

DOI: 10.7641/CTA.2018.70664

多智能体系统中基于演化博弈的群体状态控制

张建磊¹, 李智琦¹, 曹明^{2†}

(1. 南开大学 计算机与控制工程学院, 天津 300350;

2. 荷兰 格罗宁根大学 工程与技术学院, 荷兰 格罗宁根 9747AG)

摘要: 在对多智能体系统的研究中, 如何通过施加最少的控制来使某种策略在群体中占优是一个未解的难题. 本文借助演化博弈理论, 通过设置一定比例节点为指定策略作为控制手段, 分别研究了在无结构群体和随机规则网络群体中的策略演化情况. 在随机规则网络中, 本文进一步研究了在控制手段下, 一种新策略是如何演化并成功占据整个网络的. 结果表明在无结构的情况下, 强制策略对群体的影响受限于博弈的类型; 而在随机规则网络中, 在任何的博弈类型下, 只要给定足够多的强制策略就可以使其突破成功. 在理论分析的基础上, 本文进行了计算机仿真验证, 仿真结果与理论结果一致. 本文的结果揭示了如何对群体施加影响, 进而对群体中的个体状态进行控制.

关键词: 博弈论; 随机规则网络; 复制动力学

引用格式: 张建磊, 李智琦, 曹明. 多智能体系统中基于演化博弈的群体状态控制. 控制理论与应用, 2018, 35(5): 601 – 609

中图分类号: TP273 文献标识码: A

Control of states based on evolutionary games in multi-agent systems

ZHANG Jian-lei¹, LI Zhi-qi¹, CAO Ming^{2†}

(1. College of Computer and Control Engineering, Nankai University, Tianjin 300350, China;

2. Research Institute of Engineering and Technology, University of Groningen, Groningen 9747AG, Netherlands)

Abstract: When the evolution of strategies in the network is studied through game theory, it is still an open question as to how to make a strategy to take over the whole network by exerting minimum control. In this paper, we investigate the evolution of strategies in unstructured networks and random regular networks by forcing a certain proportion of nodes' strategies as a control. Furthermore, in the random regular network, we have studied how a new strategy invades and succeeds in dominating the whole network under control. The results show that the effect of the forcing strategy in the unstructured group depends strongly on the type of game; however, in the random regular network and under any game type, the intrusion can be successful as long as a sufficient number of forcing strategies are given, which is validated by our simulation results.

Key words: game theory; random regular network; replicator dynamics

Citation: ZHANG Jianlei, LI Zhiqi, CAO Ming. Control of states based on evolutionary games in multi-agent systems. *Control Theory & Applications*, 2018, 35(5): 601 – 609

1 引言(Introduction)

博弈论是一个处于理论研究前沿的热点学科, 一些经典的博弈模型经常在生物、经济、社会学等多个领域被讨论^[1-2]. 一些学者对各种博弈模型的优缺点做出了归纳, 其中主要研究的问题包括个体选择何种策略时才会获得最大的收益, 如文献[3-4], 博弈中如何达到纳什均衡等^[5-7]. 而在网络中的博弈则与此不同: 假设网络中每个节点都代表一个个体, 每个个体都使用一种策略, 网络中的连线则代表两个个体根据所使用的策略进行博弈; 当每个个体都随机选择其中

一个策略互相进行演化博弈时, 本文所关注的则是整个网络中某种策略所占的比例的变化^[8-9]. 这样, 如何使多种策略达到共存成为一个很重要的课题, 很多学者对其进行了深入的研究, 如文献[10-12]. 然而从实际应用的方面进行考虑, 在实际情况中, 有时也需要关注的问题是如何使某一种策略主宰整个网络.

在很多例子中, 都会遇到需要使某种策略独领大局的情况, 但是这并不容易达到, 博弈的类型会对其造成明显的限制. 例如在囚徒困境中^[13-15], 如果研究结果期待的是一个趋向于合作的局面, 那么这将会变

收稿日期: 2017-09-14; 录用日期: 2018-03-19.

†通信作者. E-mail: m.cao@rug.nl; Tel.: +31 50-3638712.

中国国家自然科学基金项目(61603199, 61603201), 荷兰国家自然科学基金项目(vidi-14134)资助.

Supported by the National Natural Science Foundation of China (61603199, 61603201) and the Dutch Technology Foundation (STW) (vidi-14134).

得尤为困难. 因为对于囚徒困境来讲, 背叛会使个体收益最大. 因此, 如果想要使合作策略在个体中盛行的话, 一些外部的刺激是必不可少的, 此类刺激可视作控制手段^[16-18].

如何以最小的外力来实现对网络中策略演化的控制, 这是一个非常有意义的课题^[19-20]. 本文通过强制一部分个体的策略不发生改变来对网络中的策略演化施加控制. 同样以囚徒困境为例, 如果强制一部分个体使用合作策略, 那么在不同的网络环境下, 这是否会对网络中的策略演化产生影响呢? 如果对合作有一定的促进作用的话, 那么至少需要多少的强制结点才能使合作策略充满整个网络? 在本文中, 分别研究了3种博弈环境: 囚徒困境、雪堆博弈和猎鹿博弈, 并在无结构和随机规则网络两种情况下分别对他们的演化结果进行了分析. 在数学模型的构建上, 本文选取了一些较为常见的方法^[21-22], 目的在于使本文的理论获得更大的通用性, 希望这些研究具有普适性.

在对于无结构情况下的研究, 本文主要采用了复制动态方程来模拟个体策略的演化^[23-25]. 主要方法为在使用某一种策略的个体中选择一部分, 并强制他们的策略不再发生改变, 研究不同比例的强制策略是否会对群体中的演化造成影响. 结果表明其演化结果因不同的博弈类型而不同, 受博弈类型的性质所限. 举例来说, 在一个个体都趋向于选择背叛策略的博弈类型中^[26-28], 即便是强制一部分节点为合作, 也依然无法改变合作策略的劣势. 而在一个个体趋向于选择相同策略的博弈类型中, 当强制的合作策略增加后, 群体中的其他个体也会倾向于选择合作策略, 合作策略就会增多^[29-30], 此时只要在群体中给定足够多的强制策略, 即可使合作策略主宰整个群体. 使用不同的博弈类型, 群体中策略演化的结果便不同.

而对于随机规则网络中的演化^[31-33], 本文则做了更进一步的研究, 主要问题为一个新策略是如何入侵并且占领整个网络的. 在这里所采用的策略更新方式为从邻居个体中随机选取一个来进行博弈. 博弈的收益通过费米函数计算出转换概率^[34-35]. 结果表明根据博弈类型的不同可以分为两种情况: 一种情况是所选策略内部博弈的收益大于所选策略与其他策略博弈的收益, 在本文中称其为优势策略; 另一种情况是所选策略内部博弈的收益小于所选策略与其他策略博弈的收益, 在本文中称其为劣势策略. 然而不论是那种情况, 在给定足够大的强制策略比例后均可以使所选策略充满整个群体. 此外, 本文在理论分析之外做了一些仿真, 仿真的结果与理论保持一致.

近些年来, 使用博弈理论来解决一些控制问题受到很多学者的青睐. 在对于群体策略与博弈领域的结合上, 一些面向实际问题的理论方法被提出, 并起到了积极作用^[36-37]. 如政府的引导性政策实行, 或者对

网络舆论的引导等, 如文献[38-39]. 这也证实了用博弈论的方法来研究某些控制问题是行之有效的. 本文通过引入强制策略来对群体的策略状态做出控制, 并通过博弈论来研究策略的演化过程. 希望这可以对解决一些群体行为控制的问题做出一些帮助.

2 模型介绍(Model introduction)

本文提出了两个模型分别用来描述无结构群体和随机规则网络中加入强制策略后的演化情况. 其中, 无结构群体演化模型通过复制动态方程来进行描述, 随机规则网络中的演化则通过计算策略转换概率的方法来进行研究. 在这里, 本文只考虑两种策略在群体中的演化情况, 分别用合作策略 C 和背叛策略 D 来表示. 这两种策略是竞争策略, 代表着个体可以进行的两种选择, 在任一时间内, 每个个体只能选择使用其中一种策略. 在演化初始阶段, 加入强制 C 策略来对群体施加影响, 然后研究 C 策略的比例随时间变化的情况. 其中, 强制策略的比例为 C_0 , 它不会随时间演化而改变, 其余 C 策略的比例设为 x_C , D 策略的比例设为 x_D , 这些非强制个体的策略演化情况将由收益矩阵的参数来决定. 从上述设定中可知, C 策略的总比例为 $P_C = C_0 + x_C$, 所有策略的比例之和应满足 $P_C + x_D = 1$. 在本文中, 收益矩阵分别采取3种常见的博弈类型的收益矩阵, 分别是雪堆博弈、囚徒困境和猎鹿博弈. 这里用一个统一的矩阵模型来进行表示:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array} \quad (1)$$

在上述矩阵中, a 代表 C 策略自身博弈所获得的收益, b 代表 C 策略与 D 策略博弈时所获得的收益. 类似的, c 代表 D 策略与 C 策略博弈所得到的收益, d 代表 D 策略自身博弈所得到的收益. 接下来则分别讨论博弈矩阵在两种不同结构的群体中的表现, 具体模型如下.

2.1 无结构群体中的演化(Evolution in unstructured groups)

在无结构群体内策略的演化中, 不需要考虑群体的结构特点. 这个过程可以用复制动态方程来进行描述. 根据收益矩阵, 可以得到每个策略在群体演化过程中各自的收益.

其中 C 策略的收益 S_C 为

$$S_C = a(x_C + C_0) + bx_D. \quad (2)$$

在上述方程中, 强制策略同样会产生收益且它的收益参数与 C 策略等同, 类似的, D 策略的收益 S_D 为

$$S_D = c(x_C + C_0) + dx_D. \quad (3)$$

群体中各个策略的平均收益为

$$\bar{S} = S_C(x_C + C_0) + S_D x_D. \quad (4)$$

这里,本文采用复制动力学方程来描述群体中的策略演化过程^[1].复制动力学方程是关于策略更新的确定性方程,其基本思想是收益高于整个群体平均收益的策略在群体中所占的比例会增加.根据上述方程,无结构群体中的演化模型能被建立:

$$\dot{x}_C = x_C(S_C - \bar{S}). \quad (5)$$

显然,对于上述模型中,当 $x_C = 0$ 时,则代表群体中只存在强制C策略与D策略. x_C 的取值范围为 $(0, 1 - C_0)$.可以通过求取模型的平衡点来研究C策略的演化情况,将模型(5)进行展开:

$$\dot{x}_C = x_C(1 - x_C - C_0)((a - b - c + d)(x_C + C_0) + b - d), \quad (6)$$

其中 $x_C = 0$ 和 $x_C = 1 - C_0$ 为边界平衡点,表示当 x_C 取得边界值时不再发生改变,此种情况不予考虑.现在研究其内部平衡点 x_0 :

$$x_0 = \frac{d - b}{a - b - c + d} - C_0. \quad (7)$$

可以看到, x_0 是一个与收益矩阵和强制C策略比例严格相关的值.下文将分别研究博弈矩阵和强制策略对平衡点的影响.这里,博弈的类型将通过收益矩阵中参数 a, b, c, d 的大小来表示.

2.1.1 雪堆博弈(Snowdrift game)

代入雪堆博弈的收益矩阵,它具有如下的大小关系:

$$c > a > b > d. \quad (8)$$

分别从 x_C 内部平衡点的数值和其稳定性来考虑,根据上述大小关系,可以得到

$$a - b - c + d < 0. \quad (9)$$

\dot{x}_C 在内部平衡点附近为减函数.因此其内部平衡点为稳定点,同时可以得到

$$0 < \frac{d - b}{a - b - c + d} < 1. \quad (10)$$

即群体中存在 x_C 的最大值,且当 $C_0 = 0$ 时取到最大值,此时全部的C策略均为非强制策略.而C策略的总值则保持不变,其大小与 C_0 无关,其值为

$$P_{C_{\max}} = \frac{d - b}{a - b - c + d}. \quad (11)$$

2.1.2 囚徒困境(Prisoner's dilemma)

考虑当使用囚徒博弈时,群体中策略的演化情况.收益应具有如下的大小关系:

$$c > a > d > b. \quad (12)$$

类似地,考虑其内部平衡点的数值与稳定性,此时

分两种情况考虑.第1种情况,如果

$$c - a > d - b, \quad (13)$$

那么可以得到

$$a - b - c + d < 0, \quad (14)$$

$$\frac{d - b}{a - b - c + d} < 0. \quad (15)$$

显然,在 x_C 的取值范围 $(0, 1 - C_0)$ 内不存在内部平衡点,且 \dot{x}_C 在区间内恒为负值.因此,在任意的初始条件下,最后群体中C策略的值都会仅剩强制策略 C_0 的值.

第2种情况,如果

$$c - a < d - b, \quad (16)$$

类似地,可以得到

$$a - b - c + d > 0, \quad (17)$$

$$\frac{d - b}{a - b - c + d} > 1. \quad (18)$$

与第1种情况的结果相同, x_C 的取值范围 $(0, 1 - C_0)$ 内不存在内部平衡点,且 \dot{x}_C 在区间内恒为负值.群体中C策略的值都会仅剩强制策略 C_0 的值,即 $x_C = 0$.

2.1.3 猎鹿博弈(Stag hunt game)

考虑猎鹿博弈的情况,给定收益的大小关系为

$$a > d > c > b, \quad (19)$$

此时有

$$a - b - c + d > 0, \quad (20)$$

$$0 < \frac{d - b}{a - b - c + d} < 1. \quad (21)$$

与雪堆博弈相反, \dot{x}_C 在内部平衡点附近为增函数,因此其内部平衡点为不稳定点.只要使其内部平衡点与 $x_C = 0$ 相重合,即可使在任意的初始条件下网络中充满C策略.这可以被认为是C策略入侵成功的一个典型例子,所研究的问题即当群体中只有D策略存在时,在初始时刻需要加入多少的强制C策略才可以使得新的策略C入侵并且充满整个群体.本文将在后续章节中着重讨论这个问题.此时,强制C策略的最小值可以通过如下方程得到

$$C_{0_{\min}} = \frac{d - b}{a - b - c + d}. \quad (22)$$

此时 $x_C = 0$,在 x_C 取值范围内恒有 $\dot{x}_C > 0$,直至达到 $x_C = 1 - C_0$,群体中充满C策略.

在无结构群体中,本文主要研究加入强制策略后群体中的策略演化情况,目的在于施加最小的控制,来达到某一种策略主宰网络的状态.分析表明,平衡状态受博弈类型的影响.需要指出的是受边界平衡点影响,在无结构群体中本文没有对群体策略的初始状态做出严格的限制,将主要的方向放在不同比例下两

种策略的演化情况上,即探讨内部平衡点的位置.而在接下来的随机规则网络中,本文加入了一些新的条件,对初始条件进行了更多的设置.如果说在无结构群体中研究的是两种策略的演化,并且一种策略如何才能“获胜”(占领整个网络),那么在随机规则网络中,则加入了一种策略从无到有的过程:新的策略是如何“入侵”并且“获胜”的.

2.2 随机规则网络中的演化(Evolution on random-regular network)

在这一章中,将会讨论加入强制策略后随机规则网络中的策略演化.在这里, C 策略将会被认为是一个新策略.主要研究在一个初始状态全是 D 策略的随机规则网络中,最少需要加入多少强制 C 策略才能使一种新的策略 C 成功入侵并且占领整个网络.从研究方法上来讲,两种网络群体是类似的,都是探讨加入强制策略后群体中策略演化结果.不过在随机规则网络中,本文将从另一个角度来对这个结果进行解释,将 C 策略考虑为入侵策略无疑将会是一项更有意义的研究.在随机规则网络中,假设每个节点的度数都为 n ,所有的节点的策略均为 D .在初始时刻,强制 x_c 比例的节点策略改变为强制策略 C ,并且在之后的演化过程中不会再发生改变.其他节点的策略更新方式为:从邻居中随机选取一个节点,将他的收益与自身的收益进行比较(节点的收益通过收益矩阵来计算),然后通过费米函数来计算自身策略的转换概率.收益矩阵及其他参数同无结构群体.

根据收益矩阵则可以得到每个 C 策略个体和 D 策略个体的平均收益 \bar{S}_c 和 \bar{S}_d .有

$$\bar{S}_c = an(x_c + C_0) + bn x_d, \quad (23)$$

$$\bar{S}_d = cn(x_c + C_0) + dn x_d. \quad (24)$$

然后通过费米函数计算策略的转换概率(此概率受到放大系数 k 的影响).当一个 C 策略与一个 D 策略进行博弈时, C 策略向 D 策略转换的概率为

$$\bar{u}_{c \rightarrow d} = (1 + e^{-\alpha P_c - \beta})^{-1}. \quad (25)$$

同样,当一个 D 策略与一个 C 策略进行博弈时, D 策略向 C 策略转换的概率为

$$\bar{u}_{d \rightarrow c} = (1 + e^{\alpha P_c + \beta})^{-1}, \quad (26)$$

其中:

$$\alpha = nk(c - a - d + b), \quad (27)$$

$$\beta = (d - b). \quad (28)$$

根据上述方程中每个个体的策略转换概率,可以得到整个网络中 C 策略的演化情况,具体模型如下所示:

$$\dot{P}_c = x_d P_c \bar{u}_{d \rightarrow c} - x_c P_c \bar{u}_{c \rightarrow d}. \quad (29)$$

它表示每一时刻总体 C 策略的变化率等于 C 策略的增加量减去 C 策略的减小量.这时可以通过对上述方程求解来得到 C 策略的演化趋势.其结果将通过两则定理来说明.

定理 1 在随机规则网络中,强制一定比例的 C 策略不再发生改变.如果 C 策略内部博弈的收益小于 D 策略与 C 策略博弈的收益,即博弈矩阵中 $a < c$,那么 $C_0 = 1 - e^{nk(a-c)}$,即为使整个网络中充满 C 策略的最小强制节点比例.

此定理的证明可以通过以下步骤形成:

步骤 1 将模型(29)进行展开:

$$\dot{P}_c = (1 - P_c)(P_c \bar{u}_{d \rightarrow c} - P_c \bar{u}_{c \rightarrow d} + C_0 \bar{u}_{c \rightarrow d}). \quad (30)$$

同样的,不考虑其边界平衡点,通过其内部平衡点附近的函数变化来研究网络中的策略演化趋势.其内部平衡点的值可以通过如下化简得到:

$$P_c \bar{u}_{d \rightarrow c} - P_c \bar{u}_{c \rightarrow d} + C_0 \bar{u}_{c \rightarrow d} = 0, \quad (31)$$

$$e^{-\alpha P_c - \beta} - 1 + \frac{C_0}{P_c} = 0. \quad (32)$$

步骤 2 设上式左侧为函数,并多次求导以观察平衡点附近函数的变化:

$$F(P_c) = e^{-\alpha P_c - \beta} - 1 + \frac{C_0}{P_c}. \quad (33)$$

可以看出, $F(P_c)$ 对于 C_0 是增函数.这里将其设置为无穷小,强制策略几乎为0,可以看做是仅对网络中 C 策略的演化产生一个激励.然后对其进行多次求导:

$$F'(P_c) = -\alpha e^{-\alpha P_c - \beta} - \frac{C_0}{P_c^2}, \quad (34)$$

$$F''(P_c) = \alpha^2 e^{-\alpha P_c - \beta} - \frac{2C_0}{P_c^3}. \quad (35)$$

显然可以得到

$$F''(P_c) > 0. \quad (36)$$

由前文可知 P_c 取值范围为 $(P_0, 1)$,根据上述方程可以得到 $F'(P_c)$ 必定为一个增函数.此时考虑 $F'(P_c)$ 的边界值,可以得到

$$\begin{cases} F'(C_0) = -\alpha e^{-\beta} - \frac{1}{C_0}, \\ F'(1) = -\alpha e^{a-c}. \end{cases} \quad (37)$$

因为假设 C_0 为无穷小,所以可以得知 $F'(C_0) < 0$.

步骤 3 分情况讨论结果.此时分两种情况考虑.第1种情况为

$$c - a < d - b, \quad (38)$$

即 $-\alpha > 0$,此时有 $F'(1) > 0$.在定义域 $(P_0, 1)$ 内,函

数 $F(P_C)$ 有最小值存在,考虑函数 $F(P_C)$ 的边界值,则有

$$\begin{cases} F(C_0) > 0, \\ F(1) < 0. \end{cases} \quad (39)$$

显然,在定义域内函数 $F(P_C)$ 有唯一零点存在,且零点附近函数为减函数.由于本文在随机规则网络中考虑的是新策略的入侵过程, C 策略的比例从 P_0 开始增大,所以使得 C 策略扩散至整个网络的最少强制节点数即为使得函数 $F(P_C)$ 零点为1的值,即

$$\begin{cases} F(P_C) = 0, \\ P_C = 1. \end{cases} \quad (40)$$

第2种情况为

$$c - a > d - b, \quad (41)$$

即 $-\alpha < 0$,此时有 $F'(1) < 0$,在定义域 $(P_0, 1)$ 内,函数 $F(P_C)$ 为减函数,且最小值为 $F(1)$.考虑函数 $F(P_C)$ 的边界值如方程组(39)所示,结果与第1种情况相同,函数 $F(P_C)$ 有唯一零点存在.使得 C 策略扩散至整个网络的最少强制节点数即为使得函数 $F(P_C)$ 零点为1的值,计算方式同上.

因此无论是哪一种情况,都存在唯一零点,且是稳定的(减函数).根据所得方程组(40),展开并代入参数,计算使整个网络中充满 C 策略的最小强制节点比例.可以得到

$$C_{0\min} = 1 - e^{nk(a-c)}. \quad (42)$$

证毕.

类似地,根据另一种情况,这里给出第2则定理用来概括.

定理 2 如果 C 策略内部博弈的收益小于 D 策略与 C 策略博弈的收益,即博弈矩阵中 $a > c$:

若 $a - c < b - d$,那么使得 C 策略将充满整个网络的最少强制策略比例 C_0 为无穷小.

若 $a - c > b - d$,那么使得 C 策略将充满整个网络的最少强制策略比例 C_0 可通过如下方程组求得:

$$\begin{cases} F'(P_C) = 0, \\ F(P_C) = 0. \end{cases} \quad (43)$$

此定理的证明可以通过以下步骤形成:

步骤 1 将模型(29)进行展开,与定理1中相同.

步骤 2 建立函数与求导,同定理1中步骤2.

步骤 3 最后,分情况讨论结果.类似的,此时分两种情况考虑.

第1种情况为

$$a - c < b - d, \quad (44)$$

即 $-\alpha < 0$,此时有 $F'(1) < 0$,在定义域 $(P_0, 1)$ 内,函数 $F(P_C)$ 在 $F(1)$ 处取得最小值.然而考虑函数 $F(P_C)$ 的边界值,则有

$$\begin{cases} F(C_0) > 0, \\ F(1) > 0. \end{cases} \quad (45)$$

可知导数 \dot{P}_C 在定义域内恒大于0,因此只要给定无穷小的 P_0 ,即可在演化中使得 C 策略将在网络中占据全面优势.在足够长的时间内都会使整个网络中的所有策略转化为 C 策略.

第2种情况为

$$a - c > b - d, \quad (46)$$

即 $-\alpha > 0$,此时有 $F'(1) > 0$.在定义域 $(P_0, 1)$ 内,函数 $F(P_C)$ 有最小值存在.而其边界值如方程(42)所示.显然,唯一使得强制策略 P_0 扩散到整个网络的条件为 $F(P_C)_{\min} > 0$.因此所需强制策略最小值可通过如下方程组来确定:

$$\begin{cases} F'(P_C) = 0, \\ F(P_C) = 0. \end{cases} \quad (47)$$

证毕.

尝试对上述方程组进行求解,可以得到一个关于 P_C 的方程.如下所示:

$$(\alpha P_C - 1)e^{-\alpha P_C - \beta} + 1 = 0. \quad (48)$$

此时无解析解,但可求得数值解.具体的数值解的例子将在下一章节中举出,根据上式所得的 P_C 值,代入到

$$C_{0\min} = P_C(1 - e^{-\alpha P_C - \beta}), \quad (49)$$

即可得到使得 C 策略入侵成功,即充满整个网络所需的最少的强制策略比例.接下来将会将具体的博弈代入到随机规则网络的模型中分析,一些计算与仿真的曲线将在下一章节中展示.

3 仿真结果(Simulation results)

在这一节中,本文将具体的博弈类型代入到随机规则网络中进行分析.在研究每种博弈类型时,都会给出一些实例进行展示.在理论的结果中,演化达到平衡时,强制策略比例 C_0 与 C 策略的总比例 P_C 的关系如图1(a)(c)(e)所示.此外,本文还根据实例在随机规则网络中做了一些仿真,仿真的结果将展示在图1(b)(d)(f)中.接下来本文将根据3种博弈类型依次进行分析.

3.1 雪堆博弈(Snowdrift game)

首先给出雪堆博弈的收益矩阵,不同于无结构群体时的大小关系,这里将给出更具体的矩阵参数,如下所示:

$$\begin{array}{c} C \quad D \\ C \left(\begin{array}{cc} b - \frac{c}{2} & b - c \\ b & 0 \end{array} \right). \end{array} \quad (50)$$

此时, 矩阵参数的大小关系符合定理1中所示. 固定参数 c 的值, 将参数 b 施以不同的数值. 强制策略比例与平衡时 C 策略的总比例的关系如图1(a)所示.

在图1(a)(c)(e)中, 将网络中平均度设置为 $n = 4$ 比例系数 $k = 0.2$, 固定参数 $c = 2$, 分别改变参数 b 为 $b = 3, b = 4, b = 5$, 观察 C 策略的变化情况.

可以看到随着强制策略比例的增大, 平衡状态下 C 策略的总比例 P_C 不断上升, 直至达到 $P_C = 1$. 参数 b 的增大会使同强制策略比例下演化出的 C 策略增多, 但是并不会改变使得 C 策略充满整个网络的最少 C_0 值. 这从本文的计算结果中可以体现出来. 在方程(42)中代入雪堆博弈的收益矩阵, 可以得到

$$C_{0_{\min}} = 1 - e^{-\frac{nk}{2}c}. \quad (51)$$

这是所需的最小强制节点比例, 它与参数 b 无关. 不过在实际的网络中与此略有不同, 因为在本文前述的模型中计算的收益是平均值. 其在仿真中的结果如图1(b)所示.

本文根据图1(a)(c)(e)中的例子, 以相同的参数在随机规则网络中进行了仿真, 用以验证模型以及理论计算的正确性. 首先构建一个包含2500个节点的随机规则网络, 在这个网络中每个节点随机与4个节点进行连接, 与理论模型保持一致. 每个节点都有0和1两种状态用来表示 C 和 D 两种策略. 在动态的演化过程中, 每个回合结束时, 每个节点随机选取一个邻居节点进行博弈, 根据算例中给定的收益矩阵计算当前回合节点所获得的收益(固定收益矩阵中参数 $c = 2$, 分别改变参数 b 为 $b = 3, b = 4, b = 5$), 并根据费米函数计算节点策略更新的概率(同样取比例系数为 $k = 0.2$), 然后进行策略的更新. 这里设定运行的步数为10000步, 即博弈10000个回合, 然后取最后500步的平均值为结果. 为了保证仿真的正确性, 以上所有流程重复50次取平均值为仿真结果. 然后将仿真结果与理论模型的结果作为对比放在图1(b)(d)(f)中.

对比图1(a)与图1(b), 可以看到, 在定性的结果上, 模型与仿真保持一致. 同样的初始条件下, 合作的收益越高越会促进 C 策略所占的比例. 随着外加控制 C_0 的增大, 最终 C 策略占比会达到1.

3.2 囚徒困境(Prisoner's dilemma)

类似的给出囚徒困境的收益矩阵如下:

$$\begin{array}{c} C \quad D \\ C \left(\begin{array}{cc} b - c & -c \\ b & 0 \end{array} \right). \end{array} \quad (52)$$

同雪堆博弈, 矩阵参数的大小关系符合定理1中所示. 固定参数 c 的值, 将参数 b 施以不同的数值. 强制策略比例与平衡时 C 策略的总比例的关系如图1(c)所示.

在图1(c)中, 可以看到囚徒博弈的结果与雪堆博弈有些相似, 不过使囚徒博弈中合作策略占领整个网络的难度要远远大于雪堆博弈. 而参数 b 的改变也并不会影响网络中策略的演化情况. 将其代入到模型(42)中, 可以得到

$$C_{0_{\min}} = 1 - e^{-nk}. \quad (53)$$

此为所需的最少强制策略比例. 同样, 本文使用囚徒博弈在随机规则网络中做了一些仿真, 这将被展示在图1(d)中. 与图1(c)相似的是, $b = 3, b = 4, b = 5$ 这3条线依旧在图中重合. 显然, 理论与仿真图结果定性的保持一致. 在囚徒困境中, 即使是施加了外力控制后, 达到 C 策略统治整个网络依然很艰难. 参数 b 的改变并不会对结果造成影响. 总体来说, 这与雪堆博弈结果近似, 均符合定理1中所示环境. 随着 C_0 的增大, P_C 随之增加, 直至占领整个网络.

在雪堆博弈和囚徒博弈这两种情况下, 合作策略在本文中被称为劣势策略, 随着 C_0 的增大, P_C 的上升过程是持续而缓慢的.

3.3 猎鹿博弈(Stag hunt game)

同样的方法对猎鹿博弈进行分析, 其收益矩阵如下:

$$\begin{array}{c} C \quad D \\ C \left(\begin{array}{cc} b & 0 \\ b - c & b - \frac{c}{2} \end{array} \right). \end{array} \quad (54)$$

在猎鹿博弈的收益矩阵中, 其参数的大小关系符合定理2中所示. 所需最小强制策略可以通过方程组(47)来计算. 同样改变其参数, C_0 与 C 策略的总比例 P_C 的关系如图1(e)所示. 参数的设定与之前图像保持一致.

在图1(e)中, 随着 C_0 的增大, P_C 缓慢增加. 但是存在着一个关于 C_0 的阈值, 当 C_0 达到阈值时, 网络中的合作策略增加至充满网络. 这个阈值可以通过代入到方程组(47)中求得数值解. 此外, 与雪堆博弈相反的是, 随着参数 b 的增大, 使 C 策略的总占比达到1的要求变得更高, 需要更多的强制策略 C_0 才会使 C 策略独占整个网络. 仿真的结果与此相似, 其将被展示在图1(f)中.

在图1(f)中, 参数的设置与之前相同. 可以看到, 在仿真图中同样存在突变的现象, 且参数 b 越大, 使得 C 策略占领整个网络的要求越高. 这与本文的理论模型保持一致.

此种情况下, 合作策略在本文中被称为优势策略, 只要外力影响达到某个阈值, 便可以形成突变.

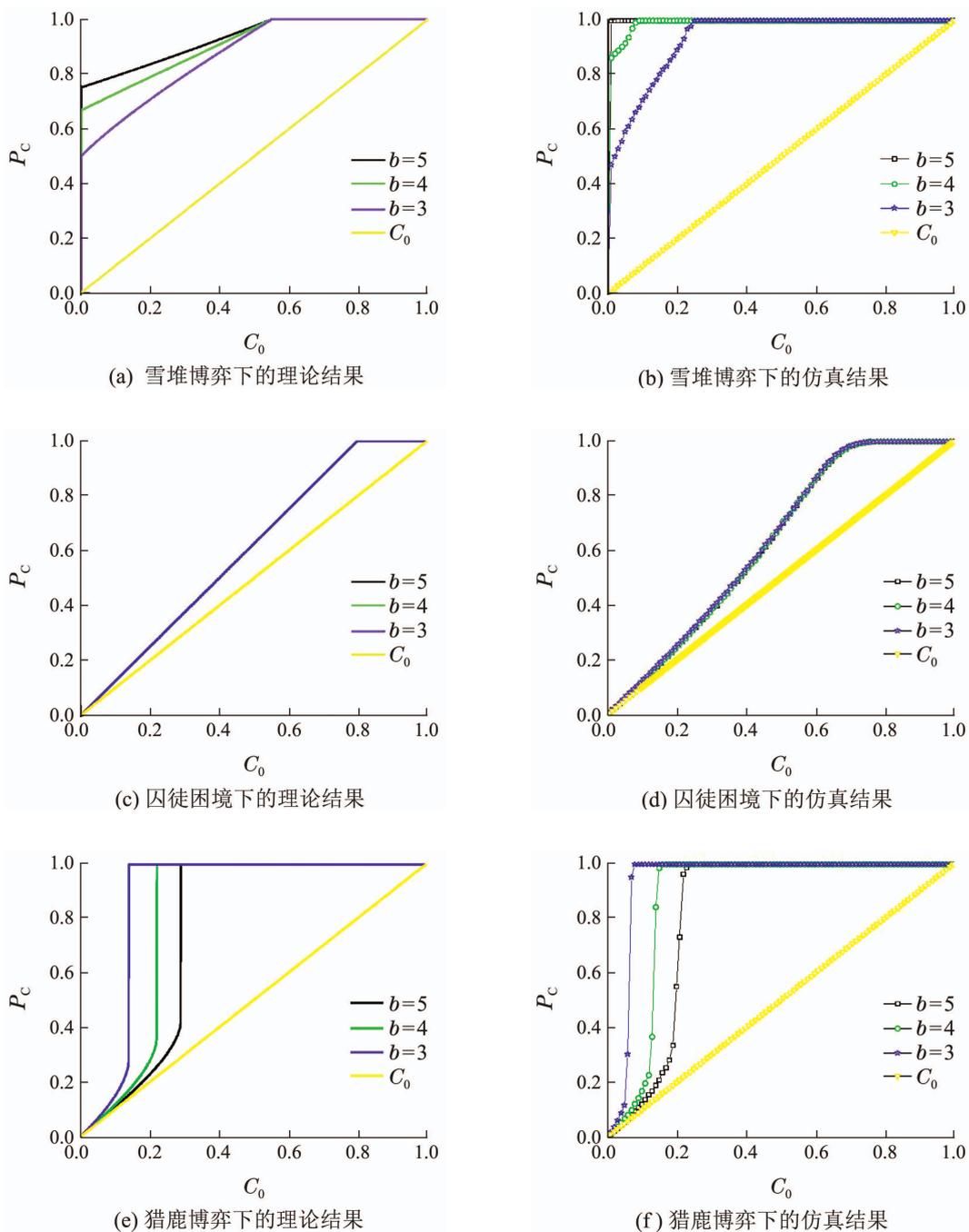


图 1 3种博弈的理论和仿真结果

Fig. 1 The theoretical and simulation results of three games

总体来说, 对于3种博弈情况, 在随机规则网络中, 都会存在一个确切的 C_0 值, 可以使得强制策略扩散至整个网络. 仿真的结果与本文的理论模型定性上保持一致. 这些结果显示了本文的分析和计算的有效性. 对比无结构的理论结果, 这些结果也说明群体结构对策略的演化产生了影响.

4 结论(Conclusions)

本文通过控制一部分节点来对网络的策略演化施加影响, 目的在于以最小的外力对网络中策略演化进行控制, 以实现某种策略主宰整个网络. 在文中, 提出

了两个模型分别来描述无结构群体中的演化和随机规则网络中的演化. 结果表明无结构群体中的演化受博弈类型自身特性的影响; 而随机规则网络内, 在足够大的外力影响下(强制节点比例) 可以使某一种策略完全主宰整个网络.

在无结构群体中, 如果采用雪堆博弈作为博弈手段, 那么不管强制节点的比例为多少, 此策略在群体中将会有一个恒定的最大值; 如果采用囚徒困境来进行博弈, 那么强制合作策略的数目对策略的演化并无帮助作用, 群体中将会只有强制策略个体保持合作策

略;而在猎鹿博弈中,只要强制策略的比例足够大,则可以带动整个群体中的节点全部转化为此策略。

在随机规则网络中,任何博弈类型下都可以通过改变强制策略的数目来使此策略充满群体,但是演化过程分为两种.根据强制策略在博弈中所占优劣,本文分别称之为优势策略或者是劣势策略.若其为优势策略,只要给一个较小的激励使其达到突变的阈值(强制策略)就可以使整个网络中充满所需策略;而若其为劣势策略,则需要一个持续的较大的激励才可以达到此策略主宰整个网络.此外,本文在随机规则网络中做了一些仿真.仿真结果在定性上与理论结果保持一致,这对本文结论的正确性做出了验证。

如何通过施加较小的控制来使某种策略独占网络是一个很有意义的话题,本文通过对施加强制策略的研究对此进行了一些解释,这对现实生活中研究通过外力控制一些群体中的策略演化有着重要的意义.本文的结果也显示了结构化群体中的策略演化和无结构群体中策略演化的不同.希望本文中的结果可以为此提供一些启示与灵感。

参考文献(References):

- [1] COLMAN A M. *Game Theory and Its Applications: In the Social and Biological Sciences* [M]. Oxford: Psychology Press, 2013.
- [2] FRIEDMAN D. On economic applications of evolutionary game theory [J]. *Journal of Evolutionary Economics*, 1998, 8(1): 15 – 43.
- [3] AKCAY E, ROUGHGARDEN J. The evolution of payoff matrices: providing incentives to cooperate [J]. *Proceedings of the Royal Society of London B: Biological Sciences*, 2011, 278(1715): 2198 – 2206.
- [4] CHENG H, DAI Q, LI H, et al. Payoff-related migration enhances cooperation in the prisoner's dilemma game [J]. *New Journal of Physics*, 2011, 13(4): 043032.
- [5] BERG J, WEIGT M. Entropy and typical properties of Nash equilibria in two-player games [J]. *EPL (Europhysics Letters)*, 1999, 48(2): 129.
- [6] BJÖRNERSTEDT J, WEIBULL J W. Nash equilibrium and evolution by imitation [R]. *IUI Working Paper*, 1994.
- [7] BRAUN D A, ORTEGA P A, WOLPERT D M. Nash equilibria in multi-agent motor interactions [J]. *Plos Computational Biology*, 2009, 5(8): e1000468.
- [8] ALLESINA S, LEVINE J M. A competitive network theory of species diversity [J]. *Proceedings of the National Academy of Sciences*, 2011, 108(14): 5638 – 5642.
- [9] ZHANG J, WEISSING F J, CAO M. Fixation of competing strategies when interacting agents differ in the time scale of strategy updating [J]. *Physical Review E*, 2016, 94(3): 032407.
- [10] ZHANG J L, CHEN Z Q. Contact-based model for strategy updating and evolution of cooperation [J]. *Physica D: Nonlinear Phenomena*, 2016, 323: 27 – 34.
- [11] ARCHETTI M, SCHEURING I. Coexistence of cooperation and defection in public goods games [J]. *Evolution*, 2011, 65(4): 1140 – 1148.
- [12] MASUDA N, KONNO N. Networks with dispersed degrees save stable coexistence of species in cyclic competition [J]. *Physical Review E*, 2006, 74(6): 066102.
- [13] ALONSO J, FERNANDEZ A, FORT H. Prisoner's Dilemma cellular automata revisited: evolution of cooperation under environmental pressure [J]. *Journal of Statistical Mechanics: Theory and Experiment*, 2006, DOI: 10.1088/1742-5468/2006/06/P06013.
- [14] ALONSO-SANZ R. Memory boosts cooperation in the structurally dynamic prisoner's dilemma [J]. *International Journal of Bifurcation and Chaos*, 2009, 19(9): 2899 – 2926.
- [15] ALONSO-SANZ R. Spatial order prevails over memory in boosting cooperation in the iterated prisoner's dilemma [J]. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2009, 19(2): 023102.
- [16] DU W B, ZHOU H, LIU Z, et al. The effect of pinning control on evolutionary prisoner's dilemma game [J]. *Modern Physics Letters B*, 2010, 24(25): 2581 – 2589.
- [17] GOLUBSKI A J, KLAUSMEIER C A. Control in mutualisms: combined implications of partner choice and bargaining roles [J]. *Journal of Theoretical Biology*, 2010, 267(4): 535 – 545.
- [18] RAMAZI P, CAO M. Asynchronous Decision-making dynamics under best-response update rule in finite heterogeneous populations [J]. *IEEE Transactions on Automatic Control*, 2018, 63(3): 742 – 751.
- [19] RAMAZI P, RIEHL J, CAO M. Networks of conforming or nonconforming individuals tend to reach satisfactory decisions [J]. *Proceedings of the National Academy of Sciences*, 2016, 113(46): 12985 – 12990.
- [20] RIEHL J R, CAO M. Towards optimal control of evolutionary games on networks [J]. *IEEE Transactions on Automatic Control*, 2017, 62(1): 458 – 462.
- [21] CRESSMAN R, TAO Y. The replicator equation and other game dynamics [J]. *Proceedings of the National Academy of Sciences*, 2014, 111(S3): 10810 – 10817.
- [22] WEITZ J S, EKSIN C, PAARPORN K, et al. An oscillating tragedy of the commons in replicator dynamics with game-environment feedback [J]. *Proceedings of the National Academy of Sciences*, 2016, 113(47): E7518 – E7525.
- [23] ALBOSZTA J, MIE J. Stability of evolutionarily stable strategies in discrete replicator dynamics with time delay [J]. *Journal of Theoretical Biology*, 2004, 231(2): 175 – 179.
- [24] ARGASINSKI K. Dynamic multipopulation and density dependent evolutionary games related to replicator dynamics. a metasimplex concept [J]. *Mathematical Biosciences*, 2006, 202(1): 88 – 114.
- [25] GALLA T. Two-population replicator dynamics and number of Nash equilibria in matrix games [J]. *EPL (Europhysics Letters)*, 2007, 78(2): 20005.
- [26] AIMONE J A, HOUSER D. Beneficial betrayal aversion [J]. *Plos one*, 2011, 6(3): e17725.
- [27] TANIMOTO J, KISHIMOTO N. Network reciprocity created in prisoner's dilemma games by coupling two mechanisms [J]. *Physical Review E*, 2015, 91(4): 042106.
- [28] CHOI W, YOON S H, KIM Y. Percolation in spatial evolutionary prisoner's dilemma game on two-dimensional lattices [J]. *Physical Review E*, 2015, 92(5): 052140.
- [29] PACHECO J M, SANTOS F C, SOUZA M O, et al. Evolutionary dynamics of collective action in N-person stag hunt dilemmas [J]. *Proceedings of the Royal Society of London B: Biological Sciences*, 2009, 276(1655): 315 – 321.
- [30] STARNINI M, SÁNCHEZ A, PONCELA J, et al. Coordination and growth: the Stag Hunt game on evolutionary networks [J]. *Journal of Statistical Mechanics: Theory and Experiment*, 2011, 52(5): 581 – 600.
- [31] SZOLNOKI A, PERC M, SZABÓ G. Phase diagrams for three-strategy evolutionary prisoner's dilemma games on regular graphs [J]. *Physical Review E*, 2009, 80(5): 056104.

- [32] SZABÓ G. Branching annihilating random walk on random regular graphs [J]. *Physical Review E*, 2000, 62(5): 7474.
- [33] POZA D J, SANTOS J I, GALÁN J M, et al. Mesoscopic effects in an agent-based bargaining model in regular lattices [J]. *PLoS One*, 2011, 6(3): e17661.
- [34] SASIDEVAN V, DHAR D. Strategy switches and co-action equilibria in a minority game [J]. *Physica A: Statistical Mechanics and its Applications*, 2014, 402(2): 306 – 317.
- [35] BARBOSA V C, DONANGELO R, SOUZA S R. Early appraisal of the fixation probability in directed networks [J]. *Physical Review E*, 2010, 82(4): 046114.
- [36] LIU Bin, CHEN Laijun, WANG Yuchen, et al. Allocating reserve cost for hedging against wind generation uncertainty: a coalitional-game-theoretic approach [J]. *Control Theory & Applications*, 2016, 33(4): 437 – 445.
(刘斌, 陈来军, 汪雨辰, 等. 应对风电出力不确定性的备用成本分摊: 联盟博弈方法 [J]. 控制理论与应用, 2016, 33(4): 437 – 445.)
- [37] MAI Junhong. A dynamic game model for a class of venture investment behaviors [J]. *Control Theory & Applications*, 2016, 33(8): 1105 – 1110.
(麦均洪. 关于一类风险投资的动态博弈模型及分析 [J]. 控制理论与应用, 2016, 33(8): 1105 – 1110.)
- [38] HU Jing. Research on incentive policy of green building based on evolutionary game theory [J]. *Value Engineering*, 2017, 36(12): 37 – 40.
(胡静. 基于演化博弈理论绿色建筑激励政策研究 [J]. 价值工程, 2017, 36(12): 37 – 40.)
- [39] QI Kai, YANG Zhi, ZHANG Zimo, et al. Analysis on the evolutionary game of public opinion guidance mechanism of internet users under the participation of government [J]. *Information Science*, 2017, 35(3): 47 – 52.
(祁凯, 杨志, 张子墨, 等. 政府参与下网民舆论引导机制的演化博弈分析 [J]. 情报科学, 2017, 35(3): 47 – 52.)

作者简介:

张建磊 (1981–), 男, 博士, 研究方向为复杂网络、演化博弈理论、多智能体系统以及群体智能理论, E-mail: jianleizhang@nankai.edu.cn;

李智琦 (1995–), 男, 硕士, 研究方向为复杂网络与演化博弈理论, E-mail: 1137692135@qq.com;

曹明 (1977–), 男, 教授, 研究方向为系统与控制理论、复杂网络、演化博弈理论、多智能体系统以及群体智能理论, E-mail: m.cao@rug.nl.