Barry Cooper, Andrew E. M. Lewis-Pye, Angsheng Li, Yicheng Pan and Xi Yong

Establishing social cooperation: the role of hubs and community structure

Article (Accepted version) (Refereed)

Original citation:
Cooper, Barry and Lewis-Pye, Andrew and Li, Angsheng and Pan, Yicheng and Yong, Xi (2017) *Establishing social cooperation: the role of hubs and community structure.* Network Science. ISSN 2050-1242

Establishing social cooperation: the role of hubs and community structure

Barry Cooper[1], Andrew E. M. Lewis-Pye[2], Angsheng Li[3],
Yicheng Pan[3] and Xi Yong[3,4],
[1] Department of Mathematics, University of Leeds,
Leeds, UK, LS29JT
[2] Department of Mathematics,
London School of Economics and Political Science,
London, UK, WC2A 2AE
[3]State Key Laboratory of Computer Science, Institute of Software,
Chinese Academy of Sciences, Beijing, P. R. China
[4]University of the Chinese Academy of Sciences, Beijing, P. R. China[*]

November 7, 2017

Abstract

Prisoner's Dilemma games have become a well-established paradigm for studying the mechanisms by which cooperative behaviour may evolve in societies consisting of selfish individuals. Recent research has focussed on the effect of spatial and connectivity structure in promoting the emergence of cooperation in scenarios where individuals play games with their neighbors, using simple 'memoryless' rules to decide their choice of strategy in repeated games. While heterogeneity and structural features such as clustering have been seen to lead to reasonable levels of cooperation in very restricted settings, no conditions on network structure have been established which robustly ensure the emergence of cooperation in a manner which is not overly sensitive to parameters such as network size, average degree, or the initial proportion of cooperating individuals. Here we consider a natural random network model, with parameters which allow us to vary the level of 'community' structure in the network, as well as the number of high degree hub nodes. We investigate the effect of varying these structural features and show that, for appropriate choices of these parameters, cooperative behaviour does now emerge in a truly robust fashion and to a previously unprecedented degree. The implication is that cooperation (as modelled here by Prisoner's Dilemma games) can become the social norm in societal structures divided into smaller communities, and in which hub nodes provide the majority of inter-community connections.

[*]Authors are listed alphabetically. Lewis-Pye was supported by a Royal Society University Research Fellowship.

Significance Statement. Using Prisoner's Dilemma games to provide a mathematical framework for the study of cooperative behaviour, we investigate how emerging levels of cooperation are effected by community structure. We establish that unprecedented levels of cooperation can be seen to reliably emerge, if the appropriate balance is struck between a) the tendency to divide a society of individuals into smaller community structures, and b) the existence of individuals with large numbers of connections between, as well as within, these communities.

Introduction

As far back as Darwin [1] it has been a major scientific task to understand how cooperative behaviour can evolve in societies where individuals act in their own self-interest. Clearly, genetic mechanisms leading to such cooperative behaviour must play a key role in any analysis, and have accordingly been the subject of intense research over many years [2, 3, 4]. An understanding of the game-theoretic considerations involved is also crucial to the development of any satisfactory theory, however [5, 6]. The game theoretic analysis both informs the biological debate, and also provides explanatory power in contexts where the time scales of evolution mean that genetic considerations are unlikely to be of relevance. This game theoretic analysis can be dated back at least as far as the 1970s, and a well-known series of experiments conducted by Robert Axelrod [7, 8], who conducted tournaments in which pairs of computer programs were pitted against each other in Iterated Prisoner's Dilemma games. As has been well reported in the popular media, the highest average score was, perhaps surprisingly, achieved by simple "Tit-for-Tat" strategies, which cooperate on the first round and thereafter repeat their opponent's previous move. A key feature of these tournaments was that players were allowed to take into account the entire history of previous interactions – high levels of cooperation cannot be maintained in any "well-mixed" society (with interactions between all individuals) without such memory allowances [6]. Since those early experiments, Prisoner's Dilemma (PD) games and variants such as the Snowdrift game, have been widely used to provide a mathematical framework for the game-theoretic analysis of emergent cooperative behaviour [9].

An important development in the theory came in 1992, with the observation by Nowak and May [10] that spatial structure (or more generally connectivity structure) can significantly impact the manner in which cooperative behaviour evolves. They considered individuals arranged in a two dimensional grid, following simple "memoryless" strategies of cooperation or defection in PD games (each individual using the same strategy at each stage in all games with neighbors), the update rule at each stage being that an individual follow the highest scoring strategy amongst their neighbours in the previous round. In this setting cooperative behaviour can be maintained indefinitely (as opposed to the doomed fate of such simple cooperative strategies in well-mixed populations). This move to consider connectivity structure was to prove key to much future research [11, 12, 13, 14, 15, 16, 17]. While very significant, the observed levels of cooperation in these experiments were not robust, however, in the sense that they relied heavily on the deterministic nature of the evolutionary process described. If the update process is altered to incorporate random elements in the decision making process, as is perhaps more realistic [18, 19], then the previously observed cooperative behaviour is no longer maintained. In this setting, it then becomes significant that there is now a substantial body of evidence establishing that social, biological and technological networks tend not to be uniform, and in fact have high levels of degree heterogeneity [20, 21, 22]. Santos and Pacheco [23, 24] showed that much higher levels of cooperation can emerge in networks with such heterogeneities, in a manner which is not sensitive to stochastic elements in the dynamic process. Networks constructed according to the preferential attachment (PA) model of Barabási and Albert, in particular, were shown to support the emergence of high degrees of cooperation around the largest hub. These observations have been verified and extensively developed for other models [15, 12, 16]. The levels of cooperation observed for these PA networks may still be seen as unsatisfactory, however, for two principal reasons. Firstly, while the levels of cooperation observed compare very favourably with those for homogeneous networks, they still do not seem sufficient to explain the high levels of such behaviour observed achieved in real world populations. Secondly (as will be expanded on later) the cooperative behaviour evolving on these networks is not at all robust to changes in network parameters such as average degree. The positive results seen in previous studies concern networks with low average degree of around 4, and cease to hold if the average degree is 8 or higher.[1]

In this paper we consider a natural random network model, the *Community-Hub* model, with parameters which allow us to vary the level of community structure in the network, as well as the number of high degree hub nodes. We shall establish that, for appropriate choices of these parameters, cooperative behaviour emerges in a truly robust fashion and to a previously unprecedented degree. Following common practice, we consider a weak version of the prisoner's dilemma game, introduced by Nowak and May [10]. Each game is specified by the *temptation* payoff b, which is normally taken to be in the real interval $[1, 2]$. Each player either plays the strategy C (cooperate) or D (defect). For mutual cooperation, both players receive the payoff 1. Mutual defection sees both players receive 0. If one cooperates and the other defects, then the cooperator gains payoff 0, while the defector gains the temptation payoff b. The motivation for considering games with standardised payoffs of this form is to reduce the size of the parameter space, and similar results are obtained if the 'sucker's payoff' given to a cooperator when the other player defects is taken to be sufficiently small and negative.

The dynamics we consider are also standard. The evolutionary process unfolds in discrete stages. Initially each individual (node) in the network chooses the strategy C with probability ζ and otherwise chooses the strategy D (choices for distinct nodes being independent). Following common practice, we shall normally take $\zeta = 0.5$, but since we are interested in establishing contexts in which cooperation emerges in a truly robust manner, we shall also sometimes consider lower values of ζ. At each stage s, each node then plays one PD game for each of its edges, the opponent in each case being the node sharing that common edge. Each node thus plays the same strategy in all of its games at a single stage, and may play multiple games with one node if there are multiple edges between them. If a cooperating node has k-many edges connected to cooperators, it will therefore receive a total payoff of k, while a defector in the same situation will receive kb. The update process, in which strategies are chosen for the next stage, then proceeds as follows at stage s. Every node u chooses one of its edges uniformly at random, and then compares its own score S_u with the score S_v of the neighbor v sharing that edge. With probability

$$\frac{1}{1 + \exp(-(S_v - S_u)/T)}$$

u will then change its strategy for stage $s + 1$ to that played by v at stage s (u's strategy

[1] It should be noted that while this holds for static networks, studies have shown cooperation being maintained in dynamic networks with high average degree, in the case that individuals are able to rewire their connections with undesirable neighbors (see, for example, [25]).

remaining unchanged otherwise). Here T is the 'temperature', taken to be 0.04 (our results being quite robust to changes in T).

Cooperation on PA networks

Before defining the Community-Hub model, we briefly consider the emergence of cooperation on PA networks, originally introduced by Barabási and Albert [20], since these are the networks for which the highest levels of cooperation have been observed in the literature. A PA network of size N, with connectivity parameter d, is constructed in stages as follows. At stage 1 we add two nodes with d edges between them. At stage s with $1 < s < N$, one further node u is added. We sample d many nodes from those added at previous stages (sampling with replacement), the probability any given node is chosen for each sample being proportion to its present degree. An edge is then added between u and each of the sampled nodes.

Figure 1 shows the outcome of simulations for the PA model. For these simulations the 'standard' input parameters are taken to be: network size 30000, $d = 4$, $b = 1.7$. Each plot then shows how resulting levels of cooperation are affected by varying one of these parameters, while other parameters take the standard values. While higher levels of cooperation may be obtained for smaller d ($d = 2$ say), or for smaller values of b, the general form of the results obtained are not overly sensitive to changes in the choice of 'standard' input parameters.

Two principal observations should be made. Firstly, it is striking that the results for each parameter set can invariably be partitioned quite cleanly into two sets, the first with levels of cooperation well above the mean, and the second with levels of cooperation well below this value. The maximum and minimum values (over 50 simulations) are reliable functions of the input parameters, varying much more smoothly than the mean values. The clear implication is that the mean value suffers from noise owing to the random chance for each simulation as to whether it belongs to the first or second partition, with high or low levels of cooperation accordingly. These features remain true when simulations are run over a much larger number of generations (of the order 10^5 or more), and when the values plotted are averaged over a larger number of stages. While these observations seem worthy of further investigation, for now the key point to be made is simply that (even when the mean levels of cooperation obtained are high), no parameter set gives high levels of cooperation in a truly reliable fashion.

The second key observation to be made is that emerging levels of cooperation show a great deal of sensitivity to the connectivity parameter d. The highest levels of cooperation are observed for $d = 2$ (the standard choice in the existing literature), and then fall rapidly as d is increased. PA networks must be of low average degree to support the emergence of even reasonably high levels of cooperation. To a lesser extent, increasing network size or decreasing the initial proportion of cooperators, also has a negative impact on cooperation levels.

The Community-Hub Model

As well as d, and network size N, the model is specified by three further parameters: an *affinity exponent* $a \in \mathbb{R}^{\geq 0}$, a *hub coefficient* $\sigma \in \mathbb{R}^{\geq 0}$, and a *cohesion coefficient* $\eta \in [0, 1]$. The network G is constructed in stages as follows:

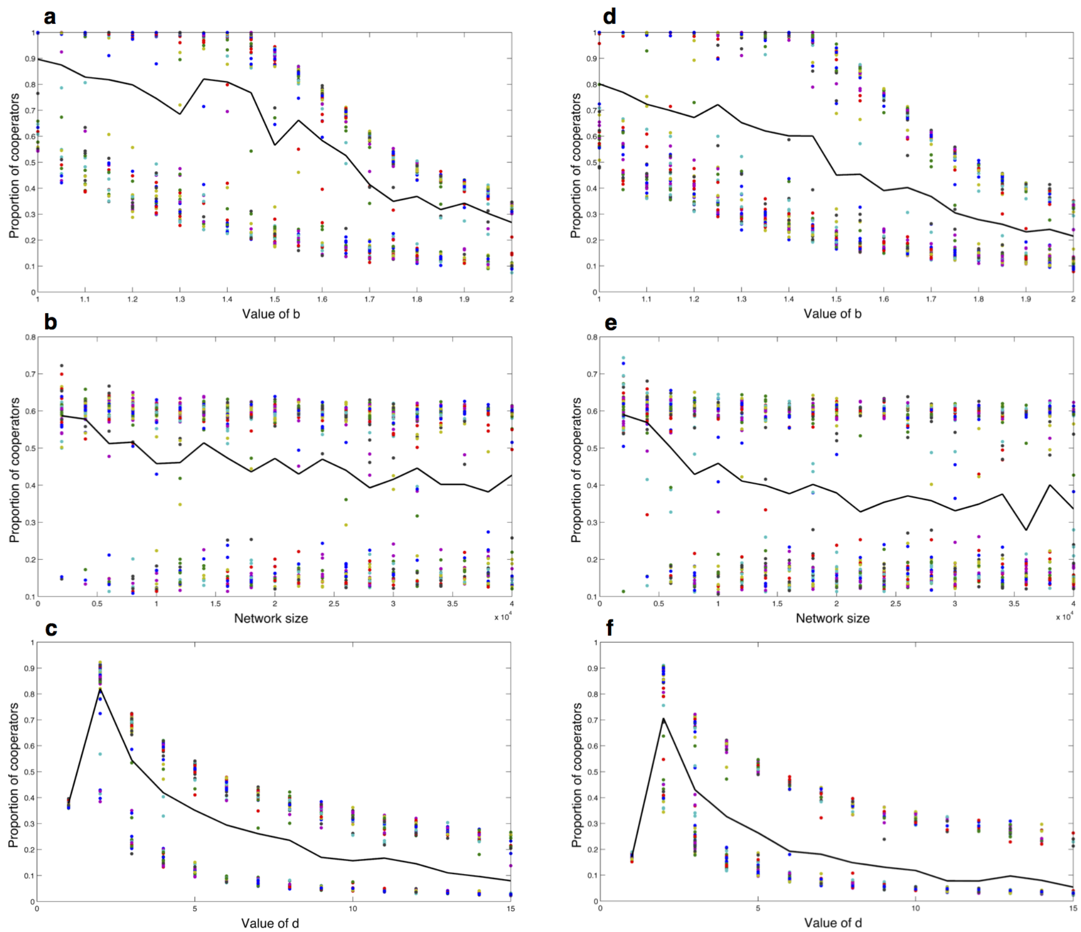

Figure 1: Cooperation on PA networks. For these simulations the 'standard' input parameters were: network size 30000, $d = 4$, $b = 1.7$. In plots (a),(b) and (c), $\zeta = 0.5$, while in plots (d),(e) and (f), $\zeta = 0.25$. In each plot one parameter is varied, while others take the standard values. 50 simulations were run for each parameter set. Each simulation was run for 5000 stages, and then the average proportion of cooperators over the last 1000 stages recorded. The outcome of each simulation is plotted as a point, while the line plots the average over all 50 simulations for each parameter set. Plots (a) and (d) show the proportion of cooperators for varying b, while (b) and (e) vary network size and (c) and (f) vary d.

Stage 0. Enumerate a single node into G, which is designated a *hub* node.
Stage s, $0 < s < N$. For e which is the base of the natural logarithm, let $p_s = \frac{1}{(\ln(s+e))^a}$. Carry out the following steps.

1. Enumerate a new node into G, u say.
2. With probability p_s, choose a new color for u. If a new color is chosen then:
   (a) Designate u a hub node.
   (b) Create $\lfloor \sigma \cdot d \rfloor$ many (undirected) edges (u, v_j), where each v_j is chosen uniformly at random amongst previous hub nodes.
3. Otherwise, if a new color was not chosen for u, then:
   (a) Choose an existing color for u uniformly at random.
   (b) Create d many edges (u, v_j). For each j, $1 \leq j \leq d$, v_j is chosen as follows. With probability η choose v_j from amongst the previous nodes with the same color as u according to the preferential attachment rule, i.e. each previous node of the same color is chosen with probability proportional to its degree. (At stage 1, this means that the first node enumerated into G must be chosen.) With probability $1 - \eta$ choose v_j from amongst *all* previous nodes according to the preferential attachment rule.

The model allows for a good deal of flexibility in the form of the resulting network. If $a = 0$, for example, then the resulting network will simply be an Erdös-Rényi random graph. For a given N and sufficiently large affinity exponent a, the resulting network will be a PA network, as considered previously. Nodes of the same color may be thought of as forming a community, while hub nodes may be thought of as playing a leadership role within communities. Varying the affinity exponent a, allows us to control the number of communities in the network (the choice of p_i is made so as to ensure a reasonable balance between the number and sizes of communities as a function of N). Varying the cohesion coefficient then allows us to control very directly the extent to which nodes of the same color do form close-knit communities, with most connections for non-hub nodes being to others within the same community. As well as increasing the number of connections between communities, the principal effect of increasing the hub coefficient is to magnify the role of the hub nodes by increasing their degree and thereby increasing the probability that other nodes in their community will follow their strategy when scores are compared. The model for the case that the cohesion coefficient $\eta = 1$, has previously been studied in the context of network security [27].

In the next section we shall examine the effect of these parameters on the levels of emerging cooperation. To give an immediate indication of the levels of cooperation that can be achieved, however, Figure 2 compares levels of cooperation achieved by Community-Hub (CH) networks and PA networks, for the same input parameters as in Figure 1. For these simulations we set $a = 1.2$, $\sigma = 2$ and $\eta = 1$. Again 50 simulations are run for each parameter set, and the outcomes of individual simulations for the CH networks are plotted as points. The solid black line shows mean values for the Community-Hub model (the proportion of cooperators normally being very close to 1). The reliability of the resulting levels of cooperation, is such that the outcomes of individual simulations are generally indistinguishable from

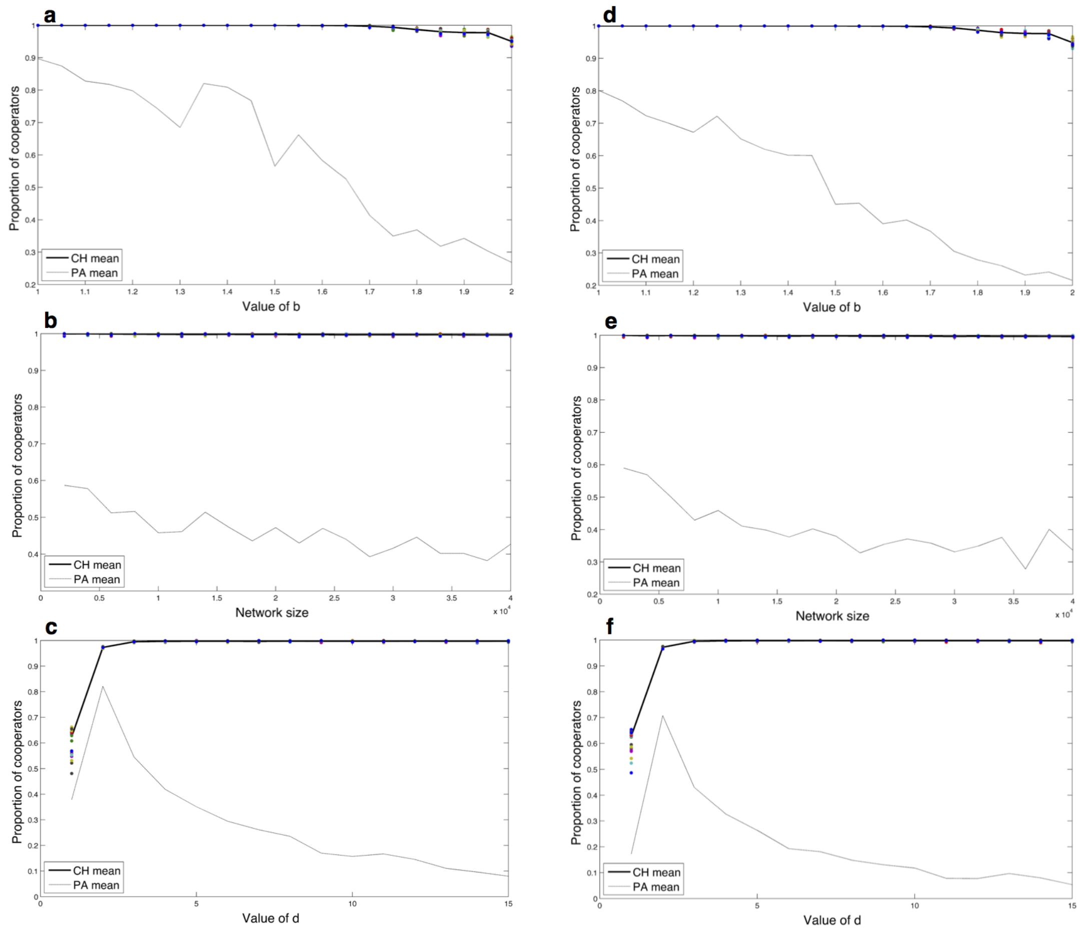

Figure 2: Comparing cooperation levels for PA and Community-Hub networks. For each plot the input parameters are as in Figure 1. For the Community-Hub model, we set $\sigma = 2$, $a = 1.2$ and $\eta = 1$. The outcome of each simulation for the Community-Hub model is plotted as a point. The solid black line shows mean values for the Community-Hub model (normally very close to 1). For comparison, the grey line shows mean values for the PA model, as plotted also in Figure 1.

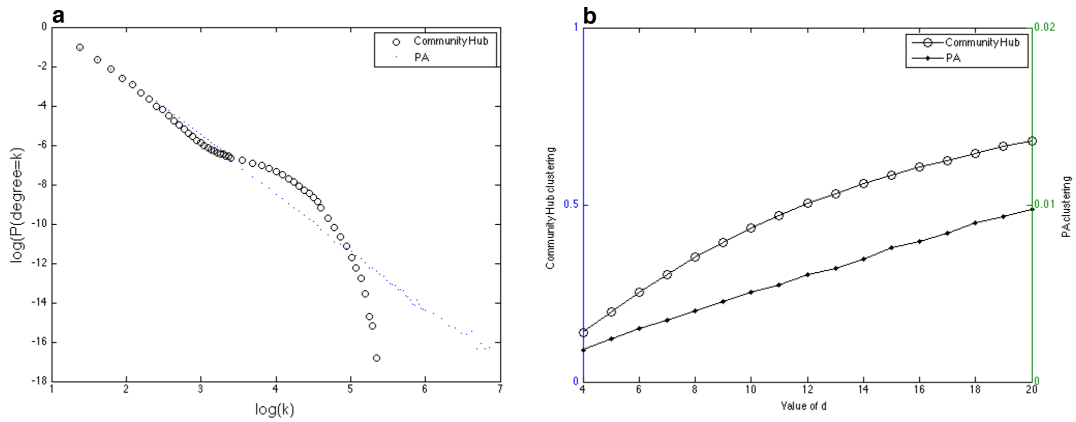

Figure 3: Graph (a) gives a log/log plot of the proportion of nodes with each degree, for the Community-Hub and PA networks and for the following 'standard' input parameters: network size 30000, $d = 4$, $a = 1.2$, $\sigma = 2$ and $\eta = 1$. Graph (b) shows how the global clustering coefficient varies with d, when other inputs still take the same standard values. The left y-axis gives the coefficient for the Community-Hub networks, while the right y-axis gives the value of the coefficient for PA networks.

the mean values. For comparison, the grey line shows mean values for the PA model, as also plotted in Figure 1.

For later reference we also describe some properties of the CH networks in terms of standard network metrics, and give comparisons to the corresponding values for PA networks. The mean degree for the CH network is easily calculated to be given by the expression $2(d + (\sigma - 1)d\alpha)$, where α is the proportion of the nodes which are designated hub nodes (actually this expression is only completely precise in the limit of network size, since the node added at stage 0 is a hub and initially has degree 0, but will be very accurate for networks of at least 100 nodes). For the 'standard' input parameters considered in Figure 2, α is $\approxeq 0.07$, meaning that the average degree for each CH network will be just slightly larger than for the corresponding PA network with the same value of d. Figure 3(a) then compares the degree frequencies for CH and PA networks, again for the 'standard' input parameters considered in Figure 2, and averaged over 500 simulations. It is notable that the CH network model *does not* have the scale-free property satisfied by PA networks. Finally, Figure 3(b) then compares the global clustering coefficients for CH and PA networks – for a definition of the global clustering coefficient see [26]. As might be expected, the CH networks have global clustering coefficients of the order 100 times those for the corresponding PA networks.

Results and Analysis

Figure 4 examines the effect of the cohesion coefficient η on resulting levels of cooperation. In these simulations the 'standard' input parameters were $N = 10000$, $d = 4$, $a = 1$, $b = 1.7$, $\sigma = 2$, $\zeta = 0.5$. Although specific values may change, the form of the results obtained is robust to changes in the choice of standard parameters. Each simulation was run for 5000 stages, and then the average proportion of cooperators over the last 1000 stages recorded. It should be noted that this generation number of 5000 is shorter than the time allowed for cooperation to emerge in much of the literature, where tests are often run over

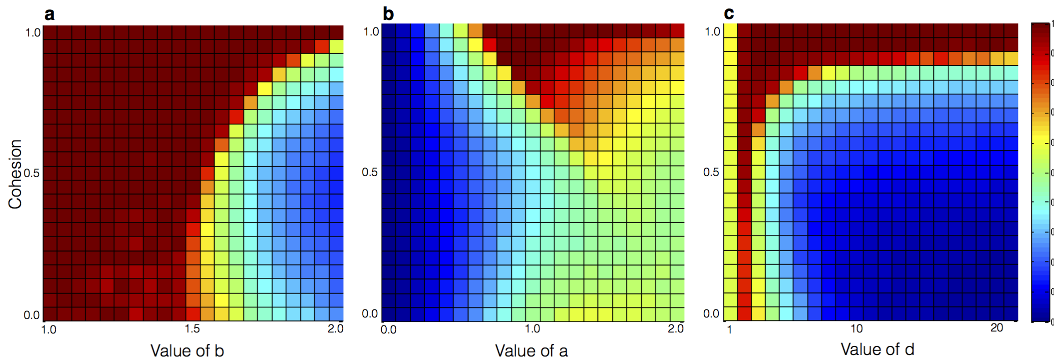

Figure 4: The effect of cohesion. In all plots the cohesion parameter is varied from 0 to 1 in increments of 0.05. In plot (a), b is varied from 1 to 2 in increments of 0.05. In (b), the affinity exponent a is varied from 0 to 2 in increments of 0.1. In (c), d is varied from 1 to 21. The color scale shows resulting cooperation levels.

10000 generations or more. While longer test lengths do not significantly effect results for the PA model, we allow a smaller number of generations, since we are interested in establishing network structures which rapidly result in high levels of cooperation. For those parameter sets resulting in high levels of cooperation, in fact a much smaller number of generations normally suffices to give the cooperation levels described here. For each parameter set 50 simulations were run, and then the level of cooperation plotted is the mean over these 50 simulations. Each plot shows the effect of varying two parameters, while other parameters take the standard values. The principal observation to be made is that, while increasing η in the interval $[0, 0.5]$ leads to decreased cooperation levels (or perhaps no change in the case that cooperation levels are close to 1), the reverse is true in the interval $[0.5, 1]$, and it is here that the highest levels of cooperation result. As a precursor to our subsequent discussion concerning the effect of changes in the affinity exponent a, one may also note that, especially for $\eta \in [0.5, 1]$, the highest levels of cooperation result for a close to 1. For η close to 1, cooperation levels are quite robust to changes in a in the interval $[1, 2]$, but not to placing a at values significantly below 1. Perhaps the most significant observation concerns the interplay between η and d. For $d = 2$ (which is, again, the case analysed in much of the literature), high cooperation levels are observed for all values of η, although with the highest levels of cooperation still occurring when η is close to 1. For larger values of d, however, high levels of cooperation *only* emerge when the cohesion coefficient η is close to 1. Since we are principally interested in establishing those network parameters which ensure high levels of cooperation, from this point on we shall therefore focus on the case $\eta = 1$.

Figure 5 examines the effect of changes in the affinity exponent a, for the same set of standard input parameters as Figure 4, but with η now fixed at 1. Once again, one sees cooperation emerging most robustly for values of a close to 1. For values of $b \in [1.2, 2]$ there appears to be a close to linear relationship between b and the minimum value of a required to give maximum cooperation levels. For all $d \geq 3$ an affinity exponent of at least 0.8 (and ≤ 1.4) suffices to achieve maximum cooperation levels. It is striking that for a around 1, there is very little dependency on d, so long as $d \geq 3$. The relationship between a and σ, however, is more subtle. For the values of a close to 1 that we are principally interested in, increasing σ from 1 to 1.75 will cause an increase in cooperation levels. For large σ, however, higher values of a are required to give maximum cooperation.

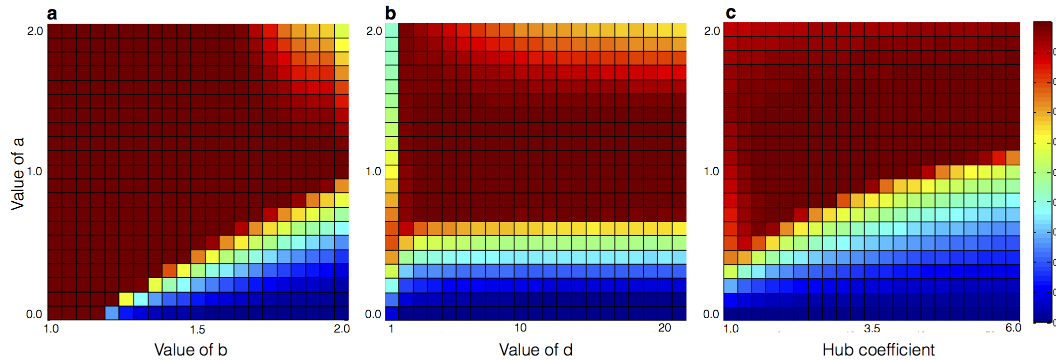

Figure 5: Varying affinity. In all plots the affinity parameter is varied from 0 to 2 in increments of 0.1. In plot (a), b is varied from 1 to 2 in increments of 0.05. In (b), d is varied from 1 to 21. In (c), σ is varied from 1 to 6 in increments of 0.25.

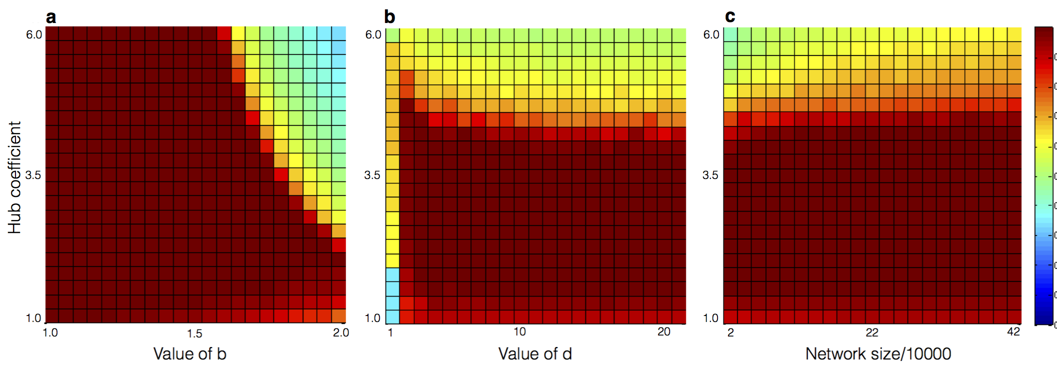

Figure 6: Varying σ. In all plots σ is varied from 1 to 6 in increments of 0.25. In plot (a), b is varied from 1 to 2 in increments of 0.05. In (b), d is varied from 1 to 21. In (c), network size is varied from 2×10^3 to 42×10^3 in increments of 2×10^3.

Figure 6 then shows the effect of varying σ. Once again, the standard set of input parameters is the same as for Figures 4 and 5, and while specific values may change, the general form of the results obtained is robust to changes in these standard parameters. We see that for values of b approaching 2, the highest cooperation levels are obtained with $\sigma = 2$. For $d \geq 3$, the values of σ required to ensure high levels of cooperation show little dependence on d, and high cooperation will result so long as σ is at least 1.75, and at most 4. Similarly, network size has little impact on the values of σ giving maximum cooperation.

In looking to explain the high levels of cooperation observed in these simulations, it is natural to begin thinking in terms of standard network metrics. In Figure 3, we saw that CH networks do not have the scale-free property satisfied by PA networks. Since it is normally this scale-free property, and the corresponding heterogeneity in terms of degree distribution, which is used to explain high levels of cooperation for PA networks, it initially seems unlikely that the degree distribution alone can be used to provide a satisfactory explanation of the observed results. While CH networks were observed to have much higher global clustering coefficients than their PA counterparts in Figure 3(b), it also seems unlikely that the global clustering coefficient alone can be used to provide a satisfactory account, for at least two reasons: (a) in PA networks an increase in d causes an increase in the global cluster-

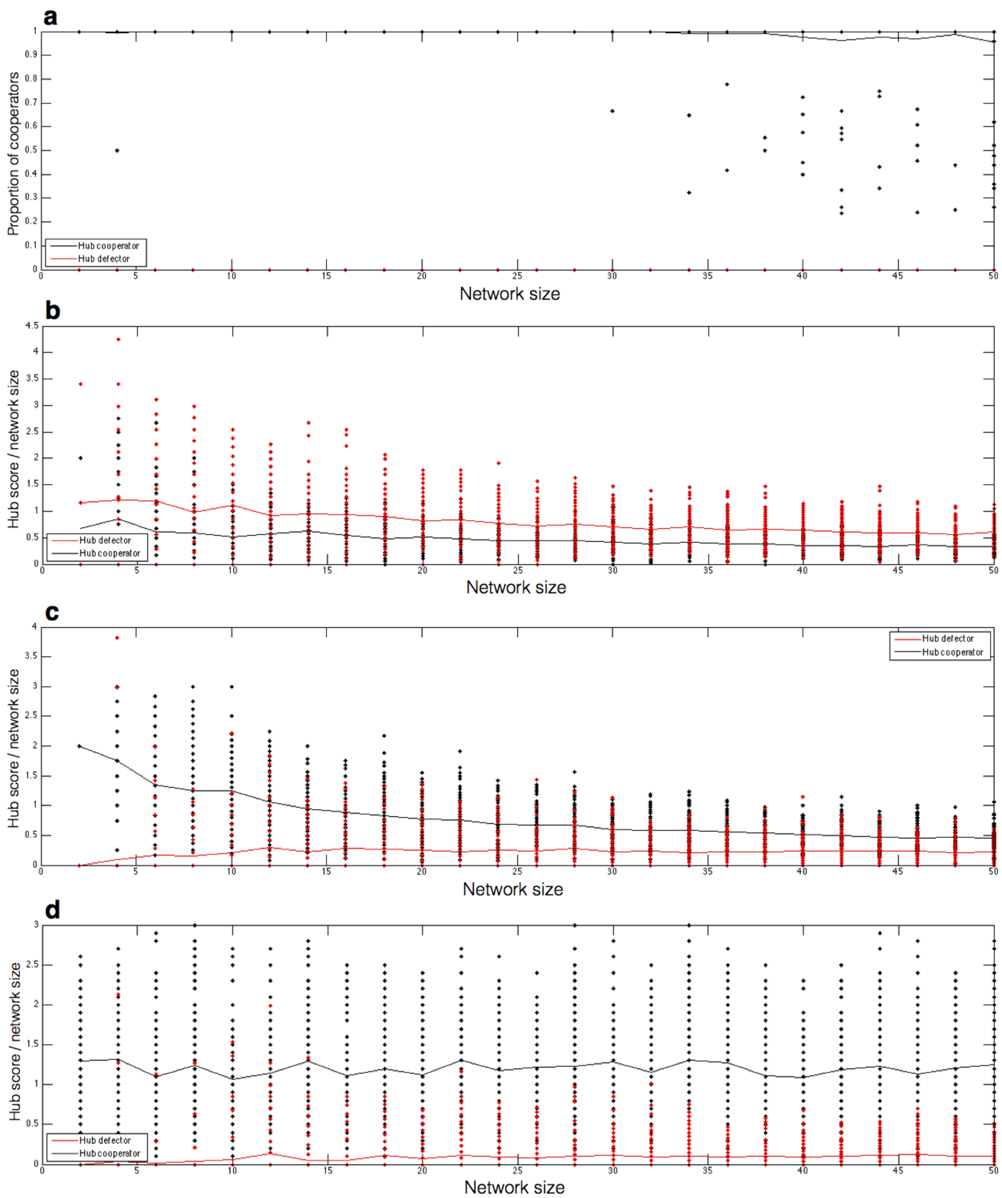

Figure 7: The effect of the hub. These plots show the outcomes of *rigid hub* simulations for PA networks. The standard input parameters were $d = 4$, $b = 1.7$, $\zeta = 0.3$. For each parameter set (and for each case that the hub is a cooperator or defector) 30 simulations were run. The outcomes of individual simulations are plotted as points, while the mean values are plotted as lines (cooperator hub values in black, defector hub values in red). Plot (a) shows the proportion of cooperators after 2000 stages (these proportions for defector hubs always being 0). Plots (b), (c) and (d), show the score of the hub node (normalised according to network size) at stages 0, 1 and 2 respectively.

ing coefficient but also tends to significantly *decrease* observed levels of cooperation, and (b) other networks with similar or higher clustering coefficients have been studied (see for example [12]) without the same high levels of cooperation being observed. The high levels of cooperation achieved by CH networks can be explained, however, modulo an interesting phenomenon that may be observed for PA networks. Consider the case that $\eta = 1$, $a = 1$ and $\sigma = 2$. Then the CH network may be understood as collection of communities which are smaller PA networks, connections between communities being via the hub nodes which will have significantly larger degree than other nodes in their community. For network sizes of the order 5×10^4, say, there is a negligible probability that more than 1% of nodes will belong to communities of size 50 or more. In order to understand the evolving process within each of these smaller PA communities *when the strategy of the hub node is fixed*, we consider a modified version of the dynamics, referred to as the 'rigid hub' model. In this model we consider a PA network, in which only the first node is designated a hub. For the purpose of preferential attachment selection, the degree of the hub node is increased by $2d$ (or more generally $\lfloor \sigma \cdot d \rfloor$) so as to accurately reflect the corresponding scenario in the CH model, i.e. $2d$ is added to the degree of the hub node when calculating sampling probabilities during edge formation. The strategy of the hub node is now fixed throughout all stages. In order to reflect the increased score achieved by hub nodes in the CH model via connections to other hubs, the dynamics are also altered such that any node comparing their score with the hub will copy the hub's strategy at the next stage. Figure 7 shows the outcomes of simulations for this model. Let us concentrate initially on plots (b), (c) and (d), which show scores for the hub node at stages 0 (prior to the point at which any strategy changes have been made), 1 and 2. At stage 0, defecting hub nodes will of course achieve higher average scores. By stage 1, however, cooperating hub nodes will have significantly higher scores on average, with this effect being strongly exaggerated by stage 2 and at subsequent stages. The implication for the CH model is clear. While cooperator hubs which compare their score with defecting hubs of similar degree at the first stage are likely to change strategy, many cooperator hubs (such as those who compare scores with other cooperators) will not change their strategy at the first stage. By the next stage cooperating hubs will, on average, be those with the highest scores. Hubs which cooperate at this stage are now liable to continue cooperating, because comparisons with defector hub nodes can be expected not to cause a strategy change, while the same is true when comparisons are made with nodes within the same community, due to the significantly higher degree of hub nodes. Defector hubs, on the other hand, are likely to change their strategy when comparing score with a cooperating hub, at any stage after the first and with increasing probability as the stages progress. Plot (a) in Figure 7 then shows the effect of this process on resulting levels of cooperation in the long term. For defecting hubs the level of cooperation resulting in the rigid hub model was 0 for all network sizes tested. For cooperating hubs, however, 100% cooperation is achieved in the majority of cases.

In the CH model a mechanism similar to the "older-get-richer" of the Barabási-Albert preferential attachment model operates on the community size: communities associated with a color attributed earlier on will comprise a larger number of nodes. The community size distribution is itself therefore heterogeneous, with older communities having more nodes. One might ask how this additional heterogeneity source influences the observed values of cooperation. Aguing in terms of the rigid hub model, as above, the conclusion would seem to be that this heterogeneity in terms of community size *is not* important for the evolution of high degrees of cooperation. Figure 7(a) indicates cooperation emerging less reliably for hub cooperators when community sizes grow too large. One might therefore expect levels of cooperation to be further improved if the model were to be adjusted so as to cap community sizes at appropriate levels and, in fact, this is precisely what one observes in simulations. For the same standard input parameters as in Figure 2(a) and for $b = 2.0$, the standard CH model gives a mean proportion of cooperators of around 0.94 (with this proportion being given, as previously, by averaging over the last 1000 stages of a 5000 stage process). Adjusting the model so that community sizes are capped at 30, for example, gives an increase in this proportion to above 0.96.

Discussion

We have investigated the effect of community structure and hub nodes on levels of emerging cooperation in PD games. While a purely random network structure is not conducive to cooperation, a 'rich get richer' social norm as exemplified by PA networks, also fails to produce cooperation in a robust fashion. High levels of cooperation will evolve robustly, however, if the appropriate balance is achieved between the tendency to divide a society of individuals into smaller community structures on the one hand, and the existence of individuals with large numbers of connections between, as well as within, these communities on the other. With a dynamics (as here and most of the literature) in which scores are not normalised by degree, hub nodes are able to adopt a leadership role within communities, since their increased degree means that their strategies are likely to be copied by others. Cooperating hub nodes will then rapidly see significantly higher utilities than their defecting equivalents, and will lead to the formation of communities with high levels of cooperation. Ultimately, their increased levels of utility mean that cooperating hubs, and the communities with high levels of cooperation that they result in, become the social norm. In a context where utilities are normalised according to degree, it may be necessarily to directly increase the influence of hub nodes and others (by implementing mechanisms increasing the chance that non-hub nodes will follow the strategy of hub nodes when scores are compared, for example) in order to maintain cooperation levels. It is an interesting direction for future research to understand such *mechanisms of influence*, and the role that they may play in the emergence of cooperation.

References

[1] Darwin, C. On the origin of species by means of natural selection, *London: John Murray*, 1859.

[2] Bell, G. Selection: the mechanism of evolution. *Oxford: Oxford University Press*, 2008.

[3] Dawkins, R. The Selfish Gene (30th Anniversary ed.). United Kingdom: *Oxford University Press*, 2006.

[4] Smith, J. M. Group Selection and Kin Selection. *Nature*, 201 (4924), 1145-1147, 1964.

[5] Smith, J. M. Evolution and the Theory of Games, *Cambridge University Press*, 1982.

[6] Pennisi, E. How did cooperative behaviour evolve. *Science*, **309**, 93-93, 2005.

[7] Axelrod, R. Effective choice in the prisoner's dilemma. *J. Confl. Resol.* 24, 3-25, 1980.

[8] Axelrod, R. More effective choice in the prisoner's dilemma. *J. Confl. Resol.* 24, 379-403, 1980.

[9] Szabó, G. & Fáth, G. Evolutionary games on graphs *Physics Reports*, **446**, 97-216, 2007.

[10] Nowak, M. A. & May, R. M. Evolutionary games and spatial chaos, *Nature*, **359**, 6398, 826-829, 1992.

[11] Lieberman, E., Hauert, C. & Nowak, M. A. Evolutionary dynamics on graphs, *Nature*, **433**, 312-316, 2005.

[12] Vukov, J., Szabó, G. & Szolnoki, A. Evolutionary prisoner's dilemma game on Newman-Watts networks, *Phys. Rev. E*, **77**, 2008.

[13] Szabó, G. & Tőke, C. Evolutionary prisoner's dilemma game on a square lattice, *Phys. Rev. E*, **58**, 69, 1998.

[14] Gómez-Gardenes, J., Campillo, M., Floría, L. M. & Moreno, Y. Dynamical organization of cooperation in complex topologies. *Physical Review Letters*, **98**, 2007.

[15] Santos, F.C., Pacheco, J.M. & Lenaerts, T. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proceedings of the National Academy of Sciences*, **103** (9): 3490-3494, 2006.

[16] Roca, C. P., Cuesta, J. A. & Sánchez, A. Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics, *Phys. of Life Rev.*, **6**, 4, 208-249, 2009.

[17] Pacheco, J. M. & Santos, F. C. Network dependence of the dilemmas of cooperation. *AIP Conf.* , **776**, 90-100, 2005.

[18] Huberman, B & Glance, N. Evolutionary games and computer simulations, *Proc. Nat. Acad. Sci.* 90, 7716-7718, 1993.

[19] Traulsen, A., Semmann, D., Sommerfeld, R.D., Krambeck, H-J. & Milinski, M. Human strategy updating in evolutionary games, *Proc. Nat. Acad. Sci.*, 107, 7, 2962-2966, 2010.

[20] Barabási, A. L. & Albert, R. Emergence of scaling in random networks. *Science*, **286**(5439):509–512, 1999.

[21] Barabási, A. L. Scale-free networks: a decade and beyond. *Science*, **325**, 5939: 412-413, 2009.

[22] Watts, D. J. & Strogatz, S. H. Collective dynamics of small world networks, *Nature*, **393**, 440-442, 1998.

[23] Santos, F. C. & Pacheco, J. M. Scale-free networks provide a unifying framework for the emergence of cooperation, *Phys. Rev. Lett.*, **95**, 9, 98-104, 2005.

[24] Santos, F. C. & Pacheco, J. M. A new route to the evolution of cooperation, *European Society for Evolutionary Biology*, **19**, 726-733, 2006.

[25] Santos, F. C., Pacheco, J. M., & Lenaerts, T. Cooperation prevails when individuals adjust their social ties. *PLoS Computational Biology*, 2(10), 2006.

[26] Newman, M. Networks: an introduction. Oxford University Press, Inc. New York, NY, USA, 2010.

[27] Li, A. & Pan, Y. A theory of network security: Principles of natural selection and combinatorics. *Internet Mathematics*, 2015.