SCIENTIFIC REPORTS

OPEN

Identification of influential invaders in evolutionary populations

Guoli Yang[1,2], Tina P. Benko[3], Matteo Cavaliere[4], Jincai Huang[1] & Matjaž Perc [3,5,6]

The identification of the most influential nodes has been a vibrant subject of research across the whole of network science. Here we map this problem to structured evolutionary populations, where strategies and the interaction network are both subject to change over time based on social inheritance. We study cooperative communities, which cheaters can invade because they avoid the cost of contributions that are associated with cooperation. The question that we seek to answer is at which nodes cheaters invade most successfully. We propose the weighted degree decomposition to identify and rank the most influential invaders. More specifically, we distinguish two kinds of ranking based on the weighted degree decomposition. We show that a ranking strategy based on negative-weighted degree allows to successfully identify the most influential invaders in the case of weak selection, while a ranking strategy based on positive-weighted degree performs better when the selection is strong. Our research thus reveals how to identify the most influential invaders based on statistical measures in dynamically evolving cooperative communities.

Group influence usually will make individuals adopt the behaviors held by their neighbours, leading to the propagation of states throughout the network. Models for the dynamics on networks have been studied in a large number of domains, including the propagation of information, idea, innovation, product, meme, failure, rumor through various networks[1–6]. It is interesting to see that even a small fraction of initial spreaders can trigger a large cascade[7,8], which can be extended and used in the control of outbreak of epidemics[9], the conduction of advertisements for products[10], and the protection of important assets in power grids or cyberspace[11]. In order to identify the influential targets, as known as influence maximization problem (IPM), two types of methods are investigated widely: one is based on operation research, and the other is based on topological centrality.

From the perspective of operation research, people aims to find a set of initial spreaders which can achieve the most prevalent propagation collectively. Kempe *et al.*[12] proposed the first provable approximation guarantees for efficient algorithms and adopted a natural greedy strategy based on sub-modular function[13] to obtain nearly 63% of the optimal solution for independent cascade model[14] and linear threshold model[15]. After that, extensive researches[16–20] have addressed the influence maximization problem in terms of scalability and uncertainty.

From the perspective of topological centrality, Lü *et al.* reviewed the progresses in the identification of vital nodes and described the state of the art in a systematic review[21]. It is known that the influence of a node is highly dependent on the network structure and its location in the network, but how to calculate the real centrality of a node is nontrivial. Despite there are more than 30 methods to characterize the *importance* of a node, there is not an ubiquitous one which can be applied to different kinds of dynamical systems. Generally speaking, four classes of ranking measurements are employed, as mentioned in[21], namely ranking based on local information[22–24], ranking based on global information[25,26], ranking based on iterative refinement[27,28], and ranking based on node removal[29,30].

Notably, the dynamics of various influence spreading models lie on the stochastic triggering of interactions from one state to another one between connecting individuals, including independent cascade model, linear threshold model, epidemic model (such as SI, SIS, SIR), and voter models[31,32]. However, when the dynamics of network is coupled to the evolution of the population (under the framework of evolutionary game theory, EGT), things dramatically change as many results from optimization theory and centrality measures cannot be used to

¹Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha, 410073, China. ²Unit 66136, Beijing, 100042, China. ³Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, SI-2000, Maribor, Slovenia. ⁴School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Manchester, UK. ⁵CAMTP – Center for Applied Mathematics and Theoretical Physics, University of Maribor, Mladinska 3, SI-2000, Maribor, Slovenia. ⁶Complexity Science Hub Vienna, Josefstädterstraße 39, A-1080, Vienna, Austria. Correspondence and requests for materials should be addressed to M.P. (email: matjaz.perc@gmail.com)

successfully identify the influential invaders, that reproduce and disappear under natural selection. Evolutionary game theory[33–36] studies a mathematical framework combining game theory[37] with evolutionary population characterized by the Darwinian process[38]. The competing strategies are formulated by the principles of game theory, where the fitness of an individual is dependent on the neighbourhood configuration and the game settings. In the framework of Prisoner's Dilemma on networks, the strategy of defection is optimal for individuals, but the strategy of cooperation is the best option for the whole population. A large number of studies[39–64] have been carried out to explore the relationship between population composition and network structure - these works have shown that the promotion of cooperation is rooted in the formation of highly connected cooperative communities (for reviews see[65–68]).

Few cheaters added into a cooperative community (just like the initial spreaders in propagation networks) will avoid the cost of contributing to the community and can occasionally spread leading to the complete collapse of cooperation and the fragmentation of the community[69]. How to identify the most influential spreaders has been widely investigated in static networks, but the problem remains open for dynamical networks, where the topology of the networks change because of node replications and/or removal[70–72]. In general, the interplay of network and evolutionary dynamics makes challenging the identification of the most influential spreaders in evolving communities. This paper, for the first time to our knowledge, explores the scenarios of the cooperation/cheating conflict in dynamical networks[69] and proposes an original methodology for the identification of the most influential cheating mutants, i.e., those that can lead to the collapse of cooperation. In the dynamical network model considered in this paper, an initial small fraction of nodes are selected to hold the strategy of defection and compete with the rest of cooperators. As in the case of ranking strategies, we propose and evaluate a weighted degree decomposition (WDD) method to identify the influential nodes which are those associated to the largest possibilities of invasion. The results obtained illustrate that a ranking strategy based on negative-weighted degree can successfully identify the most influential invaders in the case of weak selection, while a ranking strategy based on positive-weighted degree performs better when selection is strong.

Results

Model descriptions. We consider a dynamical network with a fixed number of nodes but with a non-fixed number of links which can change during the evolution of the system[69]. Each agent in the network adopts one of the two strategies of the Prisoner's Dilemma (PD). A cooperator (C) pays a cost c to provide a benefit b to each of its neighbours, where $b > c > 0$; however, the cheaters (or defectors, D) pay no cost and distribute no benefit. Specifically, if a cooperator has m cooperative neighbours and n defective neighbours, its payoff is $m(b - c) - nc$. However, a cheater in the same neighbourhood has payoff mb. The payoff matrix can be represented by:

$$\Pi = \begin{array}{c} C \\ D \end{array} \begin{pmatrix} \overset{C}{b-c} & \overset{D}{-c} \\ b & 0 \end{pmatrix} \tag{1}$$

At each update step, a node i is selected as a role-model to reproduce with a probability proportional to its effective payoff, namely the fitness f_i:

$$f_i = (1 + \delta)^{\pi_i} \tag{2}$$

where $\delta \geq 0$ specifies the strength of selection and π_i is the sum of payoffs through pairwise interactions to each neighbour. For $\delta = 0$ the selection probability is the same for all nodes, while increasing δ makes it more likely that a node with a higher payoff is selected to reproduce. In addition, a randomly chosen existing node is removed from the population, so that the number of nodes is constant. The evolutionary dynamics[73] for a system can be defined by a discrete sequence of update steps as shown in Fig. 1. The dynamics of the strategies can be depicted by differential equations, which are provided in the part of Methods.

We assume that the newcomer adopts the strategy of the chosen role-model (i.e., no mutation occurs), so the composition of population will reach a stationary state once we have $[C] = N$ or $[D] = N$. The parameters p and q are called embedding parameters, as they explicitly determine the ability of the newcomer to copy the role-model's social network, which means that the structure of network will always change even when the composition of population is steady.

Starting from an ER random network with all cooperators, we run the model for a large number of steps to remove transients and to establish a stable network topology. We then perturb the network by introducing few cheaters and the system is updated until one of the two outcomes is reached: either a case of recovery, the cheaters fail to invade and the system returns to the original state, or a case of transition, the cheaters invade successfully and the original state collapses[69]. Starting from a cooperative community, initially invaded by a single cheater, we illustrate the two possible outcomes in Fig. 2.

Network configurations. With a fixed $p = 0.6$, we vary the parameter $q \in [0, 1]$ to observe the collapse of cooperation and the max-size of invasion. The PD game parameters are $b = 10$ and $c = b/3$. Starting from an ER random network ($N = 100$) with average degree of 4, the network reaches stationary state before the introduction of invaders. 10000 independent and identical simulations are run for each q for a given size of invaders n. The network connectivity and degree diversity before cheaters' perturbation are illustrated in Fig. 3, where weak selection with $\delta = 0.01$ and strong selection with $\delta = 0.1$ are considered.

At strong selection with a larger δ, a node with a large amount of neighbours (i.e., a large payoff) at the initial stage will gain a larger fitness than that at weak selection (smaller δ) as $f_i = (1 + \delta)^{\pi_i}$ (See the more details in the Methods). In this way, a node with very high degree will be more likely to be selected as the role-model at strong selection, such that more and more newcomers will be connected to it (as $p = 0.6$) even when the embedding
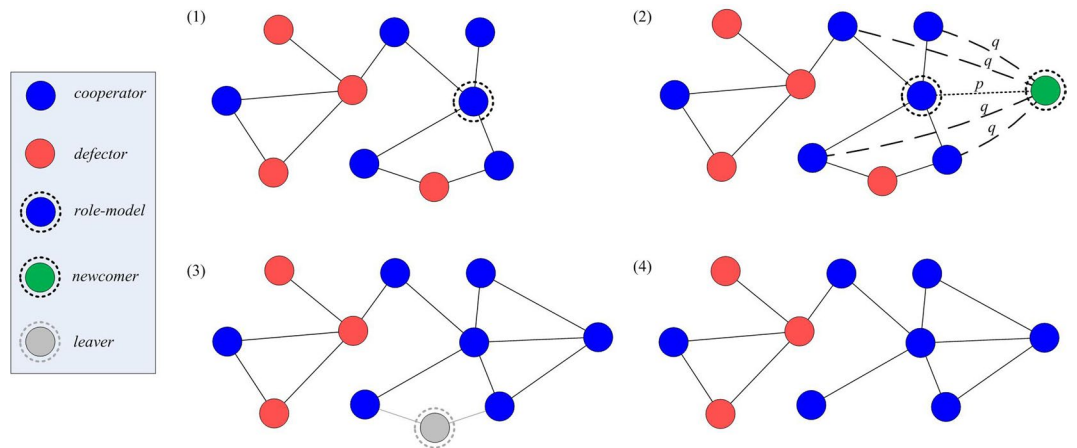
Figure 1. Co-evolution in a structured population. Each update step of the network follows these actions. (1) A role-model is selected to reproduce proportional to its effective payoff. (2) The newcomer connects to the role-model with probability p (dashed line), connects to each of its neighbours with probability q (dashed lines). p and q are called embedding parameters. (3–4) A randomly selected node and all its connections are removed from the network.
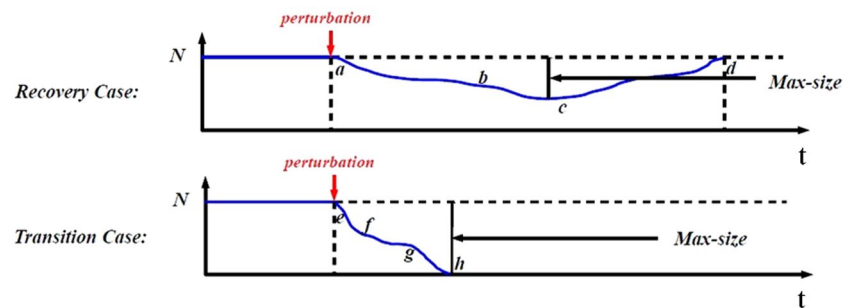

Figure 2. Perturbations lead to recoveries and transitions[69]. We show the two possible outcomes from a perturbation obtained by adding a single cheater in a network of only cooperators. The first row shows the typical case of an unsuccessful perturbation where cooperators resist, while the bottom row shows the typical case of a successful perturbation where cheaters invade gradually. The max-size of invasion is the maximal number of cheaters recorded during a perturbation.

parameter q is small. This phenomenon is the so-called *the rich get richer*. That is why we can find the relatively bigger average connectivity and degree diversity for small qs at strong selection. When the parameter q gets bigger, the network tends to be denser and less diverse for both weak and strong selections.

Ranking strategies.   We investigate several ways (*ranking strategies*) of considering n invading cheaters in the network and individuate the strategies which can facilitate the collapse of cooperation.

The initial n cheater invaders are placed in a set called *Top-n-set*. This set of selected invaders is constructed in the following way. At the beginning, we have an empty *Top-n-set*, and will insert nodes into it one by one until the size of *Top-n-set* reaches n. The nodes inserted in the *Top-n-set* are those ones selected as initial cheaters (invaders), and also called seed nodes.

In order to identify the ranking strategy which will lead to the most influential invaders, we compared the weighted degree decomposition with three classical and standard ranking methods (random ranking, max-degree ranking, and betweenness ranking). For random ranking, n random nodes in the network are selected to be the initial cheaters. For max-degree ranking, the *top-n* nodes are selected according to the ranking of their degree. For betweenness ranking, the *top-n* nodes are selected according to the ranking of their betweenness. For the ranking based on *weighted degree decomposition* (WDD), the weighted degree of a node is calculated by:

$$k^*(v) = k_p(v) + \Sigma_{u \in NS(v)}(\alpha + \beta|NS(u)|) \tag{3}$$

where $k_p(v)$ is the number of potential nodes in the neighbourhood, and $NS(v)$ is the set of neighbours who have been selected into the *Top-n-set* as seed nodes (see more details in the Methods). Specifically, two cases are considered:
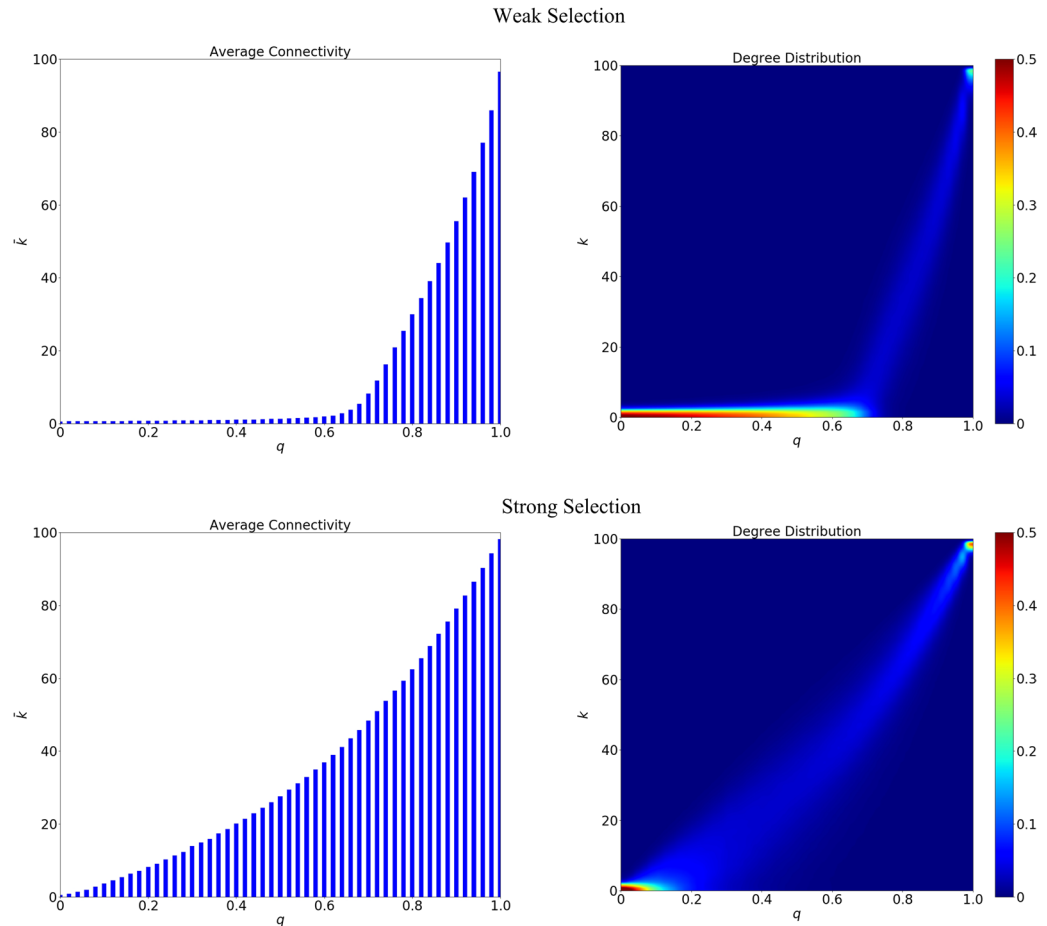
Figure 3. Network connectivity and degree distribution before the invasion of cheaters at weak selection ($\delta = 0.01$) and strong selection ($\delta = 0.1$) when q is changing from 0 to 1. For smaller qs, the network gets denser and more diverse at strong selection than that at weak selection due to the emergence of a few *hub* nodes, who are more and more likely to be selected to reproduce.

- positive-weighted degree decomposition: the seed neighbours (i.e. the neighbours who are selected as seed nodes) as well as their seed neighbours are assigned positive weights in the calculation of weighted degree, so that the nodes close to the high-degree clusters are going to be selected. In this way, the nodes with high-degree will not be missed, but the invaders are close to each other. Specifically, we implement the simulations with $\alpha = 1$ and $\beta = 0.1$.
- negative-weighted degree decomposition: the seed neighbours as well as their seed neighbours are assigned negative weights in the calculation of weighted degree, so that the nodes far away will be selected to avoid the overlap sphere of influence. In this way, some nodes with low-degree distant from the center will be selected, but some nodes with high-degree will be missed. Specifically, we implement the simulations with $\alpha = -1$ and $\beta = -0.1$.

We use different ranking strategies to select the initial cheaters in the cooperative community. After that, the population evolves according to the procedures shown in Fig. 1. As in the original network model[69] the selection of the role-model is still implemented with a probability proportional to the effective payoff (i.e., nodes with higher effective payoff have higher chances to be selected as role-models by the newcomers).

Weak selection case. At weak selection with $\delta = 0.01$, the probability for the collapse of cooperation is illustrated in Fig. 4. Here, four groups of simulations are carried out, which indicate the size of invaders is $n = 1$, $n = 5$, $n = 10$ and $n = 50$, respectively. As we can see, when the size of initial invaders n is small, the outcomes of random ranking are distinct from the others. The ranking strategies based on max-degree, betweenness, negative-weighted degree and positive-weighted degree induce a bigger loss of cooperation at bigger qs, which reveals the significance of the positions of invaders in dynamically evolving systems. When the size of n gets large, however, the differences between each other turn small. In particular, the ranking based on negative-weighted degree outperforms the rest due to the homogeneous network structure and small overlap sphere of influence. Interestingly, this result indicates that *exploration* of far away nodes should be paid more attentions in the case of weak selection.
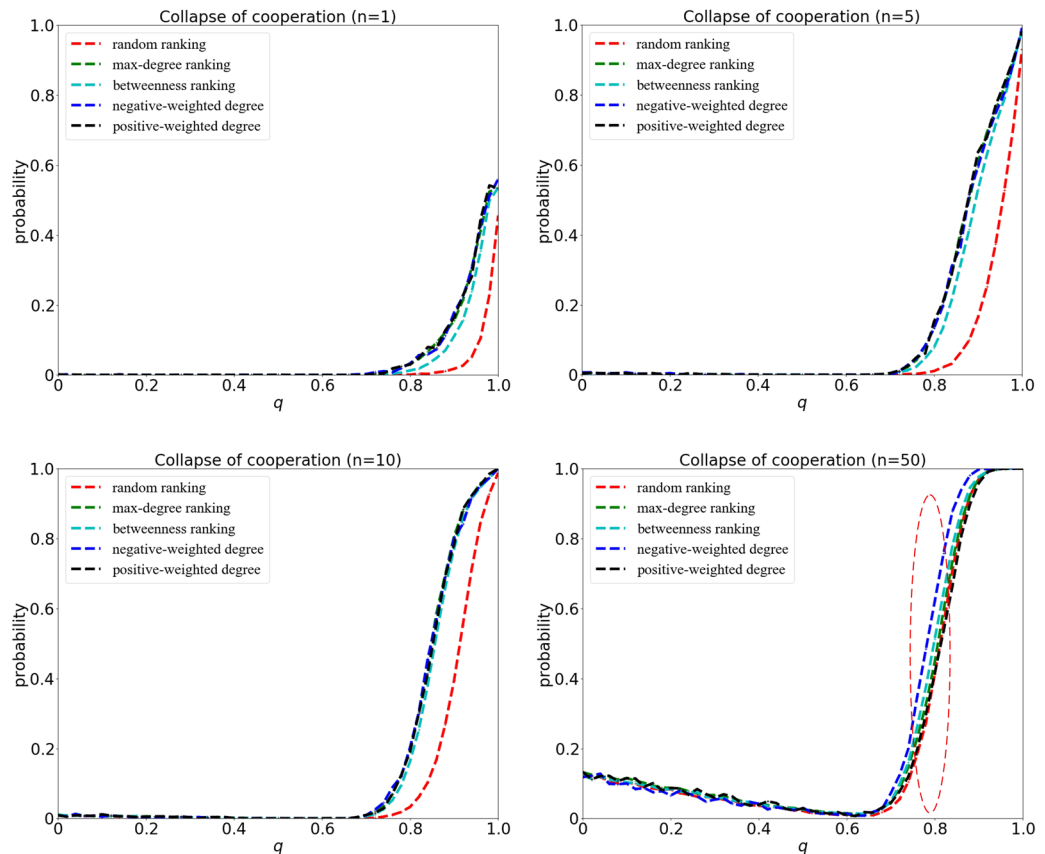
Figure 4. The collapse of cooperation for different ranking strategies at weak selection with $\delta = 0.01$. The probability for the collapse of cooperation is computed by the fraction of perturbations that lead to the successful invasion of cheaters out of 10000 perturbation experiments. The x-axis represents the embedding parameter q ranging from 0 to 1. The invaders identified by those heuristic and statistical ranking strategies (such as max-degree ranking, betweenness ranking, weighted degree ranking) can lead to a more evident shift of regime than the random one, especially when the value of n is not very big. However, the ranking based on negative-weighted degree outperforms others when n is large.

For a given $q = 0.8$, the embedding level of the newcomer agent is high, and we have the collapse of cooperation and max-size of invasion at weak selection in Fig. 5. As we can see, when the size of invaders n gets large enough, the ranking strategy based on negative-weighted degree can outperform other ranking strategies in the collapse of cooperation.

To take a deep insight into the difference caused by positive-weighted degree ranking and negative-weighted degree ranking, we obtained the degree distributions for cooperators and defectors at weak selection in Fig. 6. As we can see, a big distinction of degree distributions (k) for cooperators and defectors is present under the ranking of positive-weighted degree, while the degree distributions for cooperators and defectors are overlapping under the ranking of negative-weighted degree. As for the distributions of the size of cooperative neighbours (k_C), we can find a manifest difference under the ranking of negative-weighted degree, where the defectors hold more cooperative neighbours than the cooperators. On the contrary, for the distributions of the size of defective neighbours (k_D), the cooperators hold more defective neighbours than the defectors. As a consequence, the high-degree defectors are distributed over the network, which is dense and well-connected, leading to a bigger loss of cooperation under negative-weighted degree ranking. The snapshots illustrate that the defectors (red nodes) under the ranking of positive-weighted degree tend to be located in the *core*, while the defectors are more likely to keep distant from each other under the ranking of negative-weighted degree.

Strong selection case. At strong selection with $\delta = 0.1$, the probability for the collapse of cooperation is illustrated in Fig. 7, which is far from a linear pattern. As we can see, the collapse of cooperation is enhanced by the increase of selection, especially under the ranking of positive-weighted degree and max-degree. As the network structure is more heterogeneous before the invasion at strong selection (see Fig. 3), a few nodes will gain a quite large number of neighbours, but the degree of others is very small. When the size of initial invaders is sufficiently large, the ranking based on negative-weighted degree tries to avoid the gathering of invaders, which will put some low-degree and even isolated nodes far away from the *core* into the *Top-n-set*. However, the ranking based on positive-weighted degree will pay more attentions to the high-degree nodes close to the *core*, so the defectors located in the *interface* can benefit from the connections to the cooperators.
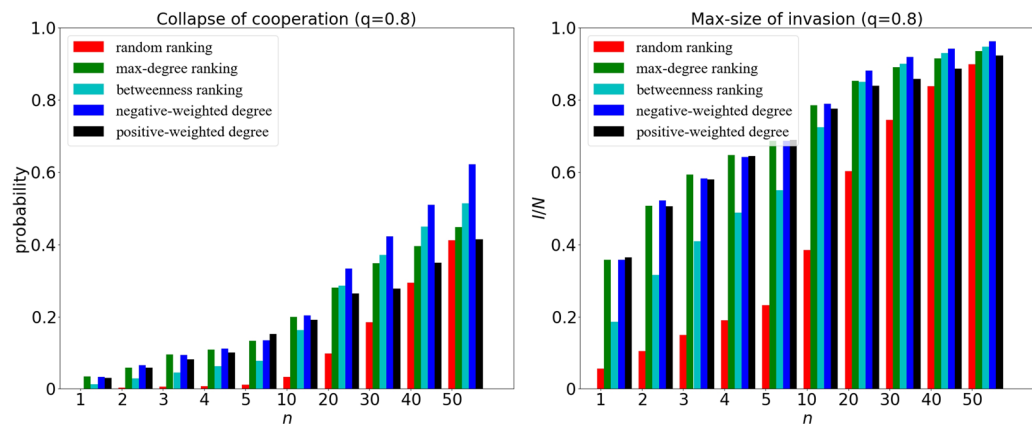
Figure 5. The collapse of cooperation (left panel) and max-size of invasion (right panel) for different ranking strategies at $q = 0.8$ and $\delta = 0.01$. The x-axis represents the size of initial cheaters n, varying from 1 to 50. Clearly, the ranking strategy based on negative-weighted degree makes the initial invaders distant from each other, leading to a bigger loss of cooperation, especially when n is large.
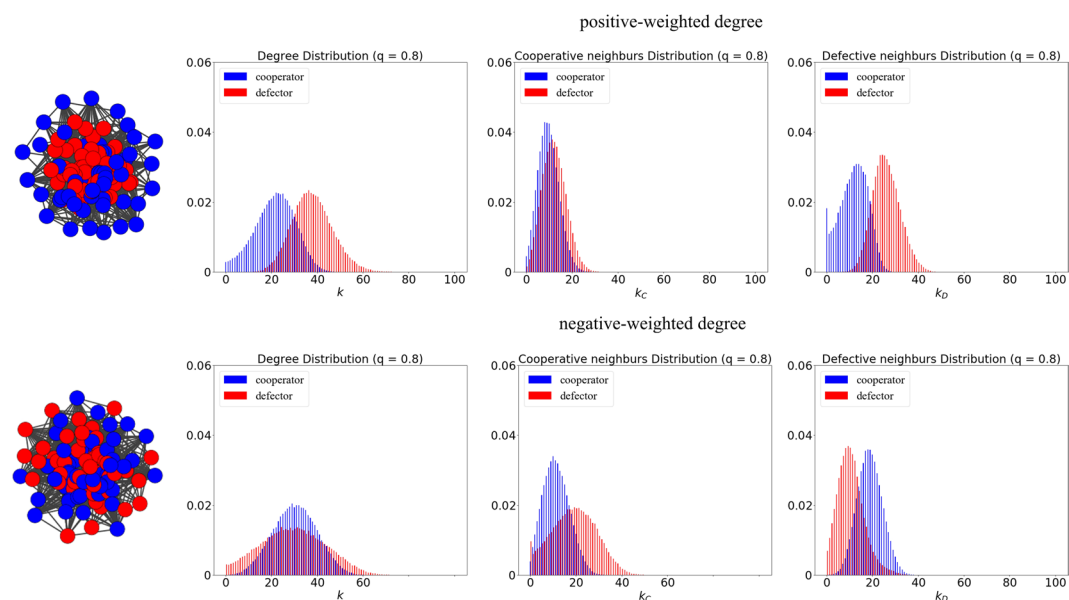

Figure 6. The degree distributions for cooperators and defectors at weak selection ($\delta = 0.01$) and high embedding ($q = 0.8$) based on positive-weighted degree ranking (upper panels) and negative-weighted degree ranking (lower panels). Two snapshots are obtained after the invasion of $n = 50$ cheaters. The left panels illustrate the distributions of degree (k) for cooperators (blue histograms) and defectors (red histograms), and the middle panels present the distributions of the sizes of cooperative neighbours (k_C), and the right panels show the distributions of the sizes of defective neighbours (k_D).

For a given $q = 0.2$, the network connectivity is low as a result of the small embedding level of the newcomer. We can also observe the collapse of cooperation and max-size of invasion at strong selection in Fig. 8. As we can see, when the size of invaders n is sufficiently large, the ranking strategy based on positive-weighted degree can outperform the other ranking strategies.

To take a deep insight into the differences caused by positive-weighted degree ranking and negative-weighted degree ranking, we obtained the degree distributions for cooperators and defectors at strong selection in Fig. 9. As we can see, a big distinction of degree distributions (k) for cooperators and defectors is present under the ranking of positive-weighted degree, while the degree distributions for cooperators and defectors are overlapping under the ranking of negative-weighted degree. This indicates that high-degree defectors are well identified through the ranking of positive-weighted degree. As the network connectivity is low and heterogeneity is high due to the small q and big δ, the CC connections are preserved better under the ranking of negative-weighted degree (see the distribution of the sizes of cooperative neighbours), and many defectors with low-degree have no access to the cooperators. On the contrary, the defectors under the ranking of positive-weighted degree have sufficient access to the cooperators, especially for those located in the *interface* between cooperators and defectors. As a consequence, a
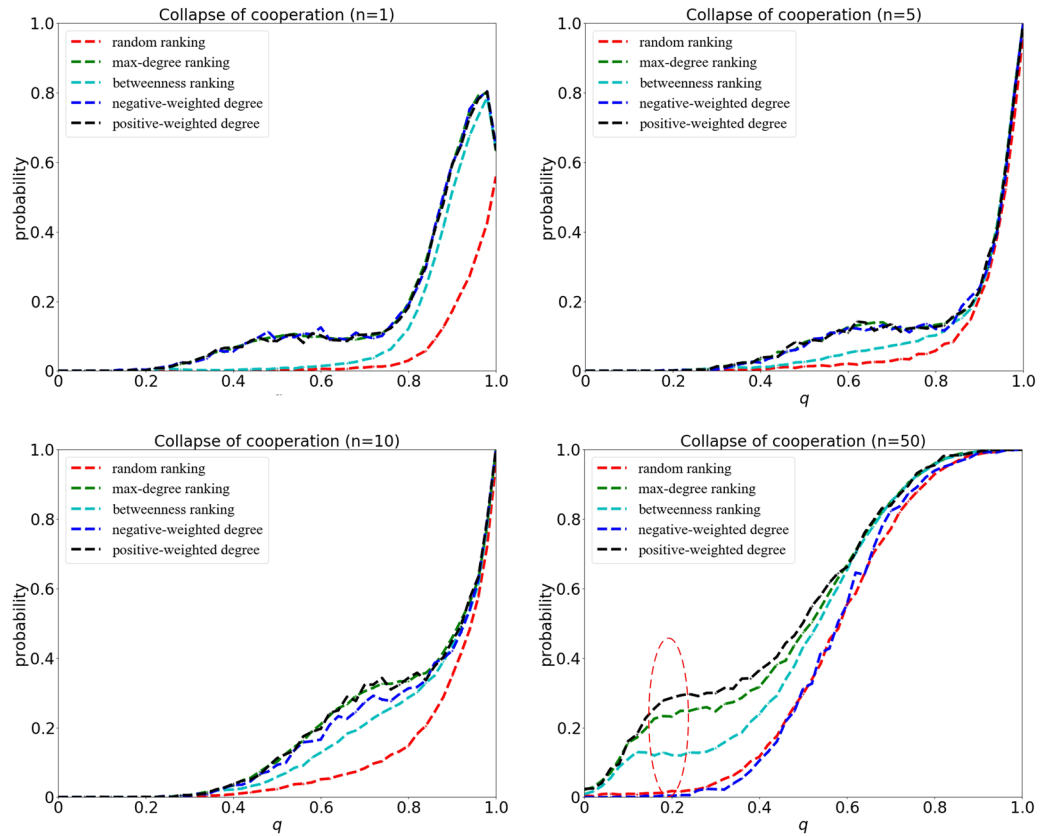
Figure 7. The collapse of cooperation for different ranking strategies at strong selection with $\delta = 0.1$. The x-axis represents the embedding parameter q, and the y-axis represents the probability for the collapse of cooperation. As we can see, a bigger loss of cooperation is obtained at strong selection even when the embedding parameter q is not large. The ranking based on positive-weighted degree works better than the others due to the formation of a heterogeneous network.
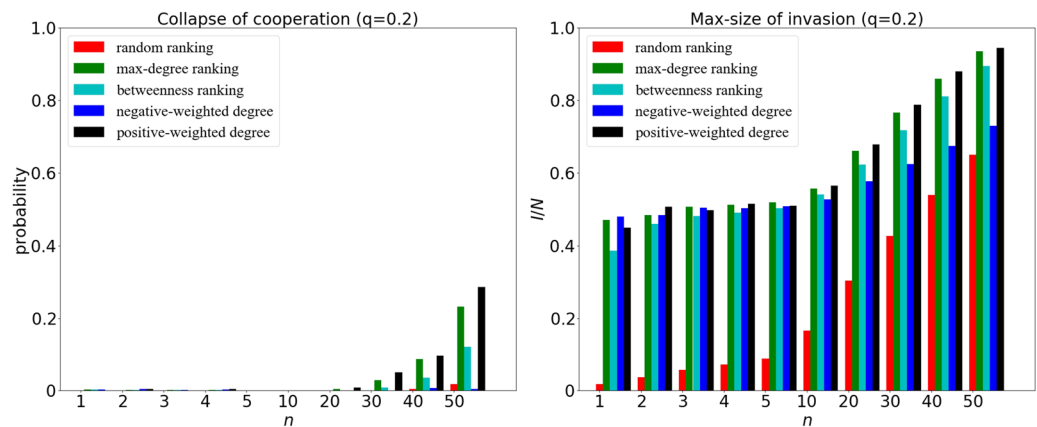

Figure 8. The collapse of cooperation (left panel) and max-size of invasion (right panel) for different ranking strategies at $q = 0.2$ and $\delta = 0.1$. The x-axis represents the size of initial cheaters n, varying from 1 to 50. Due to the strong selection and low embedding, the ranking strategy based on positive-weighted degree makes the initial invaders close to the *core*, leading to a bigger loss of cooperation especially when n is large.

bigger loss of cooperation under positive-weighted ranking is achieved at strong selection and low embedding. The snapshots illustrate that the cooperators (blue nodes) under the ranking of positive-weighted degree tend to be located on the periphery and maybe even isolated, while the defectors are too far away from the *core* and will more likely to be sparse or isolated under the ranking of negative-weighted degree.
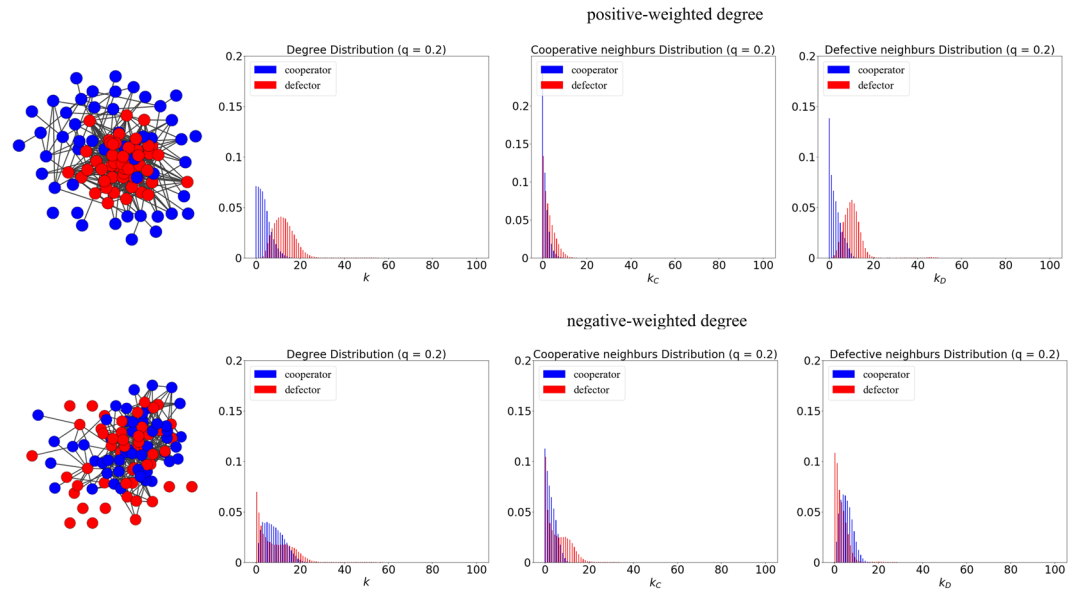
positive-weighted degree

negative-weighted degree

Figure 9. The degree distributions for cooperators and defectors at strong selection ($\delta = 0.1$) and low embedding ($q = 0.2$) based on positive-weighted degree ranking (upper panels) and negative-weighted degree ranking (lower panels). Two snapshots are presented after the invasion of $n = 50$ cheaters. The left panels illustrate the distributions of degree (k) for cooperators (blue histograms) and defectors (red histograms), and the middle panels present the distributions of the sizes of cooperative neighbours (k_C), and the right panels show the distributions of the sizes of defective neighbours (k_D).

Discussion

In dynamically evolving structured population, both the positions and the size of the initial invaders are crucial to the stability of the system. In this paper, we have considered the scenario of perturbation invasions where few initial invaders (cheaters) are added to the system (a cooperative community) and selected based on some proposed ranking strategies. As usual, under natural selection, nodes with bigger payoff are more likely to be selected to reproduce, especially when the strength of selection is bigger. In this way, the few initial cheaters can reproduce more often than the cooperators and can ultimately lead to the collapse of cooperation. In this paper weighted degree decomposition (WDD) is proposed as a way to rank the nodes and identify the most influential invaders in the cooperative population.

Interestingly, the ranking strategy based on negative-weighted degree works better at weak selection and for large embedding values. However, the ranking strategy based on positive-weighted degree outperforms the others at strong selection and for low embedding values due to the emergence of heterogeneous networks with low average degrees. This method incorporates the balance between dispersing and gathering into the identification of a set of influential nodes through a general framework, which indicates a trade-off between exploration and exploitation.

To conclude, this work provides a degree discount or reward heuristic for the identification of influential invaders in structured evolutionary populations, which paves a new way to study the influence maximization for dynamically evolving systems in the fields of biology, ecology and sociology. Despite this area is still in its infancy we expect to be a very promising future research line.

Methods

Analytical calculation. Let's derive the dynamics of the evolutionary population by approximate master equations (AME)[31], which are investigated to construct a system of closed star-like motifs. To capture the dynamics of the stars, we need to consider the following 8 types of dynamics, which are demonstrated through a motif $C_{m,n}$ (a cooperator with m cooperative neighbours and n defective neighbours).

- (a) the removal of $C_{m,n}$, leading to the decrease of $C_{m,n}$;
- (b) the removal of a neighbour C or D from $C_{m,n}$, leading to the decrease of $C_{m,n}$;
- (c) the removal of a neighbour C from $C_{m+1,n}$ or a neighbour D from $C_{m,n+1}$, leading to the increase of $C_{m,n}$;
- (d) the newcomer of $C_{m,n}$ connecting to it, leading to the decrease of $C_{m,n}$;
- (e) the newcomer of $C_{m-1,n}$ connecting to it, leading to the increase of $C_{m,n}$;
- (f) the newcomer of $C_{m,n}$'s neighbour C or D connecting to it, leading to the decrease of $C_{m,n}$;
- (g) the newcomer of $C_{m-1,n}$'s neighbour C or $C_{m,n-1}$'s neighbour D connecting to it, leading to the increase of $C_{m,n}$;
- (h) the newcomer of $C_{x,y}$ (where $x \geq m-1$, $y \geq n$) connecting to m cooperators and n defectors, leading to the increase of $C_{m,n}$;

Clearly, we have $[C] = \sum_x [C_x] = \sum_x \sum_y [C_{x,y}]$, where C_x means the set of cooperators who have x cooperative neighbours and $[C]$ is the total number of cooperators. Similarly, we have $[D] = \sum_x [D_x] = \sum_x \sum_y [D_{x,y}]$. As for the number of various edges, we can obtain them as follows:

$$[CC] = \sum_x \sum_y x \times [C_{x,y}] \qquad\qquad [CD] = \sum_x \sum_y y \times [C_{x,y}]$$

$$[DC] = \sum_x \sum_y x \times [D_{x,y}] \qquad\qquad [DD] = \sum_x \sum_y y \times [D_{x,y}]$$

$$[CCC] = \sum_x \sum_y x \times (x-1) \times [C_{x,y}] \qquad [CCD] = \sum_x \sum_y x \times y \times [C_{x,y}]$$

$$[CDC] = \sum_x \sum_y x \times (x-1) \times [D_{x,y}] \qquad [CDD] = \sum_x \sum_y x \times y \times [D_{x,y}]$$

$$[DCC] = \sum_x \sum_y x \times y \times [C_{x,y}] \qquad\qquad [DCD] = \sum_x \sum_y y \times (y-1) \times [D_{x,y}]$$

$$[DDC] = \sum_x \sum_y x \times y \times [D_{x,y}] \qquad\qquad [DDD] = \sum_x \sum_y y \times (y-1) \times [D_{x,y}]$$

Here, it is supposed that all neighbours holding same state are indifferent. For a motif $C_{m,n}$, the cooperative neighbour can be approximated by the configuration $C_{1+[CCC]/[CC],[CCD]/[CC]}$, and the defective neighbour can be approximated by $D_{1+[CDC]/[CD],[CDD]/[CD]}$. Similarly, for a motif $D_{m,n}$, the cooperative neighbour can be approximated by the configuration $C_{[DCC]/[DC], 1+[DCD]/[DC]}$, and the defective neighbour can be approximated by $D_{[DDC]/[DD],1+[DDD]/[DD]}$. When $\delta \to 0$, we have $(1+\delta)^\pi \approx 1 + \delta\pi$, so the sum of fitness can be calculated by:

$$\Sigma = \sum_i f_i = \sum_i (1+\delta)^{\pi_i} \approx \sum_i 1 + \delta\pi_i = N + \delta([CC] + [CD])(b-c) \tag{4}$$

Therefore, the dynamics for the motif $C_{m,n}$ can be depicted by the following equation:

$$\begin{aligned}
\frac{d}{dt}[C_{m,n}] =\ & -1/N[C_{m,n}] - (m+n)/N[C_{m,n}] + (m+1)/N[C_{m+1,n}] + (n+1)/N[C_{m,n+1}] \\
& - [C_{m,n}]\frac{1+\delta\pi(C_{m,n})}{\Sigma}(1-(1-p)(1-q)^{m+1}) + [C_{m-1,n}]\frac{1+\delta\pi(C_{m-1,n})}{\Sigma}p \\
& - m[C_{m,n}]\frac{1+\delta\pi(\overline{C})}{\Sigma}q + (m-1)[C_{m-1,n}]\frac{1+\delta\pi(\overline{C})}{\Sigma}q \\
& - n[C_{m,n}]\frac{1+\delta\pi(\overline{D})}{\Sigma}q + (n-1)[C_{m,n-1}]\frac{1+\delta\pi(\overline{D})}{\Sigma}q \\
& + \sum_{x \ge m-1, y \ge n}[C_{x,y}]\frac{1+\delta\pi(C_{x,y})}{\Sigma}\left(\binom{m}{x}q^m(1-q)^{x-m}(1-p)\binom{n}{y}q^n(1-q)^{y-n}\right. \\
& \left. + \binom{m-1}{x}q^{m-1}(1-q)^{x-m+1}p\binom{n}{y}q^n(1-q)^{y-n}\right)
\end{aligned}$$

Similarly, we can write the differential equations for $D_{m,n}$. The above derivations provide some analytical calculations for the dynamics of evolutionary population, by which we can estimate the stationary distributions of cooperators or defectors.

Now, let's investigate the influence of the strength of selection (δ) on the evolution of cooperation. Given a network with all nodes holding the state of cooperation, for $\delta_1 > \delta_2$, the probability that node i will be selected as the role-model at strong selection (namely δ_1) can be approximated by $\frac{1+\delta_1\pi_i}{N+\delta_1[CC](b-c)}$. Similarly, it will be selected as the role-model at weak selection (δ_2) is around $\frac{1+\delta_2\pi_i}{N+\delta_2[CC](b-c)}$. That node i will be more likely to be selected at strong selection if and only if:

$$\frac{1+\delta_1\pi_i}{N+\delta_1[CC](b-c)} > \frac{1+\delta_2\pi_i}{N+\delta_2[CC](b-c)} \tag{5}$$

Derive the above formula, we have that:

$$\begin{aligned}
\delta_2[CC](b-c) + \delta_1 N\pi_i &> \delta_1[CC](b-c) + \delta_2 N\pi_i \\
\frac{\delta_2}{\delta_1}[CC](b-c) + N\pi_i &> [CC](b-c) + \frac{\delta_2}{\delta_1}N\pi_i \\
\left(1 - \frac{\delta_2}{\delta_1}\right)N\pi_i &> \left(1 - \frac{\delta_2}{\delta_1}\right)[CC](b-c) \\
\pi_i &> [CC](b-c)/N
\end{aligned}$$

As we can see, at the initial state of evolution, a node with a payoff bigger than $[CC](b-c)/N$ will be more likely to be selected as the role-model at stronger selection.

Weighted degree decomposition. In evolving communities, the spreading of invaders is highly dependent on the large-size size of cooperative neighbours, meanwhile the positions of the invaders shouldn't be too close together in order to avoid the overlap sphere of influence. Here we will try to propose a new method, namely weighted degree decomposition (WDD), to identify the influential nodes in triggering the shift of regimes in structural population. Just like the method of K-shell[22] or mixed degree decomposition[24], we decompose the network gradually according to the weighted degree. The aim of this method is to find a *n*-element set, which will lead to the largest collapse of cooperation.

In general, two types of nodes will be in the network according to the weighted degree decomposition (WDD) method, and they are seed nodes and potential nodes respectively. In detail, we define them by:

- a seed node, who has been selected into the *Top-n-set* as one of the initial invaders;
- a potential node, who is not in the *Top-n-set*, that is to say all the rest are potential nodes to be selected.

Generally speaking, the potential neighbours usually provide positive feedback to the ego agent as they can be used for further spreading. As for the seed neighbours as well as the seed neighbours of seed neighbours, they are usually the nodes holding higher degree and more connections than the ego agent. Two situations need to be considered when we are counting the connections linking the seed nodes: one is *exploitation*, namely taking the links from the seed neighbours into consideration by assigning a positive weight, so that the high-degree seed nodes will get together; and the other is *exploration*, namely discounting the links from the seed neighbours by assigning a negative weight, so that the seed nodes will keep away from each other. In this paper, we compute the weighted degree of a node by the formula:

$$k^*(v) = k_p(v) + \sum_{u \in NS(v)} (\alpha + \beta |NS(u)|)$$

Interestingly, we have two weight coefficients α and β to make a trade-off between exploitation and exploration. As we can see, once a node has been selected as a seed, it will have a positive or negative influence on its neighbours by changing α and β. In order to avoid the overlap of sphere of influence, α and β will set to be negative, and the neighbours' degree of the seed nodes will be discounted, as a result some remote nodes will be explored. In order to get together to make a strong influence, α and β will set to be positive, and the neighbours' degree of the seed nodes will be overrated, as a result the nodes close to the cluster of high-degree seeds will be excessive exploited. It is α and β that control the weighted degree and the positions of the seeds in the network.

The detailed decomposition is implemented with the following procedure: (1) the node with the maximum weighted degree is selected into the *Top-n-set*; (2) update the states of nodes in the neighbourhood and recalculate the weighted degree for the remaining nodes; (3) repeat the above steps until the size of *Top-n-set* reaches *n*.

References
1. Asavathiratham, C., Roy, S., Lesieutre, B. & Verghese, G. The influence model. *IEEE Control. Syst.* **21**, 52–64 (2001).
2. Valente, T. W. Network models of the diffusion of innovations. *Comput. & Math. Organ. Theory* **2**, 163–164 (1996).
3. Young, H. P. The diffusion of innovations in social networks. *Gen. Inf.* **413**, 2329–2334 (2000).
4. Domingos, P. & Richardson, M. Matt Mining the network value of customers. In *International Conference on Knowledge Discovery and Data Mining*, 57–66 (2001).
5. Richardson, M. & Domingos, P. Mining knowledge-sharing sites for viral marketing. In *International Conference on Knowledge Discovery and Data Mining*, 61–70 (2002).
6. Budak, C., Agrawal, D. & El Abbadi, A. Limiting the spread of misinformation in social networks. *Games & Econ. Behav.* **70**, 194–227 (2010).
7. Aral, S. & Walker, D. Identifying influential and susceptible members of social networks. *Sci.* **337**, 337 (2012).
8. Leskovec, J. *et al.* Cost-effective outbreak detection in networks. In *International Conference on Knowledge Discovery and Data Mining*, 420–429 (2007).
9. Pastorsatorras, R. & Vespignani, A. Immunization of complex networks. *Phys. Rev. E* **65**, 036104 (2002).
10. Leskovec, J., Adamic, L. A. & Huberman, B. A. The dynamics of viral marketing. *ACM Transactions on Web* **1** (2007).
11. Albert, R., Albert, I. & Nakarado, G. L. Structural vulnerability of the north american power grid. *Phys. Rev. E* **69**, 025103 (2004).
12. Kempe, D., Kleinberg, J. & Tardos, E. Maximizing the spread of influence through a social network. In *International Conference on Knowledge Discovery and Data Mining*, 137–146 (2003).
13. Nemhauser, G. L., Wolsey, L. A. & Fisher, M. L. An analysis of approximations for maximizing submodular set functions. *Math. Program.* **14**, 265–294 (1978).
14. Goldenberg, J. & Libai, B. Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. *Mon. Labor Rev.* **31**, 8–11 (2001).
15. Granovetter, M. Threshold models of collective behavior. *Am. J. Sociol.* **83**, 1420–1443 (1978).
16. Chen, W., Wang, C. & Wang, Y. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *International Conference on Knowledge Discovery and Data Mining*, 1029–1038 (2010).
17. Chen, W., Wang, Y. & Yang, S. Efficient influence maximization in social networks. In *International Conference on Knowledge Discovery and Data Mining*, 199–208 (2009).
18. Goyal, A., Lu, W. & Lakshmanan, L. V. S. Celf++: optimizing the greedy algorithm for influence maximization in social networks. In *International Conference on World Wide Web*, 47–48 (2011).
19. Tang, Y., Xiao, X. & Shi, Y. Influence maximization:near-optimal time complexity meets practical efficiency. In *International Conference on Management of Data*, 75–86 (2014).
20. Tang, Y., Shi, Y. & Xiao, X. Influence maximization in near-linear time: A martingale approach. In *International Conference on Management of Data*, 1539–1554 (2015).
21. Lü, L. *et al.* Vital nodes identification in complex networks. *Phys. Rep.* **650**, 1–63 (2016).
22. Kitsak, M. *et al.* Identification of influential spreaders in complex networks. *Nat. Phys.* **6**, 888–893 (2010).
23. Chen, D., Lu, L., Shang, M., Zhang, Y. & Zhou, T. Identifying influential nodes in complex networks. *Phys. A* **391**, 1777–1787 (2012).
24. Zeng, A. & Zhang, C. J. Ranking spreaders by decomposing complex networks. *Phys. Lett. A* **377**, 1031–1035 (2013).
25. Freeman, L. C. A set of measures of centrality based on betweenness. *Sociom.* **40**, 35–41 (1977).
26. Newman, M. E. J. A measure of betweenness centrality based on random walks. *Soc. networks* **27**, 39–54 (2003).
27. Bonacich, P. Some unique properties of eigenvector centrality. *Soc. networks* **29**, 555–564 (2007).

28. Brin, S. The anatomy of large-scale hypertextual web search engine. *Comput. networks* **56**, 3825–3833 (2012).
29. Albert, R., Jeong, H. & Barabasi, A. L. Error and attack tolerance of complex networks. *Nat.* **340**, 378–382 (2000).
30. Cohen, R., Erez, K., Ben-Avraham, D. & Havlin, S. Breakdown of the internet under intentional attack. *Phys. Rev. Lett.* **86**, 3682 (2001).
31. Durrett, R. *et al.* Graph fission in an evolving voter model. *Proc. Natl. Acad. Sci.* **109**, 3682–3687 (2012).
32. Li, Y., Chen, W., Wang, Y. & Zhang, Z. L. Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships. In *International Conference on Web Search and Data Mining*, 657–666 (2013).
33. Smith, J. M. & Price, G. The logic of animal conflict. *Nat.* **246**, 15 (1973).
34. Weibull, J. W. *Evolutionary game theory*. (MIT press, 1997).
35. Hofbauer, J. & Sigmund, K. *Evolutionary games and population dynamics*. (Cambridge University Press, 1998).
36. Nowak, M. A. *Evolutionary Dynamics*. (Harvard University Press, Cambridge, MA, 2006).
37. Fudenberg, D. & Tirole, J. *Game theory*. (MIT Press, 1991).
38. Darwin, C. & Bynum, W. F. *The origin of species by means of natural selection*. (AL Burt, 2009).
39. Nowak, M. A. & May, R. M. Evolutionary games and spatial chaos. *Nat.* **359**, 826–829 (1992).
40. Abramson, G. & Kuperman, M. Social games in a social network. *Phys. Rev. E* **63**, 030901(R) (2001).
41. Masuda, N. & Aihara, K. Spatial prisoner's dilemma optimally played in small-world networks. *Phys. Lett. A* **313**, 55–61 (2003).
42. Fu, F., Liu, L.-H. & Wang, L. Evolutionary prisoner's dilemma on heterogeneous Newman-Watts small-world network. *Eur. Phys. J. B* **56**, 367–372 (2007).
43. Santos, F. C., Pacheco, J. M. & Lenaerts, T. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl. Acad. Sci. USA* **103**, 3490–3494 (2006).
44. Gómez-Gardeñes, J., Campillo, M., Floría, L. M. & Moreno, Y. Dynamical organization of cooperation in complex networks. *Phys. Rev. Lett.* **98**, 108103 (2007).
45. Rong, Z., Li, X. & Wang, X. Roles of mixing patterns in cooperation on a scale-free networked game. *Phys. Rev. E* **76**, 027101 (2007).
46. Masuda, N. Participation costs dismiss the advantage of heterogeneous networks in evolution of cooperation. *Proc. R. Soc. B* **274**, 1815–1821 (2007).
47. Assenza, S., Gómez-Gardeñes, J. & Latora, V. Enhancement of cooperation in highly clustered scale-free networks. *Phys. Rev. E* **78**, 017101 (2008).
48. Santos, F. C., Santos, M. D. & Pacheco, J. M. Social diversity promotes the emergence of cooperation in public goods games. *Nat.* **454**, 213–216 (2008).
49. Poncela, J., Gómez-Gardeñes, J. & Moreno, Y. Cooperation in scale-free networks with limited associative capacities. *Phys. Rev. E* **83**, 057101 (2011).
50. Tanimoto, J., Brede, M. & Yamauchi, A. Network reciprocity by coexisting learning and teaching strategies. *Phys. Rev. E* **85**, 032101 (2012).
51. Pinheiro, F., Pacheco, J. & Santos, F. From local to global dilemmas in social networks. *PLoS One* **7**, e32114 (2012).
52. Zimmermann, M. G., Eguíluz, V. M. & San Miguel, M. Coevolution of dynamical states and interactions in dynamic networks. *Phys. Rev. E* **69**, 065102(R) (2004).
53. Szabó, G. & Vukov, J. Cooperation for volunteering and partially random partnerships. *Phys. Rev. E* **69**, 036107 (2004).
54. Pacheco, J. M., Traulsen, A. & Nowak, M. A. Coevolution of strategy and structure in complex networks with dynamical linking. *Phys. Rev. Lett.* **97**, 258103 (2006).
55. Santos, F. C., Pacheco, J. M. & Lenaerts, T. Cooperation prevails when individuals adjust their social ties. *PLoS Comput. Biol.* **2**, 1284–1290 (2006).
56. Fu, F., Chen, X., Liu, L. & Wang, L. Promotion of cooperation induced by the interplay between structure and game dynamics. *Phys. A* **383**, 651–659 (2007).
57. Tanimoto, J. Dilemma solving by coevolution of networks and strategy in a 2×2 game. *Phys. Rev. E* **76**, 021126 (2007).
58. Fu, F., Hauert, C., Nowak, M. A. & Wang, L. Reputation-based partner choice promotes cooperation in social networks. *Phys. Rev. E* **78**, 026117 (2008).
59. Fu, F., Wu, T. & Wang, L. Partner switching stabilizes cooperation in coevolutionary prisoner's dilemma. *Phys. Rev. E* **79**, 036101 (2009).
60. Lee, S., Holme, P. & Wu, Z.-X. Emergent hierarchical structures in multiadaptive games. *Phys. Rev. Lett.* **106**, 028702 (2011).
61. Gómez-Gardeñes, J., Romance, M., Criado, R., Vilone, D. & Sánchez, A. Evolutionary games defined at the network mesoscale: The public goods game. *Chaos* **21**, 016113 (2011).
62. Pinheiro, F., Santos, M. D., Santos, F. & Pacheco, J. Origin of peer influence in social networks. *Phys. Rev. Lett.* **112**, 098702 (2014).
63. Mitrović, M. & Tadić, B. Dynamics of bloggers' communities: Bipartite networks from empirical data and agent-based modeling. *Phys. A* **391**, 5264–5278 (2012).
64. Tadić, B., Gligorijević, V., Mitrović, M. & Šuvakov, M. Co-evolutionary mechanisms of emotional bursts in online social dynamics and networks. *Entropy* **15**, 5084–5120 (2013).
65. Perc, M. & Szolnoki, A. Coevolutionary games – a mini review. *BioSystems* **99**, 109–125 (2010).
66. Pacheco, J. M., Vasconcelos, V. V. & Santos, F. C. Climate change governance, cooperation and self-organization. *Phys. Life Rev.* **11**, 573–586 (2014).
67. Wang, Z., Wang, L., Szolnoki, A. & Perc, M. Evolutionary games on multilayer networks: a colloquium. *Eur. Phys. J. B* **88**, 124 (2015).
68. Perc, M. *et al.* Statistical physics of human cooperation. *Phys. Rep.* **687**, 1–51 (2017).
69. Cavaliere, M., Guoli, Y., Vincent, D. & Vasilis, D. Detecting the collapse of cooperation in evolving networks. *Sci. Rep.* **6**, 30845 (2016).
70. Pastor-Satorras, R., Smith, E. & Solé, R. V. Evolving protein interaction networks through gene duplication. *J. Theor. Biol.* **222**, 199–210 (2003).
71. Wagner, A. How the global structure of protein interaction networks evolves. *Proc. R. Soc. B* **270**, 457–466 (2003).
72. Chung, F., Lu, L., Dewey, T. G. & Galas, D. J. Duplication models for biological networks. *J. Comput. Biol.* **10**, 677–687 (2003).
73. Cavaliere, M., Sedwards, S., Tarnita, C. E., Nowak, M. A. & Csikász-Nagy, A. Prosperity is associated with instability in dynamical networks. *J. Theor. Biol.* **299**, 126–138 (2012).

Acknowledgements

Author Contributions

G.Y., T.P.B., M.C., J.H. and M.P. designed research, performed research, analyzed data, and wrote the paper.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.