复杂网络上的演化博弈研究

王旭文　汪秉宏

（中国科学技术大学近代物理系　230026）

DOI:10.13405/j.cnki.xdwz.2015.03.006

合作行为常见于生物系统和社会系统中，而博弈论为研究自私个体间的合作行为提供了一个有力的工具。博弈论又称对策论，指的是研究多个个体之间在特定条件制约下的对局里利用相关方的策略，而实施对应策略的学科。如美苏核武竞争博弈，如果双方都选择加入核武裁撤，那么双方的总收益是最高的，然而诱惑却是等待对方摧毁核武设施而自己保留一些。由此看出，合作能够获得较高的集体收益，然而个体由于自私通常会选择背叛以获得较高的自身收益。因此维持和促进合作行为的涌现一直是博弈论的一个挑战。其中一些典型的模型，如囚徒困境博弈（prisoner's dilemma game），雪堆博弈（snow-drift game）和公共物品博弈（public goods game）等被广泛研究。

一、博弈模型

1. 囚徒困境博弈

囚徒困境博弈例子：假定两个小偷 A 和 B 合伙犯罪，被捕后被独立审讯。如果两人彼此合作（合作 -C），即都拒不认罪，两人会被判 1 年徒刑。然而，如果 A 揭发 B（背叛 -D），B 拒不认 A 的罪行，则 A 将无罪释放而 B 将被判 5 年徒刑；如果 A、B 都揭发对方罪行，则二者均判 3 年。在这个模型中，相互合作可为全体带来最佳利益，但在不明对手策略的情况下，因为出卖同伙可为自己带来利益，同时同伙出卖自己会为同伙带来利益，因此彼此出卖虽违反最佳共同利益，反而是自己最大利益所在。抽象成更一般的模型：两个个体同时从合作（C）和背叛策略（D）中选择其一作为策略。如果两者都选择合作策略则二者皆获得相互合作收益 R；相反，两者都选择背叛策略则二者皆获得惩罚收益 P。如果一方选择合作而一方选择背叛，则合作者获得傻瓜收益 S，背叛者获得诱惑收益 T。这四个收益满足关系 $T>R>P>S$ 和 $2R>T+S$。博弈中一个常用概念即纳什均衡是指博弈个体单方面改变策略不能增加自己的收益。可以看出，囚徒困境模型中的纳什均衡是纯策略（D，D）。因为如果将自身策略由背叛转变为合作，则自己得收益 S，而对手得收益 T。因此单方面改变策略反而降低了收益。在标准的囚徒困境博弈中，个体的策略为完全合作或者完全背叛，所以我们称为纯策略；同时，在一些改进模型中，个体的策略为介于 0 到 1 之间的一个数 P，它代表个体合作的概率。0 代表个体合作的概率为 0，即完全背叛，而 1 代表个体合作的概率为 1，即完全合作。所以策略称之为混合策略，因为此时个体策略中既有合作成分也有背叛成分。

2. 雪堆博弈

雪堆博弈例子：在一个风雪交加的夜晚，两人相向而来，被一个雪堆所阻，假设铲除这个雪堆使道路通畅需要的代价为 c，如果道路通畅则带给每个人的好处量化为 b。如果两人一齐动手铲雪（即相互合作），则他们得相互合作收益 $R=b-c/2$；如果只有一人铲雪（即一方选择合作而一方选择背叛），虽然两个人都可以回家，但是背叛者逃避了劳动，他获得诱惑收益 $T=b$，而合作者获得傻瓜收益 $S=b-c$；如果两人都选择不合作（即相互背叛），两人都被雪堆挡住而无法回家，他们获得惩罚收益 $P=0$。这四个收益大小关系不同于囚徒困境博弈，其满足 $T>R>S>P$。雪堆博弈模型中的纳什均衡是纯策略（C，D）和（D，C）。假定个体 A 采用合作策略 C 而个体 B 采用背叛策略 D，此时个体 A 得收益 S，个体 B 得收益 T。如果个体 A 单方面将其策略改变为 D，则二者获得惩罚收益 P，而 $P<S$。所以单方面改变策略不能增加自身收益。同样，假定个体 A 采用背叛策略 D 而个体 B 采用合作策略 C，此时个体 A 得收益 T，个体 B 得收益 S。如果个体 A 单方面将其策略改变为 C，则二者获得相互合作收益 R，而 $R<T$。因此单方面改变策略同样不能增加自身

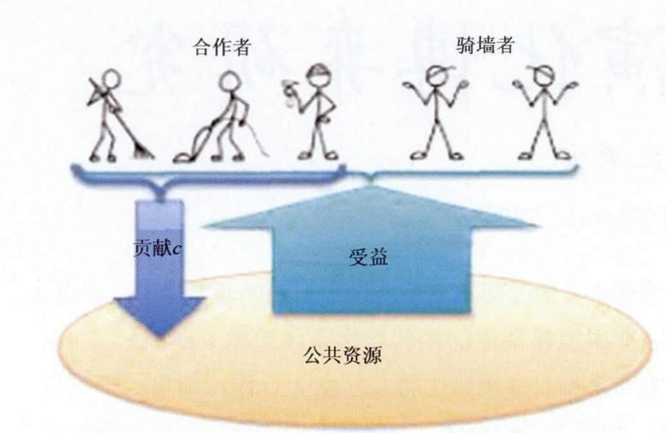

图 1 公共物品博弈

收益。

3. 公共物品博弈

在典型的公共物品博弈模型中，个体同时决定是否愿意对一个公共资源贡献 c 或者不贡献。如图 1，若个体贡献我们称之为合作者，反之，不贡献者为背叛者。初始的贡献会被放大 r 倍，进而，被放大后的总投资会被平分给每个个体，无论个体贡献与否。因此合作者和背叛者都能够从公共资源中受益，而背叛者由于没有贡献也能获得收益，因此称之为骑墙者。如在 5 个个体中有 3 个为合作者（贡献），而其余 2 个为背叛者（不贡献），则每个合作者的收益为：$\frac{3r}{5}-c$，而每个背叛者的收益为 $\frac{3r}{5}$。很明显，自私的个体会选择背叛策略，因为不投资也能获得收益。囚徒困境博弈和雪堆博弈是两个体之间的对－对作用，而公共物品博弈是群组行为，即多个体参与的行为。

二、复杂网络上的博弈

复杂网络是研究大规模复杂系统的有力工具。许多系统都可以简化为网络的形式，其中节点代表实际系统中的个体，连边代表两个个体之间能够发生作用。当博弈模型应用于网络上时，每个个体与其邻居进行博弈，其收益为与邻居博弈的收益总和。下一时步，个体通过一定策略更新规则进行策略学习，直至系统达到稳定。以图 2 正格子网络为例，如采用周期性边界条件，每个个体均有四个邻居。其中绿色代表合作者而红色代表背叛者。如中间的绿色节点为例我们进行收益计算。因为该节点自身为合作者，其邻居中有两个合作者和两个背叛者。根据囚徒困境博弈的收益

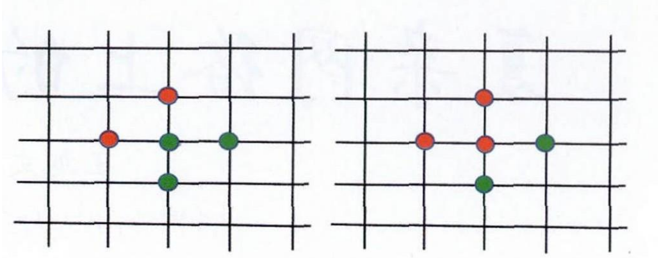

图 2 规则正格子网络上的博弈

矩阵，该节点从每个合作者邻居得收益 R，而从背叛者邻居得收益 S，因此其总收益为 $2R+2S$。下一时步该节点对自身收益和邻居收益进行比较，以一定的概率学习邻居的策略，如从合作者变为背叛者。

复杂网络上的博弈研究主要集中在三个方面。(1) 策略更新规则如何影响网络上的合作行为；（2）网络拓扑结构对合作行为的影响；（3）结构和策略更新同时作用对合作的影响。

1. 策略更新规则

（1）学习最优者：个体下一时间步采取邻居中收益最高者的策略作为其下一时步的策略。

（2）模仿优胜者：个体在策略更新时，同时参考那些收益比自身高的邻居的策略，以正比于他们所得收益的概率进行策略转变。

（3）配对比较：个体随机选择一个邻居进行收益比较，并以概率 W_{ij} 接受其策略。

$$W_{ij}=\frac{1}{1+\exp[-(U_j-U_i)/\kappa]},$$

其中 W_{ij} 是个体 i 采取随机选择的邻居 j 策略的概率。U_i 和 U_j 分别是个体 i 和 j 的收益，κ 是噪声项。当 κ 趋于无穷时，个体随机选择邻居策略，当 κ 趋于零且邻居收益大于个体收益时，个体确定性采取邻居策略。同时整个系统的策略更新方式也有同步更新和随机更新两种方式。同步更新指每个时间步所有个体都进行策略学习，随机更新是指仅随机选取的一个个体进行策略学习。

2. 规则网络上的囚徒困境博弈

对标准的囚徒困境博弈，其收益矩阵为：$A=\begin{pmatrix}R & S\\ T & P\end{pmatrix}$，为了简化模型且不失一般性，诺瓦克（M. A. Nowak）等将以上收益矩阵化简为：$A=\begin{pmatrix}1 & 0\\ b & 0\end{pmatrix}$，

因此参数b成了唯一参数，且$1<b<2$，代表背叛的诱惑。

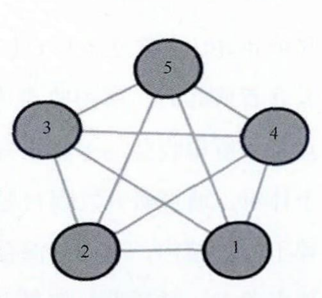

图 3　全连通网络示意图

当群体是均匀混合（任意两个节点作用的概率相同）或者全连通，也就是每个个体都与其余个体发生作用时，如图 3 所示。假定群体规模为N，则每个节点有N-1 个邻居。设群体中有N_C个合作者，那么，每个合作者收益为（N_C-1）R（因为总共有N_C个合作者，所以每个合作者有N_C-1 个合作邻居），背叛者收益为N_CT。可以看出只要N_C>0，背叛者收益必然大于合作者，因此合作者会学习背叛策略从而使得演化稳定时群体中个体均为背叛者。相比均匀混合群体，结构群体更加符合真实系统，因为个体更有可能与邻居进行博弈，而不是一些距离较远的个体，在结构群体中个体仅与其邻居进行博弈。在图 4 中，每个个体仅与上下左右四个邻居博弈，而不是所有个体。在结构群体中，由于网络互惠使得合作能够出现。1992 年诺瓦克和梅奥（R. M. May）研究了正格子网络上的囚徒困境博弈，发现在网络中合作者能够形成紧密的团簇以抵御背叛者的入侵。在此类团簇中，合作者能够获得较高的收益，因此合作者才能稳定存在。如图 4 所示，椭圆圈内 9 个合作者形成了紧密的团簇，位于这类团簇中间的节点，由于其邻居均为合作者，因此它们下一时步策略的学习对象也是合作策略，而位于合作团簇与背叛者交界处的合作者，由于其邻居中合作者居多可以有较高的收益，导致它们学习背叛策略的概率也很小。因此此类合作者团簇能够稳定存在，从而使得合作者得以生存。

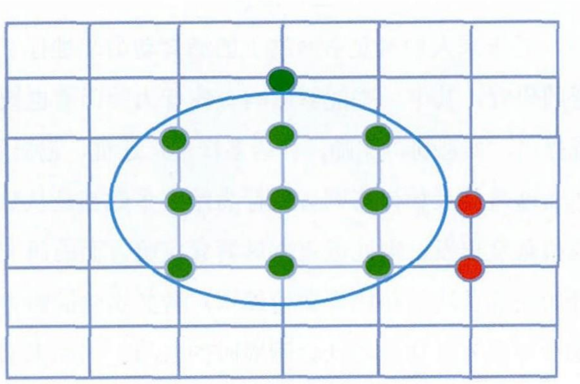
图 4　正格子网络上合作者的成簇现象

3. 无标度网络上的博弈

节点的度指网络中一个节点的邻居个数。对实际网络研究发现大多数网络的度分布遵循幂律分布，即$p(k) \propto k^{\gamma}$，其中k为度，$p(k)$表示度为k节点的比例。我们把度分布服从幂律分布的网络称为无标度网络。随后巴拉巴希（A-L. Barabási）和艾尔伯特（R. Albert）提出了基于线性择优生长机制的无标度网络模型。其规则为：（1）生长：初始时刻网络由m_0个节点开始生长，每个时间间隔添加一个具有m条边的新节点。（2）择优：当选择与新节点连接时，新节点连接到节点i的概率正比于其度k_i，即$\pi(k_i)=\dfrac{k_i}{\sum_j k_j}$。该模型生成的网络度分布服从幂律分布，且幂律指数为 3。随后桑托斯（F. C. Santos）等人研究了无标度网络上的囚徒困境博弈和雪堆博弈，如图 5 所示。模拟中囚徒困境博弈收益矩阵为：$A=\begin{pmatrix}1 & 0\\ b & 0\end{pmatrix}$，其中$b$代表背叛的诱惑；雪堆博弈的收益矩阵取$A=\begin{pmatrix}1 & -r\\ 1+r & 0\end{pmatrix}$，其中$r$代表相互合作的收益与代价之比。

可以看出虽然规则网络与无标度网络同为结构群体，然而相比于规则网络，无标度网络能够极大地促进合作行为的涌现。规则网络是通过每个节点以相同的连接规则构建而成，因此每个节点具有相同的度。如图 4 是一个规则的正格子网络，每个节点与上下左右四个邻居相连，度均为 4。而无标度网络因为度分布遵从幂律分布$p(k) \propto k^{\gamma}$，所以网络中节点度的分布范围较大，既有度很小的节点也有度很大的节点。图 5（a）是规则网络上的囚徒困境博弈。可以看出随着b的增加，合作者比例迅速降低。当b趋于 1.15 时合作者已经不能在规则网络中存在。而图 5（b）是无标度网络上不同平均度下的囚徒困境博弈，可以看出当b趋于 2 时，合作者比例依然保持在一个很高的数值。类似对比图 5（b）和图 5（d）的雪堆博弈模型

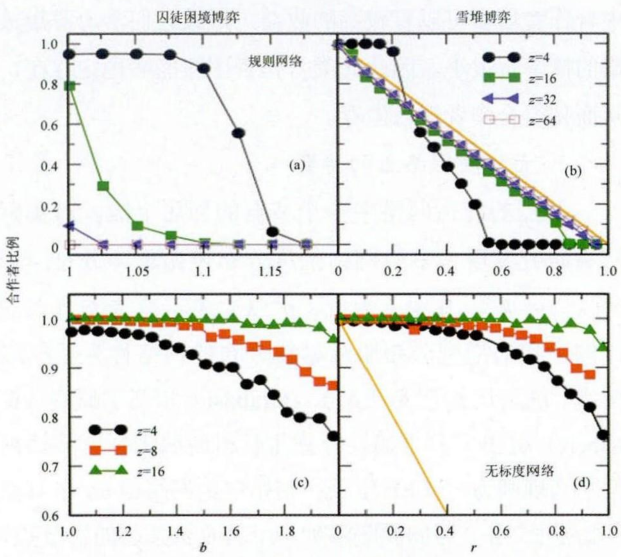

图 5　无标度网络上的博弈

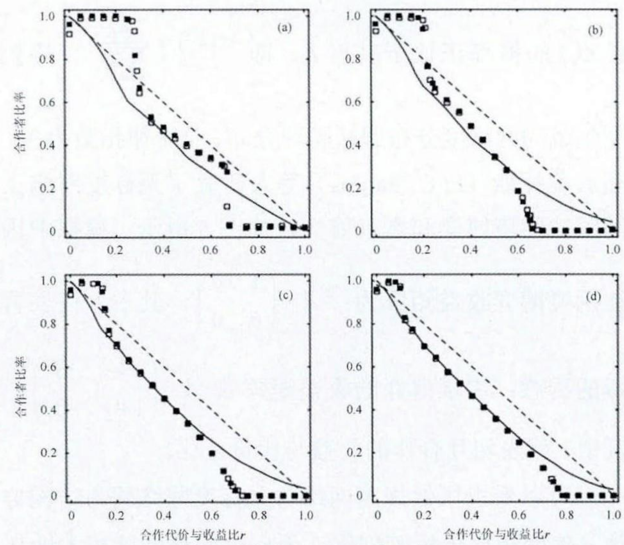

图 6　规则网络上的雪堆博弈

可以看出，随着 r 的增加，无标度网络上的合作水平依然是高于规则网络。这是由于无标度网络中存在一定比例的度极大的节点，这些节点能够带动这个系统中其他节点的合作。

4. 带有自愿者参与的囚徒困境博弈

上述博弈模型中，个体可以选择合作（C）或者背叛（D）中的一种策略。而萨博（G. Szabo）和奥埃尔（C. Hauert）等引入一种新的个体即孤立者到囚徒困境博弈中。此时收益矩阵变为：

$$M=\begin{pmatrix} 0 & b & \delta \\ 0 & 1 & \delta \\ \delta & \delta & \delta \end{pmatrix}$$

其中 $0<\delta<1$。孤立者与合作者、背叛者或者另外一个孤立者博弈时，所得收益皆为 δ。因此孤立者及其交互个体所得收益高于相互背叛个体对而低于相互合作个体对。通过研究发现自愿参与的合作影响极大地依赖于群体结构。对均匀混合群体，最后会趋于全同的孤立者态。循环支配使得正格子网络上能产生自组织斑图，而对随机规则网络上则会导致不同的震荡行为。

5. 规则网络上的雪堆博弈

类似于囚徒困境博弈，雪堆博弈的收益矩阵也可以简化为以下形式：

$$A=\begin{pmatrix} 1 & -r \\ 1+r & 0 \end{pmatrix},$$

其中 $0<r<1$，表示合作代价与合作的净收益的比率。

虽然空间结构有助于囚徒困境博弈中合作的产生（对均匀混合群体，合作者不能存在），但是 2004 年奥埃尔和多贝尔利（M. Doebeli）的工作则发现，对于雪堆博弈，空间结构效应会阻碍合作行为的产生。如图 6 所示，（a）（b）（c）（d）分别是三角格子（邻居个数为 3），正格子（邻居个数为 4），heagnal 格子（邻居个数为 6）和正方格子（邻居个数为 8）上合作者比率随参数 r 的关系。其中实线为均匀混合群体中合作者比率的变化。可以看出在 r 较小时，空间结构网络上的合作者比例比均匀混合群体要高，而对中间和数值较大的 r，其合作者比例小于均匀混合群体中的合作者比例。因此空间结构反而不利于雪堆博弈中合作行为的产生。

三、展望

近年来人们对复杂网络上的博弈动力学进行了广泛的研究，其中一些能够影响合作行为的因素也被相继提出，如惩罚，奖励，社会多样性，迁徙，记忆等。然而这些因素如何影响合作行为绝大多数都是从数值模拟角度出发，因此近年来对博弈实验方面的研究也开始受到关注。相比于数值模拟，博弈实验能够更好地反应真实个体面对社会困境时的决策，因此未来博弈论也将朝此方向发展。