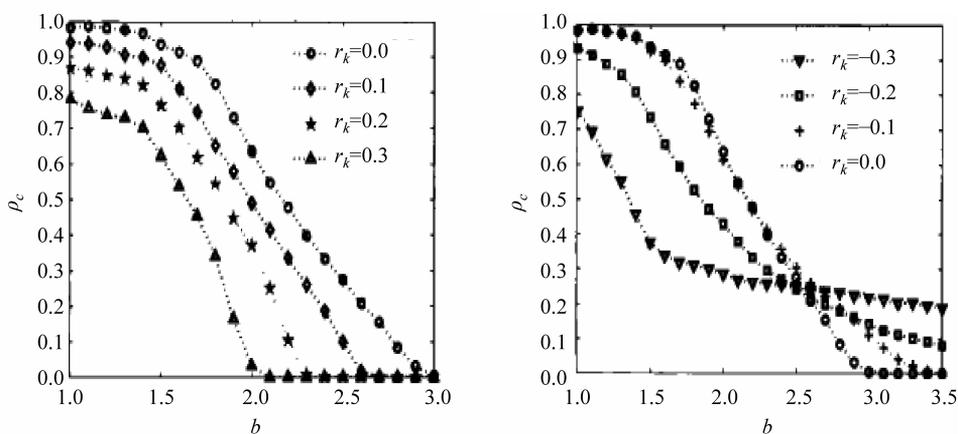


图 5-8 雪堆博弈的度相关性

5.2.3 多层网络演化博弈

关于网络人口中 PDG (Prisoner's Dilemma Game) 合作的研究也被扩展到多层系统的案例中, 其中一种方法是考虑每一层都由一组代理组成, 所以每个人在同一个网络层中同时与相邻的邻居一起玩。考虑两个相互关联的网络的情况, 其中每个代理根据邻居是否属于同一个网络, 或者对另一个网络进行了不同的游戏。特别地, 考虑两个可能的收益矩阵, 一个用于相同网络层节点之间的交互 \prod_{intra} , 采用的是相同的 E_q 形式, 下面是用于层间交互的。

$$\prod_{\text{intra}} = \begin{pmatrix} 1 & 0 \\ b > 1 & \epsilon < 0 \end{pmatrix} \quad (5-5)$$

因此, 如果 A_1 和 A_2 是两个网络的邻接矩阵, 则 $C_{12} = C_{21}$ 是包含层间连接的矩阵, 在 α 层的节点 i 的收益:

$$f_i^\alpha(t) = \sum_{j=1}^{N_1} (A_\alpha)_{ij} \cdot s_j(t)^T \prod_{\text{intra}} s_j^\alpha(t) \quad (5-6)$$

$$f_i^\alpha(t) = \sum_{j=1}^{N_i} (A_\alpha)_{ij} \cdot s_j^\alpha(t)^T \prod_{\text{intra}} s_j^\alpha(t) + \sum_{j=1}^{N_2} (C_{12})_{ij} \cdot s_j^\beta(t)^T \prod_{\text{intra}} s_j^\beta(t), \quad \alpha \neq \beta \quad (5-7)$$

s_i^α 是一个有两个组件的向量, 表示在 t 时刻 α 层节点 i 的策略, s_i^α 中读取 $(1, 0)^T$ 为合作, $(0, 1)^T$ 为背叛。因此, 当在同一层中与邻居进行博弈时, 节点会根据 PDG 获得收益, 而对于层间来说, 除了 2 个背叛者相互之间的作用外, 收益是相同的。当惩罚变成负的时, 这个矩阵最小可能收益也就是收益排行演变成 $T > R > S > P$, 后一个变化使 PDG 变成 Snowdrift game (SDG)。通过这种方式, 背叛者对合作者的利用, 不管他们在哪一层上, 都受到另一个网络层背叛者利益损失的阻碍。对于混合网和网络层, 可以找到在其中一层中合作占主导地位的极化状态, 而在另一层中则相反。

这里提到一个两个网络的互联系统, 但是考虑到内部和层间的交互是由 E_q 控制的, 也就是 PDG 的收益矩阵。在这种情况下, 我们把合作的出现看作层间连接的密度和每一层拓扑结构的作用, 一般来说, 由于层间连接的低密度, 系统的整体合作得到了增强。在这个系统中发现, 连接两个网络层的代理比只共享内部层的连接更有可能进行合作。此外, 尽管层间的相互联系促进了整个系统的合作, 但当考虑到单个层的层次时, 相互连接的效果可以在一层中减少合作, 这是由另一个的增长补偿的。

除了通过与其他网络节点博弈实现收益外, 另一种可能的方法是允许在层之间直接进行信息传输或交换。在最近的研究中, PDG 和 SDG 在不同层上分别实现, 提出了有偏见的模仿: 节点以概率 p 模仿层内的邻居在同一网络的策略, 在这样的框架下, 试图探索合作特质是如何随着偏差概率 p 的函数变化的。有偏差的概率揭示了多重效应, 它对人口结构的变化是有利的。从 $p=1$ 开始的偏差概率的轻微下降促进 PDG 中的合作行为, 相反, 这一减少损害了 SDG 群体中合作的发展, 一些文献也验证了定性有效的观测可以