Integrating emotion-imitating into strategy learning improves cooperation in social dilemmas with extortion

Ji Quan [a,*], Yawen Zhou [a], Xiaojian Ma [a], Xianjia Wang [b,**], Jian-Bo Yang [c]

[a] School of Management, Wuhan University of Technology, Wuhan 430070, China
[b] School of Economics and Management, Wuhan University, Wuhan 430072, China
[c] Alliance Manchester Business School, The University of Manchester, Manchester M15 6PB, UK

Abstract. Extortion strategy can play the role of a Trojan horse for cooperators and act as a catalyst for the evolution of cooperation in social dilemma games. Based on the prisoner's dilemma game model with extortion, an emotion-imitating rule for strategy updating is proposed in this study, which takes into account the diverse attitudes of a player towards a strong or weak opponent. Furthermore, by employing evidential reasoning, we establish an improved rule for strategy updating, which integrates the emotion-imitating rule into strategy learning as well as the original myopic best response rule, and via a fusion weight, the role of emotion-imitating can be subtly adjusted in the process of strategy learning. The effect of emotion-imitating on cooperation is investigated in the presence of extortioners. Through Monte Carlo simulations, it is shown that the new rule can significantly boost cooperation when the fusion weight is close to 0.5. In strategy learning, the emotion-imitating rule relieves the direct exploitation from defectors for cooperators and also facilitates cooperators to cluster and spread in the population. The extortion strategy provides shelter for cooperators to survive even at the extremely high temptation to defect. In addition, the myopic best response rule can also improve the performance of the emotion-imitating rule, especially in the region of high temptation to defect. Based on evidential reasoning, the new rule integrates sophisticatedly the advantage of the two single rules, enhancing the emergence of cooperation as well as the average payoffs of the whole system. The robustness of the shelter effects of extortioners under the new learning rule model is also studied by considering action errors and the cognition cost of extortioners respectively. The results further support the catalyst roles of extortioners, provided that the dilemma strength is not too strong. We also demonstrate the enhancement of cooperation when the model is extended to four competing strategies where the tit for tat (TFT) strategy is also included. The coexistence of extortioners and TFT players restricts the expansion of defectors, and the extortioners' catalyst role still works prominently in the interval of low tempt values.

Keywords: Evolutionary games; cooperation; evidential reasoning; strategy learning

1. Introduction

Cooperation has been one major and popular way in individual interactions and behaviors from animals to humans [1]. However, since Darwin, there remains a long-standing conundrum to be unraveled: how can cooperative behavior evolve among selfish individuals? [2]. Intuitively, since maximizing one's fitness is the basis for rational individuals to survive, one's cooperative behavior,

* Corresponding author.
** Corresponding author.
E-mail addresses: quanji123@163.com (J. Quan), wangxj@whu.edu.cn (X.J. Wang)

giving up one's benefit to help others, runs counter to Darwin's natural selection [3]. Over the past few decades, this conundrum has attracted numerous researchers, ranging from sociologists, biologists, and mathematicians, to physicists [2, 4-7].

A highlight in the evolution of cooperation is the successful strategies for inducing cooperation, such as classical tit-for-tat and so on in Axelrod's popular computer tournament. Recently, Press and Dyson unexpectedly discovered a special class of ultimatum strategies in a two-person iterated prisoner's dilemma game (PDG) [8], called zero-determinant (ZD) strategies, whereby players can unilaterally enforce a claim to an unfair share of rewards. Extortion, as a subset of ZD strategies, has recently attracted considerable attention [9-13]. Hilbe et al. [14] proved that extortion can act as catalysts for the evolution of cooperation in large populations, but they are not the stable outcome of natural selection in a well-mixed population. In the field of evolutionary games, it is revealed that ZD strategies, including extortion, are evolutionarily unstable [9]. However, the evolutionary viability of extortion in structured populations has been explored by Szolnoki and Perc [11, 15], finding that extortion can be evolutionary stable if the strategy updating is governed by the myopic best response rule. Hence, the evolutionary dynamics of extortion in structured populations are still worthy of attention.

In addition to the introduction of extortion, strategy update rules also play a crucial role in the emergence of cooperation, such as the above myopic best response. The kernel of the strategy update rule is a kind of learning mechanism that describes how players learn to act to maximize their payoffs in repeated interactions. They can learn by themselves or learn from the surrounding environment to reduce uncertainty for making decisions [16]. Up to now, many strategy update rules have been proposed, such as the win-stay-lose-shift rule [17-19], Fermi function-based imitation rule [20, 21], aspiration-driven rule [22], myopic best response rule [15, 23], Moran process-based rule [24-27], and PSO-based rule [28-30].

However, in the PDG without extortion, the traditional rule of strategy-imitating has little help for the emergence of cooperation, and cooperators can hardly survive even at a very low temptation to defect [15]. Emotions are ineluctably related to individuals' actions. Recent research has clearly emphasized the role of emotion in the promotion of prosocial behavior [31, 32]. In this paper, we propose an emotion-imitating rule in the context of a three-strategy PDG model with extortion. In contrast with the traditional strategy learning rule, it has been revealed that in spatial games, imitating emotion profiles like goodwill and envy, instead of pure strategies, elevates social welfare [33], and in random graphs and scale-free networks, the resolution of social dilemmas is facilitated by imitating not strategies but more successful players' emotion [34]. Emotion decisions also work well in spatial public goods games [35]. Inspired by this conception, we integrate emotion-imitating into strategy learning, and further employ evidence theory to fuse information from both the myopic best response rule [15] and the emotion-imitating rule, to build a new rule for strategy updating.

Recent research has also focused on the role of reasoning in the behavioral updating process [36]. Especially, individuals' behavioral evolution driven by the cognition process has been studied [37-39] and its promoting role in the evolution of cooperation has also been demonstrated [40]. Inspired by these studies, we further consider the integration of reasoning into individuals' decision-making process in cooperative dilemmas. Evidence theory, which is a generalization of Bayesian inference [41, 42] and features in modeling imprecision and uncertainty is adopted. Associated with the evolutionary games [43-50], it has recently been introduced to fuse local environment, global information, or both for players to enhance the effectiveness of decisions [43-45]. Aimed at

problems of multiple attribute decision-making, evidence theory provides a valuable method for tackling complex uncertainty [51-54]. Thereby, evidence theory is instrumental for individuals to update their strategies. Specifically, by learning strategy respectively from the emotion-imitating rule and myopic best response rule, individuals can acquire two aspects of information as two separate pieces of evidence from two different sources. And by evidential reasoning, the two pieces of evidence can be fused into a new piece of evidence to provide more reliable information. Via the new piece of evidence, individuals can handle uncertainty in the game to adopt more effective strategies. Moreover, to surmount some inherent defects in traditional Dempster's combination rule, we adopt a special evidence reasoning rule [55]. Our work focuses on the effect of emotion-imitating on cooperation in the presence of extortioners, comparing with the case of the single update rule of myopic best response [15].

The rest of this paper is skeletonized as follows. Firstly, we introduce the spatial PDG model with extortion and integrate emotion-imitating into strategy learning via evidential reasoning to construct a new strategy learning rule. Then in Section 3, we demonstrate our results of simulations and immediately discuss the main arguments about the new rule. Finally, we summarize our main work and conclusions.

2. Model

2.1 Spatial PDG with extortion

The extortion strategy is introduced in the spatial PDG model. In a $L \times L$ square lattice network with periodic boundary conditions, each player is located on a node connected with 4 directed neighbors (denoted as Ω), and the links between them represent the interaction relationship[56, 57]. Initially, individuals play one role of defector (D), cooperator (C), or extortioner (Eχ) at random with equal probability. For one game, a cooperator will pay cost c and finally produce benefit b for himself and other players ($b > c > 0$), while a defector takes a free ride. Especially, extortioners can impose a linear unilateral relation to the payoffs of their opponents with a certain exploitation strength, resulting in forcibly sharing the benefit of cooperators. As such, cooperators have to passively accept extortion from extortioners. Following the game parameterization of previous studies [10, 15, 58], the payoff π_{ij} for a focal player i against an opponent j ($j \in \Omega_i$) can be given as the following matrix:

$$M_1 = \begin{pmatrix} & \mathrm{D} & C & E_\chi \\ D & 0 & b & 0 \\ C & -c & b-c & \dfrac{b^2-c^2}{b\chi+c} \\ E_\chi & 0 & \dfrac{(b^2-c^2)\chi}{b\chi+c} & 0 \end{pmatrix}, \tag{1}$$

where the linear relation of benefit-sharing between a cooperator and an extortioner is formulated as $\frac{b^2-c^2}{b\chi+c}$ and $\frac{(b^2-c^2)\chi}{b\chi+c}$. $\chi\,(\chi > 1)$ adjusts the intensity of an extortioner's extortion against a cooperator. According to Ref. [59-61], the strength of dilemma can be classified as the gamble-intending dilemmas (GID) and risk-averting dilemmas (RAD), and the dilemma strength is closely

related to benefit b and cost c. Apart from that, it is generally assumed that $b-c=1$ for simplification without loss of generalization. Thus, for the PDG model with extortion, there are only two main parameters, namely, b and χ controlling the game model. The interaction between a cooperator and an extortioner is somehow like a snowdrift game, but the extortioner has a silver bullet to defeat its opponent. Instead, once the extortioner encounters a defector, they both get nothing. In addition, at every stage of evolution, it is not just one round of game in the spatial game, but player i will play the game with all directed neighbors (Ω_i) simultaneously to reap his final payoff, i.e., $\pi_i = \sum_{j\in\Omega_i} \pi_{ij}$.

2.2 Learning rules

Under the strategy update rule of myopic best response, it has been recently proved that introducing extortion into traditional PDG can induce the emergence of cooperation. Nevertheless, the individual-based rule for strategy learning also plays a vital role in the evolution of cooperation. We propose a new rule via evidential reasoning, which integrates emotion-imitating into the original rule of myopic best response for individuals' strategy learning. Specifically, emotion-imitating considers an individual's emotional profile, not just traditional simple strategy-imitating. From the perspective of evidence theory, the information of strategy learning from the two rules are regarded as two pieces of evidence, and by evidential reasoning, individuals can depend on both of the two aspects of information to make a more reliable decision in the game.

2.2.1 Emotion-imitating based rule

We first extract the first piece of evidence from the emotion-imitating rule. In the traditional pairwise imitation, each player x has a pure strategy ($s_x = C \text{ or } D$). Here, instead of pure strategies, we introduce an emotional profile $e_x = \{S_x, \alpha_x; W_x, \beta_x\}$ to each player, which covers the different attitudes T_x toward its opponent y, depending upon the opponent's success quantified by the payoff value. Particularly, for player x and its opponent y in a round of the game, let π_x and π_y denote their corresponding payoffs respectively. Based on the above three-strategy model, we suppose that there are three attitudes T_x for player x towards its opponent y: defection or extortion, cooperation or extortion, and cooperation or defection, coded by 0, 1, and 2 respectively. Each attitude contains a set of two strategies for player x to choose from. When facing a strong or a weak opponent, player x may take different attitudes to make a decision. Let S_x and W_x represent the different attitudes toward a strong opponent if $\pi_x < \pi_y$ and a weaker opponent if $\pi_x \geq \pi_y$, respectively, where $S_x, W_x \in \{0, 1, 2\}$. Thus, the emotional profile e_x can also be expressed in the form of T_x

$$e_x = \{S_x, \alpha_x; W_x, \beta_x\} \Rightarrow T_x \Rightarrow \begin{cases} \{S_x, \alpha_x\}, & \text{if } \pi_x < \pi_y \\ \{W_x, \beta_x\}, & \text{if } \pi_x \geq \pi_y \end{cases}, \tag{2}$$

where a_x denotes the possibility, with which player x adopting the first strategy for the corresponding attitude (a set of two strategies) or adopting the other strategy with the possibility of

$1-\alpha_x$ towards a strong opponent, and β_x towards a weak opponent. Take $e_x = \{1, 0.2; 0, 0.3\}$ for example. If $\pi_x < \pi_y$, then player x encounters a strong opponent. For a strong opponent with $S_x = 1$, player x takes the second attitude (to choose whether to cooperate with its opponent or to extort its opponent) and will adopt the first strategy (cooperation) with the possibility of $\alpha_x = 0.2$ and adopt another strategy (extortion) with the possibility of $1-\alpha_x$. If $\pi_x \geq \pi_y$, then player x takes the attitude $\{W_x, \beta_x\}$ towards a weak opponent. Under attitude $T_x = \{W_x = 0, \beta_x = 0.3\}$, he will defect his opponent with the possibility of $\beta_x = 0.3$ and extort his opponent with the possibility of $1-\beta_x$.

Similar to the process of the traditional pairwise imitation, the emotion-imitating rule also calculate the possibility P_1 for imitation based on the difference of payoffs between player i and one randomly selected neighbor j, described as (here for clarification we let i and j denote the focal player and its neighbor as a mock object, not the above x and its opponent y)

$$P_1(e_j \to e_i) = \frac{1}{1+\exp(-\frac{\pi_j - \pi_i}{\kappa})}, \tag{3}$$

where π_i and π_j denote the payoff of player i and the selected neighbor j, and their corresponding emotion profiles are e_i and e_i. $\kappa = 0.05$ indicates noise intensity quantifying uncertainty in the strategy imitation [15]. Thus, player i will imitate the selected neighbor with the above possibility P_1.

Differently, in the process of imitating, player i can copy a partial or whole emotion profile of the neighbor j at random, depending upon two randomly generated numbers γ_1 and γ_2 ($\gamma_1, \gamma_2 \sim U(0,1)$). After player i imitates the neighbor j, there are four results for the emotion profile of i as

$$e_i \Rightarrow \begin{cases} \{S_i, \alpha_i; W_i, \beta_i\}, & if\ P_1 < \gamma_1\ and\ P_1 < \gamma_2 \\ \{S_i, \alpha_i; W_j, \beta_j\}, & if\ \gamma_1 < P_1 < \gamma_2 \\ \{S_j, \alpha_j; W_i, \beta_i\}, & if\ \gamma_2 < P_1 < \gamma_1 \\ \{S_j, \alpha_j; W_j, \beta_j\}, & if\ \gamma_1 \leq P_1\ and\ \gamma_2 \leq P_1 \end{cases}. \tag{4}$$

The above results mean that if $\gamma_1 < P_1 < \gamma_2$, player i will only copy the attitude towards a weak opponent ($\{W_j, \beta_j\}$) of neighbor j; if $\gamma_2 < P_1 < \gamma_1$, player i will only copy the attitude towards a strong opponent ($\{S_j, \alpha_j\}$) of neighbor j; if $\gamma_1 \leq P_1\ and\ \gamma_2 \leq P_1$, player i will copy the emotion profile $\{S_j, \alpha_j; W_j, \beta_j\}$ of neighbor j. If $P_1 < \gamma_1\ and\ P_1 < \gamma_2$, the emotion profile of player i will remain unchanged.

Based on the emotion-imitating rule, we can refine the information of strategy learning used for evidential reasoning, as listed in Table 1. In the game model of three strategies, the decision of player i is driven by its emotional profile. Encountering a strong or weak opponent, player i will take different attitudes T_i. Under a certain T_i, the strategy that player i will choose and its corresponding possibility constitute the effective BPA, namely Body of Evidence.

Table 1 Evidence extracted from the emotion-imitation rule

π_y vs π_i	T_i	s_i	Body of Evidence
	$\{S_i = 0, \alpha_i\}$	D or E_χ	$m_1(D) = \alpha_i$, $m_1(E_\chi) = 1 - \alpha_i$
$\pi_y > \pi_i$	$\{S_i = 1, \alpha_i\}$	C or E_χ	$m_1(C) = \alpha_i$, $m_1(E_\chi) = 1 - \alpha_i$
	$\{S_i = 2, \alpha_i\}$	C or D	$m_1(C) = \alpha_i$, $m_1(D) = 1 - \alpha_i$
	$\{W_i = 0, \beta_i\}$	D or E_χ	$m_1(D) = \beta_i$, $m_1(E_\chi) = 1 - \beta_i$
$\pi_y \le \pi_i$	$\{W_i = 1, \beta_i\}$	C or E_χ	$m_1(C) = \beta_i$, $m_1(E_\chi) = 1 - \beta_i$
	$\{W_i = 2, \beta_i\}$	C or D	$m_1(C) = \beta_i$, $m_1(D) = 1 - \beta_i$

i denotes the focal player and y denotes its opponent.

2.2.2 Myopic best response rule

The first piece of evidence for strategy learning is learnt from others to acquire information. The second piece of evidence is learnt by self-reflection to explore other strategies according to the myopic best response rule. Actually, this rule is one of the heuristic strategies based on stochastic learning theory [62]. The core of the rule is that it allows individuals to adopt strategies that are not even present in their opponents when they are not satisfied with the benefit reaped by adopting the current strategy. In the process of Monte Carlo simulations, a randomly selected player i with strategy s_i reaps its payoff π_i, and then switches its strategy s_i to another strategy $\tilde{s}_i$ ($\tilde{s}_i$ is drawn randomly from the remaining two strategies) with probability

$$P_2(\tilde{s}_i \to s_i) = \frac{1}{1 + \exp(-\frac{\tilde{\pi}_i - \pi_i}{\kappa})}, \tag{5}$$

where $\tilde{\pi}_i$ is the payoff of the same player gained by adopting $\tilde{s}_i$ within the same neighborhood, and κ is the same as that in the emotion-imitating rule.

Based on the above information for the strategy learning, the body of this piece of evidence for learning can be extracted as illustrated in Table 2, which describes all possible situations that player i may encounter, including possible strategies and the corresponding probability, namely the body of evidence.

Table 2 Evidence profiled from the myopic best response rule

$\tilde{s}_i$	s_i	Body of Evidence
D or E_χ	C	$m_2(D) = P_2(\tilde{s}_i \to s_i)$ or $m_2(E_\chi) = P_2(\tilde{s}_i \to s_i)$, $m_2(C) = 1 - P_2(\tilde{s}_i \to s_i)$

$C\ or\ E_\chi$	D	$m_2(C)=P_2(\tilde{s}_i \to s_i)\ or\ m_2(E_\chi)=P_2(\tilde{s}_i \to s_i)$, $m_2(D)=1-P_2(\tilde{s}_i \to s_i)$
$C\ or\ D$	E_χ	$m_2(C)=P_2(\tilde{s}_i \to s_i)\ or\ m_2(D)=P_2(\tilde{s}_i \to s_i)$, $m_2(E_\chi)=1-P_2(\tilde{s}_i \to s_i)$

2.3 Strategy updating based on evidential reasoning

At last, we utilize evidential reasoning to refine the above two pieces of evidence. Two sources of strategy learning provide comprehensive information for individual-based strategy updating. Actually, individuals' strategy updating is a process of decision-making, finally choosing one strategy of C, D, or Eχ in a game. Aimed at the decision problem with uncertainty, two separate pieces of evidence are listed in Table 3, representing the basic belief distribution of different strategies; m_1 is from the emotion-imitating rule and m_2 is from the myopic best response rule.

Table 3 Evidence combination

Strategy s_i	Emotion-imitating m_1	Myopic Best Response m_2	Combination m_{12}
C	$m_1(C)$	$m_2(C)$	$m_{12}(C)$
D	$m_1(D)$	$m_2(D)$	$m_{12}(D)$
E_χ	$m_1(E_\chi)$	$m_2(E_\chi)$	$m_{12}(E_\chi)$

Fig. 1 presents the schematic of the proposed strategy learning rule in structure populations. Considering the importance and reliability of evidence m_1 and m_2, we let μ_1 and μ_2 denote their reliability, while λ_1 and λ_2 are for their importance, respectively. In order to integrate the two pieces of evidence, m_{12}, the combination of m_1 and m_2 is formulated as

$$m_{12}(\theta)=[m_1 \oplus m_2](\theta)=\begin{cases} 0 & \theta=\varnothing \\ \dfrac{\tilde{m}_{12}(\theta)}{\sum_{\theta\subseteq\Theta}\tilde{m}_{12}(\theta)} & \theta\neq\varnothing \end{cases}, \tag{6}$$

Our focus is the belief of the three strategies, C, D, and Eχ. Thus, the effective proposition θ belongs to $\{C, D, E_\chi\}$. Let $\tilde{m}_1(\theta)=\lambda_1 m_1(\theta)$ and $\tilde{m}_2(\theta)=\lambda_2 m_2(\theta)$. As a result, $\tilde{m}_{12}(\theta)$ is as follows.

$$\tilde{m}_{12}(\theta)=\left[(1-\mu_2)\tilde{m}_1(\theta)+(1-\mu_1)\tilde{m}_2(\theta)\right]+\sum_{A\cap B=\theta}\tilde{m}_1(A)\tilde{m}_2(B)\ . \tag{7}$$

To a large extent, this special rule is superior to Dempster's combination rule, for it disposes of the Zadeh paradox [63].

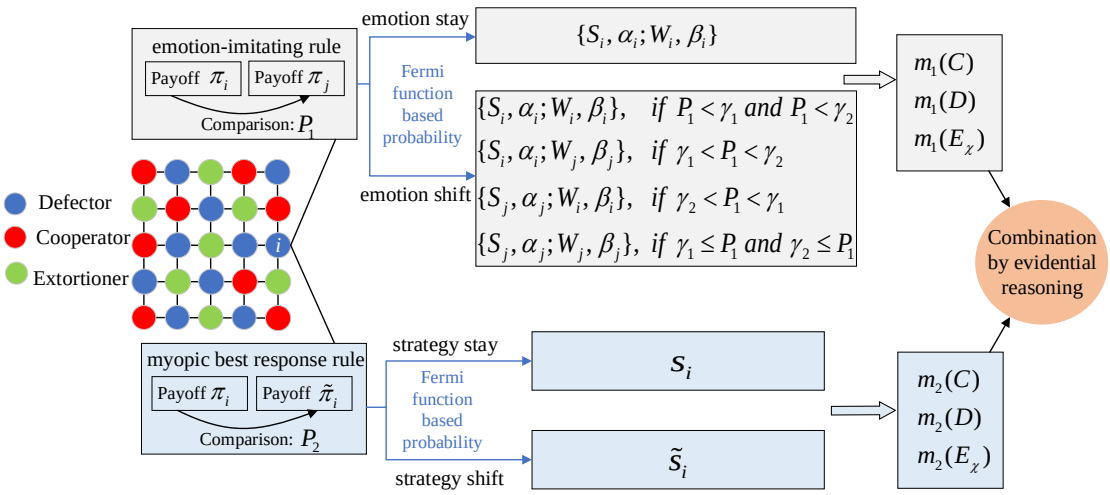

Fig. 1 Schematic diagram of the model. Individuals play the spatial PDG with roles of defector, cooperator, and extortioner. The strategy updating of each individual is driven by two learning rules. The emotion-imitating rule regulates individuals' attitudes toward a stronger or a weaker opponent, the attitudes contain the possibility to adopt different strategies. The myopic best response rule regulates the possibility to adopt one random remaining strategy. Evidential reasoning is utilized to provide comprehensive information for individuals' strategy updating.

In addition, for the sake of reasonable simplification, we make the following assumption: $\lambda_1 = \mu_1$, $\lambda_2 = \mu_2$ and $\lambda_1 + \lambda_2 = 1$. Furthermore, we use a brief symbol ω to represent both importance λ_1 and reliability μ_1 for the evidence from the emotion-imitating rule, i.e. $\omega = \lambda_1 = \mu_1$; we regard ω as a comprehensive fusion weight for the evidential reasoning to tune the result of strategy learning. The final refined piece of evidence m_{12} for strategy updating is given as

$$m_{12}(C) = \frac{\tilde{m}_{12}(C)}{\tilde{m}_{12}(C) + \tilde{m}_{12}(D) + \tilde{m}_{12}(E_\chi)}, \tag{8}$$

$$m_{12}(D) = \frac{\tilde{m}_{12}(D)}{\tilde{m}_{12}(C) + \tilde{m}_{12}(D) + \tilde{m}_{12}(E_\chi)}, \tag{9}$$

$$m_{12}(E_\chi) = \frac{\tilde{m}_{12}(E_\chi)}{\tilde{m}_{12}(C) + \tilde{m}_{12}(D) + \tilde{m}_{12}(E_\chi)}, \tag{10}$$

where according to equation (7) and above assumptions for simplification, the combination $\tilde{m}_{12}(\theta)$ of $m_1(\theta)$ and $m_2(\theta)$ are $\tilde{m}_{12}(C) = \left[(1-\omega)^2 m_1(C) + \omega^2 m_2(C)\right] + \omega(1-\omega) m_1(C) * m_2(C)$,

$\tilde{m}_{12}(D) = \left[(1-\omega)^2 m_1(D) + \omega^2 m_2(D)\right] + \omega(1-\omega) m_1(D) * m_2(D)$, and

$\tilde{m}_{12}(E_\chi) = \left[(1-\omega)^2 m_1(E_\chi) + \omega^2 m_2(E_\chi)\right] + \omega(1-\omega) m_1(E_\chi) * m_2(E_\chi)$, respectively.

3. Results and discussion

According to the above model, we conducted sufficient Monte Carlo simulations to investigate the effect of the new strategy update rule on the evolution of cooperation, especially the strategy learning based on emotion-imitating. Particularly, for the initial status of the population, we adopt the random initialization configuration, whereby $L \times L$ agents of the random role of defector, cooperator, or extortioner are distributed uniformly on the square lattice at the beginning of the

evolution. Subsequently, they start to play the PDG with extortion in their neighborhood step by step as the evolution of the population develops gradually. The results and their related discussion are reported as follows. It is worth mentioning that we uniformly display the results on a system size of $L = 100$, for the sake of clear representation, but in fact, we have also verified the validation of our arguments on larger systems of $L = 200$ *and* 500 [64]. The frequencies of the strategies are obtained by averaging over the final 1,000 steps of the entire 10,000 Monte Carlo simulations on the premise that the system is at the equilibrium state. In most cases, 10,000 Monte Carlo steps are enough to stabilize the system in our simulation. If the number of any strategy in the last 1000 rounds changes more than 1%, the iteration will continue until the change in the proportion of strategies in the last 1000 rounds is less than 1%. We take the average of the last 1000 rounds of iteration as the result after the system is stable. Furthermore, all results are taken from an average of 10 independent realizations of simulation, with each realization repeated over 10 times, to get rid of the influence of accidental error or noise from numerical simulations of Monte Carlo.

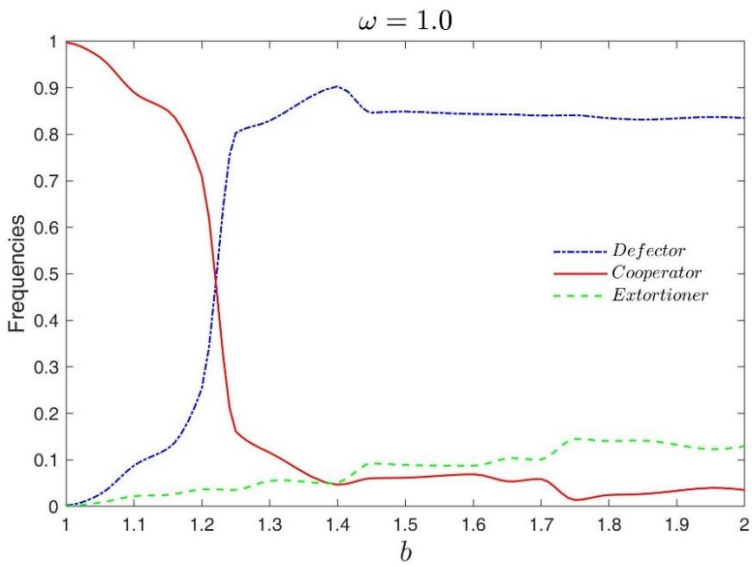

Fig. 2 With $\chi = 1.5$ and $\omega = 1.0$, average frequencies of defectors, cooperators, and extortioners vary as b increases, in which the model degenerates to the emotion-imitating based learning. In the presence of extortioners, cooperators can survive in the whole region of b. Emotion-imitating strategy induces cooperators to form clusters and prevail in the population at a low b but high temptation to defect can still disintegrate them.

Firstly, we probe into the evolutionary behavior of the pure emotion-imitating rule. We let the system evolve towards a steady-state, in which the average frequencies of strategies become time-independent. As is graphed in Fig. 2, as b increases, defectors soar sharply and hold the upper hand in the population, for defectors can gain a higher and higher payoff without any contribution in the game, so that at a very high temptation to defect, more and more players will turn into defectors, while the curve for the frequency of cooperators drops sharply. But it is worth noting that in the interval of low b, cooperators unexpectedly occupy the entire population even though the temptation to defect still exists. In the model of two strategies (cooperation and defection), cooperators have a slim chance to survive under the temptation of high payoffs. However, with the introduction of extortioners, it turns out that cooperation emerges and persists throughout the whole interval of b. Extortioners seemingly extort cooperators by an enforced linear relation between their payoffs, but in a sense, they can act as a Trojan horse to shield cooperators from the direct exploitation of defectors. Besides, we also notice that the frequency of extortioners rises steadily

with the increase of b. Defectors have little influence on extortioners while cooperators are passively extorted when the frequency of defectors keeps nearly unchanged. Thus, in the emotion-imitating rule, individuals can imitate different attitudes to avoid the strong exploitation of defectors. Extortion strategy can also play a role in inducing cooperation, and in the presence of extortioners, cooperators are also able to prevail within the interval of low b. But under a high temptation to defect, blind imitation accelerates the spread of defectors in the population. When we fix the fusion weight to 0, the mode of strategy learning degenerates into the original myopic best response rule in ref [15]. Fig. 3 shows that our result is consistent with that of Szolnoki and Perc. Thus, the new rule is compatible with the original myopic best rule. Actually, the myopic best response rule is a kind of self-learning mechanism, whereby an individual explores other strategies to seek a higher payoff. Hence, in the shelter of extortioners, cooperators are not extinct because they can timely switch to extortioners, or even expose themselves and spread once numerous extortioners appear among the defectors like a Trojan horse. In addition, in the region of high temptation to defect, the frequency of cooperators is higher than that in Fig. 2, and the myopic best response rule performs a little better than the emotion-imitating rule.

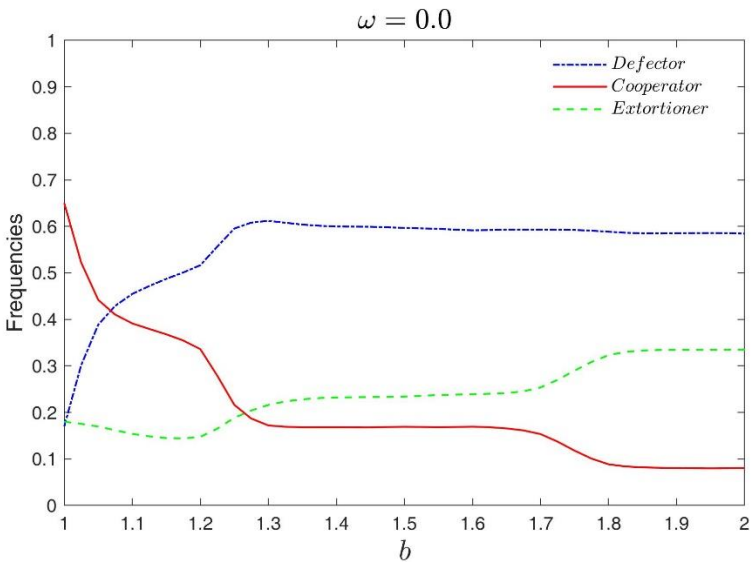

Fig. 3 With $\chi = 1.5$ and $\omega = 0.0$, average frequencies of defectors, cooperators, and extortioners vary as b increases. The result of the myopic best response rule in the previous study [15] is reproduced here. Compare with the emotion-imitating rule, the myopic best response rule performs a little better in the region of high b.

When we gradually integrate the emotion-imitating rule into strategy learning by increasing the fusion weight ω, it is shown that cooperation can be improved remarkably, as plotted in Fig. 4. From 0.1 to 0.5, we notice that the curve for the frequency of cooperators is going up while the curve for the frequency of defectors declines continuously. Especially, when $\omega = 0.5$, in most parts of the whole interval of b, cooperators dominate the whole population. At the same time, extortioners also strengthen their power by extorting cooperators. On the contrary, defectors lose their advantage. Although defection strategy helps defectors gain a high payoff without a cost when encountering cooperators, this situation will not last for a long time. Because their payoffs are rooted in the contribution of cooperators and the return of the game, and if all defect or extort, no one will be better. Therefore, with the help of extortioners who indirectly alleviate the exploitation from greedy defectors and the emotion-imitating rule that enables cooperators to spread quickly in the population, prosperous cooperation emerges and prevails as in the real world. And cooperation is

boosted best when the two single rules are fused with a weight near 0.5. As for the case of $\omega = 0.7$ where the weight of the emotion-imitating rule is higher than that of the myopic best response rule, the frequencies of defectors, cooperators, and extortioners tend to approach those of $\omega = 1.0$ in Fig. 2. But the myopic best response rule that accounts for a weight of 0.3 also plays a good role in improving the overall situation. Therefore, the new rule based on evidential reasoning can combine the advantages of two single rules and make them promote each other in the emergence of cooperation.

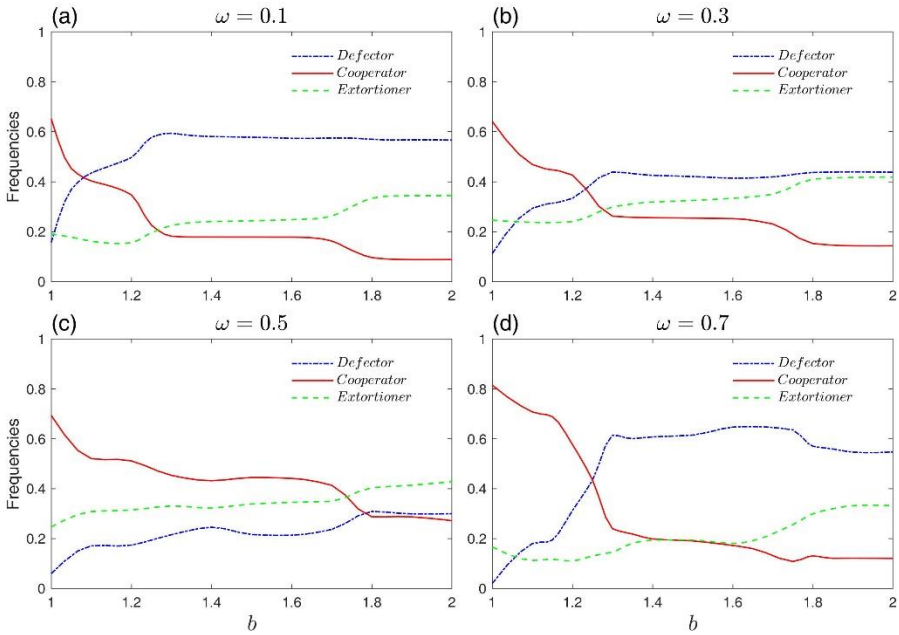

Fig. 4 Average frequencies of defectors, cooperators, and extortioners dependent on b, with $\chi = 1.5$ for different fusion weights $\omega = 0.1, 0.3, 0.5$ *and* 0.7. Unexpectedly, integrating emotion-imitating into strategy learning improves cooperation but suppresses defection significantly even at a high b.

Furthermore, we demonstrate more details of the evolutionary process from the perspective of strategy distribution on the network and the fluctuation of the frequency of cooperators. In order to examine the impact of the fusion weight on dynamic patterns, we select three different typical values of $\omega = 0.1, 0.3$ *and* 0.5, as displayed in Fig. 5 and Fig. 6. As the fusion weight increases step by step, the stationary frequency of cooperators rises. In the process of evolution, defectors always first take the upper hand but cooperators catch up finally. Moreover, in the view of strategy distribution on the network, defectors locally gather together into clusters, while cooperators and extortioners are twisted together. Especially, when $\omega = 0.5$ at a time step of 10000, in the local area, the snowdrift game relation between cooperators and extortioners results in a checkerboard ordering [15], where extortioners do not have to interact with players of their kind, and cooperators and extortioners are neighbors to each other. In this way, cooperators can completely avoid the exploitation of defectors. Apart from that, aimed at different cases of $b = 1.1, 1.5$ *and* 1.9, we examine the influence of b on evolutionary behavior, as graphed in Fig. 7 and Fig. 8. As b increases from 1.1 to 1.5, the frequency of cooperators is declining. At the beginning of evolution, defectors, cooperators, and extortioners are clustering respectively, and after a sufficiently long relaxation time, they interweave with each other, forming the checkerboard order. But a high temptation to defect facilitates defectors and extortioners to form clusters, and subsequently destroys the checkerboard order. Especially, when $b = 1.9$, extortioners occupy the entire population.

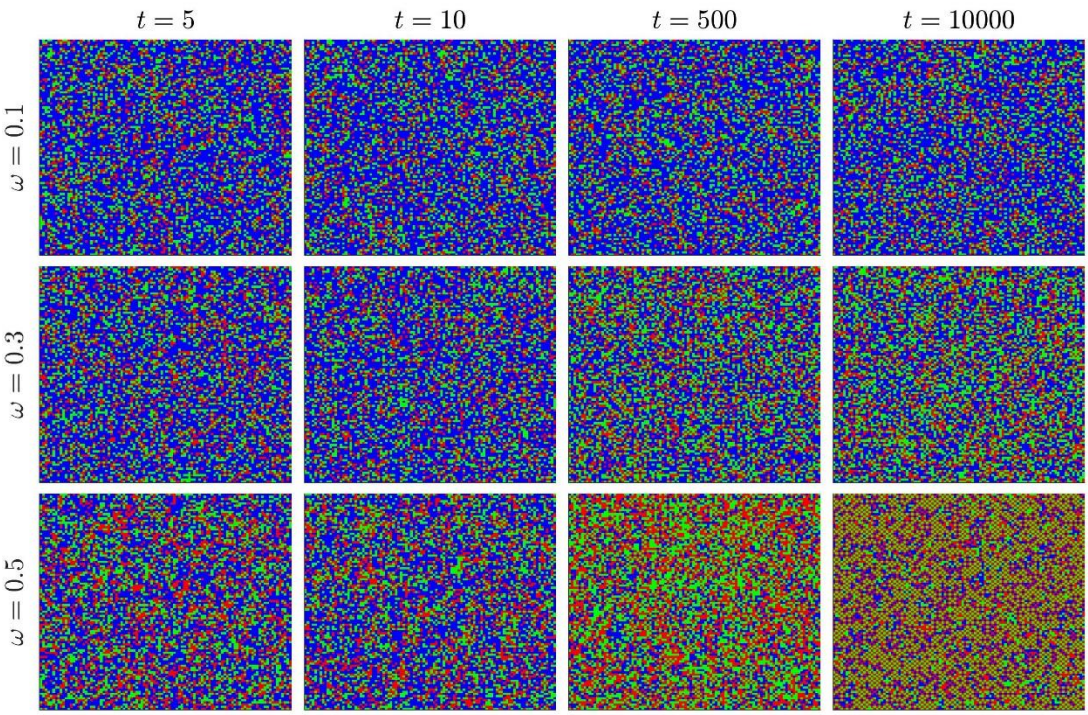

Fig. 5 Snapshots of strategy distribution in the network at time steps $t = 5$, 10, 500, and 10000, with different fusion weight values ω of 0.1, 0.3, and 0.5 for $b = 1.5$ and $\chi = 1.5$. Defectors, cooperators, and extortioners are represented by blue, red, and green colors respectively. The snowdrift game relation between cooperators and extortioners gives rise to a checkerboard order when $\omega = 0.5$ at the time step of 10000.

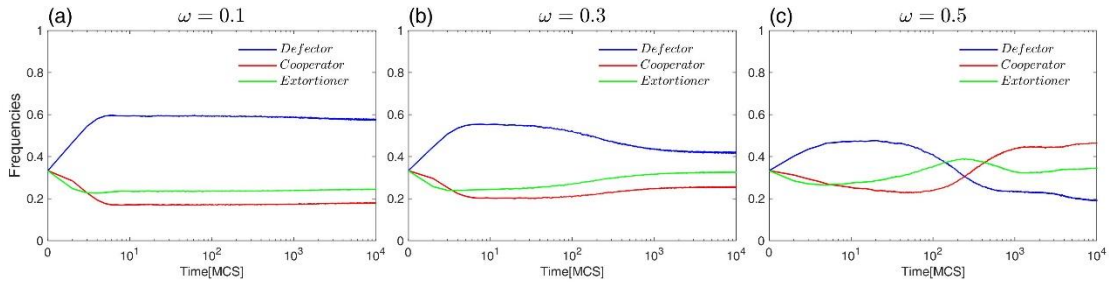

Fig. 6 Frequencies of defectors, cooperators, and extortioners towards its steady-state with different values of fusion weight $\omega = 0.1, 0.3$ *and* 0.5 for $b = 1.5$ and $\chi = 1.5$.

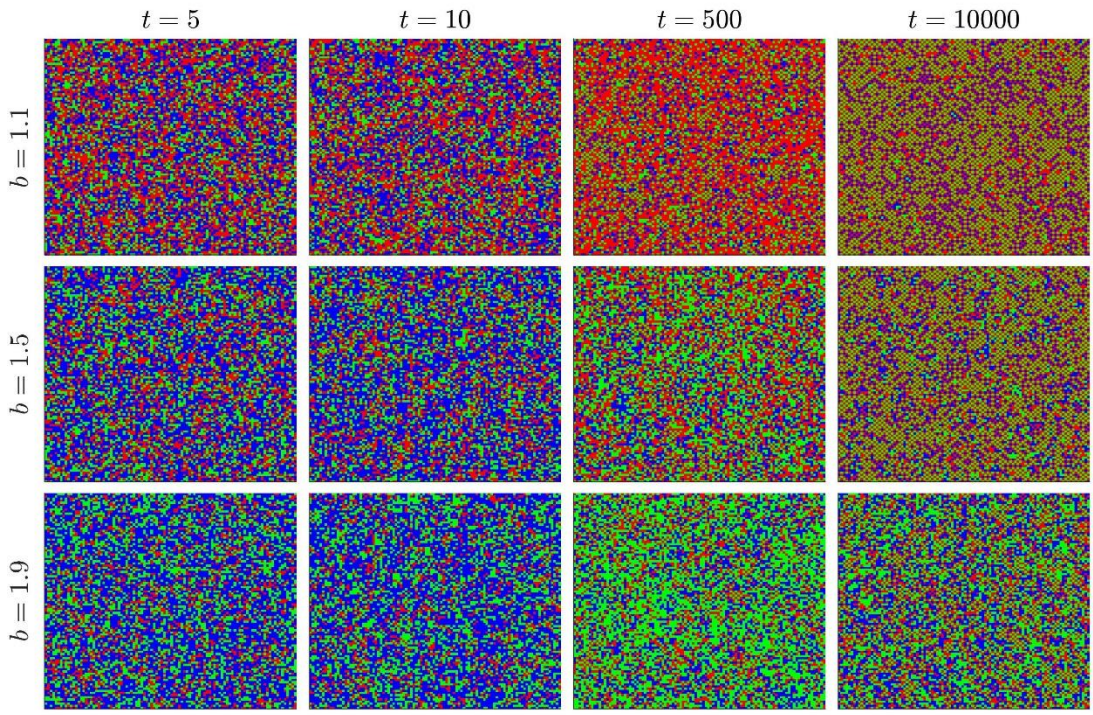

Fig. 7 Snapshots of strategy distribution in the network at time steps $t = 5, 10, 500$ *and* 10000 with different values of $b = 1.1, 1.5$ *and* 1.9 for $\omega = 0.5$ and $\chi = 1.5$. Defectors, cooperators, and extortioners are represented by blue, red, and green colors respectively. A high b facilitates defectors and extortioners to form clusters and then undermines the checkerboard order.

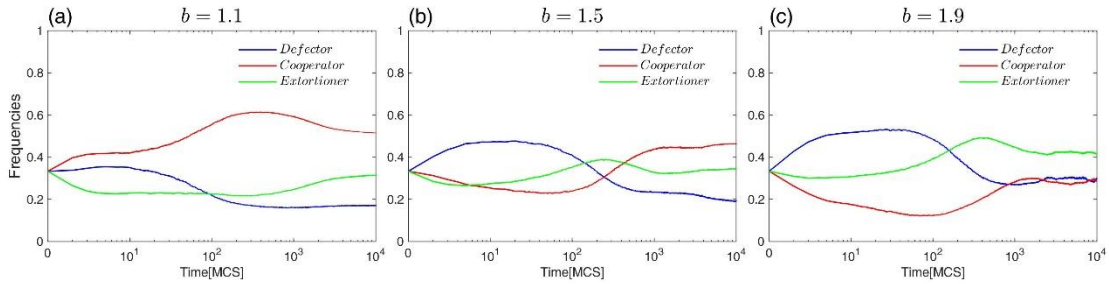

Fig. 8 Frequencies of defectors, cooperators, and extortioners towards its steady-state, with different values $b = 1.1, 1.5$ *and* 1.9 for $\omega = 0.5$ and $\chi = 1.5$.

Next, we display the effect of the new strategy learning that combines the two single rules, as visualized in Fig. 9 and Fig. 10. $\omega = 0$ represents the result of the myopic best response rule, consistent with that of Szolnoki and Perc [15]. We discover that with the increase of b, the frequency of cooperators decreases slowly. In addition, the panel of parameter space can be divided into three parts (cyan area, blue area, and deep blue area), and in the same area, the changes of parameter pairs do not have a great influence on the cooperation level. As for $\omega = 1$, the panel presents the situation of two opposing extremes (deep red area for low b and deep blue area for high b), because the emotion-imitating rule expands the influence of a successful strategy, which means a successful strategy can be strengthened and stronger. Therefore, in the region of low b, cooperators with the shelter of extortioners can quickly occupy the whole population. But in the region of high b, cooperators are very hard to survive under the fierce exploitation of defectors but never disappear. The results are consistent with that of $\omega = 1$ in Fig. 2, displaying the features of

the emotion-imitating rule. In Fig. 10, we can see that the new rule can improve the frequency of cooperators in the whole panel of parameter space when ω is from 0.1 to 0.5, with the deep blue area shrinking and the green area expanding. In addition, comparing the situation of $\omega = 0.7$ with that of $\omega = 1$ in Fig. 9, the cooperation level of $\omega = 0.7$ is still enhanced in the region of high b. Evidential reasoning can subtly fuse the two single rules into strategy learning and make them complement each other.

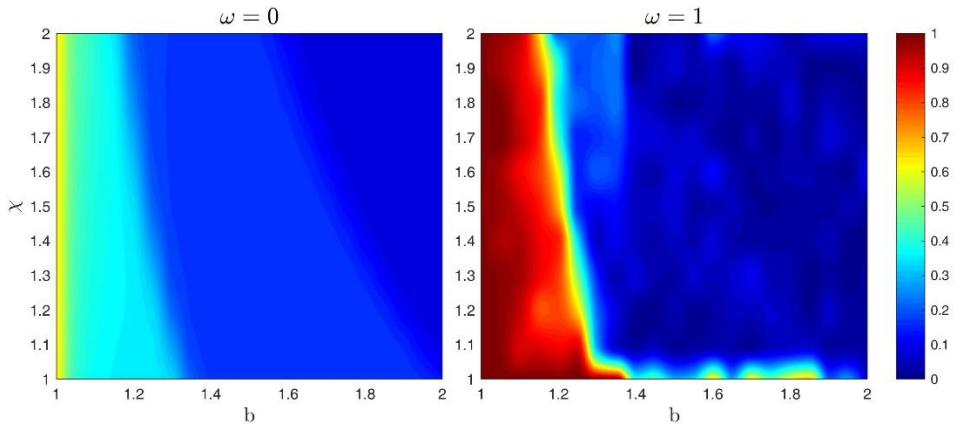

Fig. 9 Color-coded average frequency of cooperators on the $\chi - b$ parameter plane for $\omega = 0.0$ (the myopic best response rule) and $\omega = 1.0$ (the emotion-imitating rule).

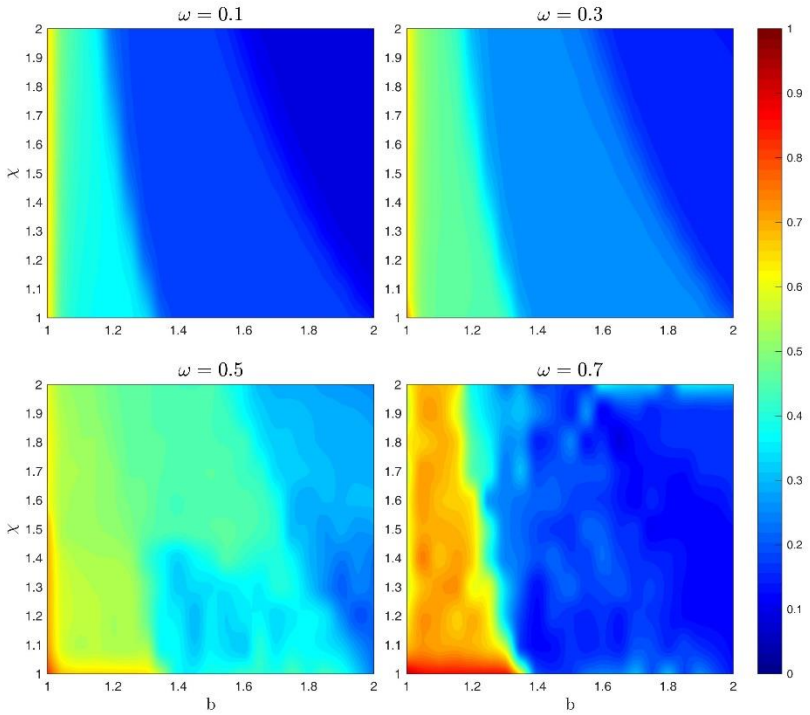

Fig. 10 Color-coded average frequency of cooperators on the $\chi - b$ parameter plane for different fusion weights ω of 0.1, 0.3, 0.5, and 0.7. The new rule based on evidential reasoning can obtain an intermediate transition state, and improve cooperation in a large area apace if the fusion weight is tuned well.

To further evaluate the effect of the new strategy learning rules on the promotion of cooperation, Fig. 11 characterize the Social Efficiency Deficit (SED) [65-67] of the full system for different parameters sets of extortion χ and the temptation b under different ω. The SED is defined as the difference between an actor's expected utility under the circumstance of the social optimum and

the evolutionary equilibrium. The results show that as ω increases from 0.1 to 0.5, the gap between social optimum payoff and the payoff attained at equilibrium gradually decreases in the whole $\chi - b$ panel. Especially, when $\omega = 0.5$, the SED is almost less than 1.0 regardless of the values of χ and b, which further illustrates that the effectiveness of the combination of the two single rules on the improvement of the average payoffs of the whole system. It should be noted that the SED remains at relatively small levels even in the region of high b when $\omega = 0.7$.

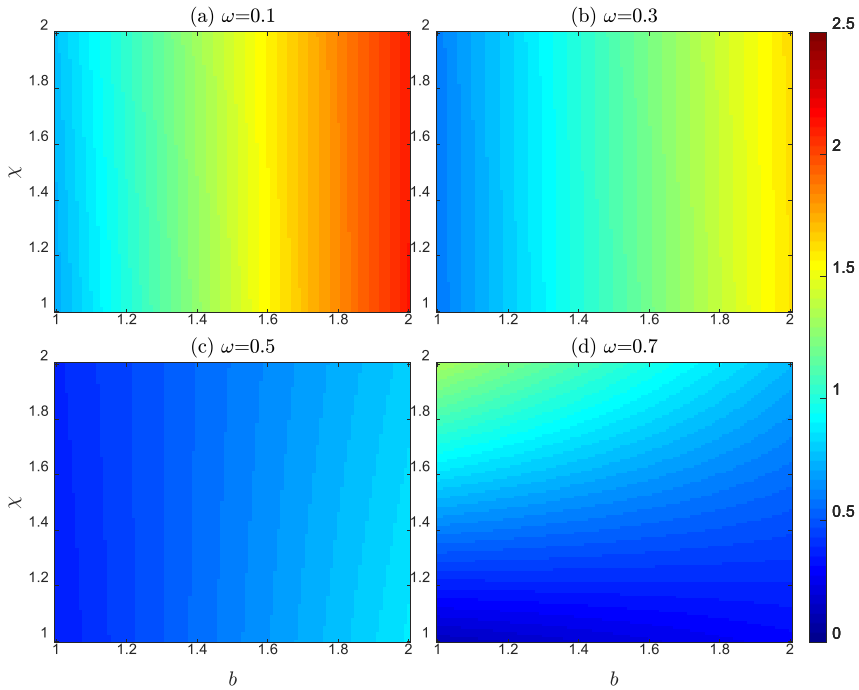

Fig. 11 Color-coded social efficiency deficit on the $\chi - b$ parameter plane for different fusion weights ω of 0.1, 0.3, 0.5, and 0.7.

Furthermore, considering that errors (or noise) are unavoidable in human interactions, the effect of error for extortion strategy is concerned under the proposed learning rule. Our analysis here focuses primarily on conditional strategy's action errors, that is, extortioners may execute actions incorrectly with a small probability. For instance, an extortion player gains zero benefit when repeatedly interacting with a defection player. However, due to the 'trembling hand', the extortion implements the other action with probability ε; thus the expected payoff he encounters with defectors becomes $-\varepsilon c$. Similarly, the expected payoffs of individuals with other strategies will also be influenced by extortioners' action errors. The payoff matrix considering action errors can be generalized as

$$M_2 = \begin{pmatrix} & \mathrm{D} & C & E_\chi \\ D & 0 & b & \varepsilon b \\ C & -c & b-c & (1-\varepsilon)\dfrac{(b^2-c^2)}{b\chi+c} - \varepsilon c + \varepsilon b/2 \\ E_\chi & -\varepsilon c & (1-\varepsilon)\dfrac{(b^2-c^2)\chi}{b\chi+c} + \varepsilon b - \varepsilon c/2 & 0 \end{pmatrix}, \tag{11}$$

Fig. 12 shows the frequencies of the three strategies when the probability ε of extortioners' action error is set reasonably to 0.01. The extortion strategy shows its robustness against the action error

when the value of b is not too large, under which cooperators can still dominate the whole population since the frequency of extortioners remains at a considerable level, which supports some level of cooperation. However, as b eventually increases near 2.0, compared to the results present in Fig. 4(c) where action error is absent, under the interference of the error, extortioners almost lose their chances of emerging among defectors, and its frequency tends to decrease, which leads to a sharp decline in cooperators. The intervention of the action error weakens the robustness of the extortion strategy as the dilemma strength increases.

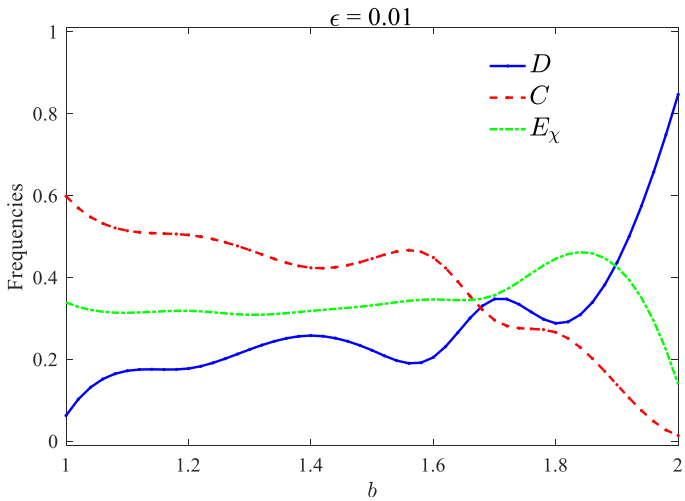

Fig. 12 With $\chi = 1.5$ and $\omega = 0.5$, average frequencies of defector (D), cooperator (C), and extortioner (E_χ) vary as b increases when the probability of action error for the extortioner is set as 0.01.

As the application of extortion strategy requires continuous inspection, we consider that the pure benefits of extortioners are reduced by cognition cost θ. We further explore how the evolution results will be when cognition cost is taken into account. Fig. 13 shows the evolutionary outcomes when θ is set as 0.1. Compared to the results depicted in Fig. 4(c) where monitoring cost is absent, the frequencies of the three strategies seem to be strongly influenced by the cognition cost. Especially, defectors tend to dominant the whole population under a wide range of tempting values b. The only source of payoffs for extortioners is to extort from cooperators, yet their amount can be directly burned by cognition cost. In this way, defectors can invade extortioners easily since it will not pay to stick to the current extortion. As a consequence, the shelter effect provided by extortioners for cooperators is getting weaker and weaker as the deterioration of conditions (for large b). The results indirectly highlight the role of extortioners in the evolution of cooperation, and once they become weak, cooperators may turn to be infeasible as well.

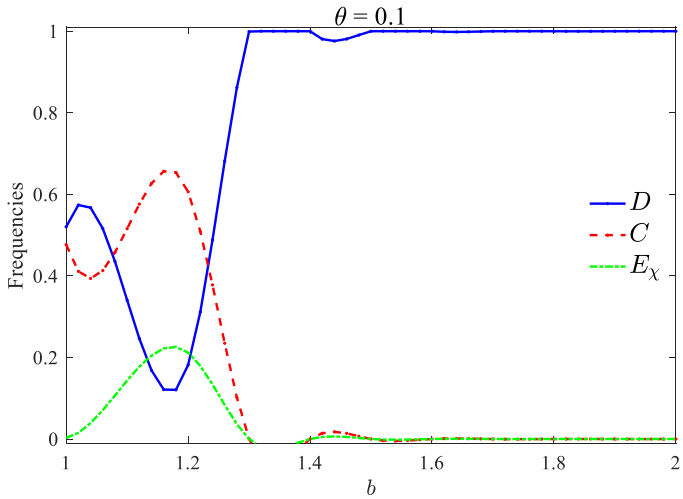

Fig. 13 With $\chi = 1.5$ and $\omega = 0.5$, average frequencies of defector (D), cooperator (C), and extortioner (E_χ) vary as b increases when the cognition cost θ of the extortioners is set as 0.1.

Last, we extend the study to four competing strategies, where the tit for tat (TFT) strategy is also taken into account. For the four competing strategies, the game parametrization is adopted the same as Ref. [14, 68]. The values in the matrix are the averaged payoffs given by infinite repeated games between the strategies, which is legitimate provided that the strategy adoption is rare in comparison with the frequency of games [69]. The expected payoff for any two role players is given by the following matrix

$$M_3 = \begin{pmatrix} & D & C & E_\chi & TFT \\ D & 0 & b & 0 & 0 \\ C & -c & b-c & \dfrac{b^2-c^2}{b\chi+c} & b-c \\ E_\chi & 0 & \dfrac{(b^2-c^2)\chi}{b\chi+c} & 0 & 0 \\ TFT & 0 & b-c & 0 & (b-c)/2 \end{pmatrix}. \tag{12}$$

To examine the effect of the new strategy learning rule on the promotion of cooperation, we set ω as 0.5 for the presented result in Fig. 14, which shows that cooperative strategies dominant the population regardless of the strength of the social dilemma. Compared to the results present in Fig. 4(c), the inclusion of the TFT enhances the level of cooperation and realizes better evolutionary outcomes for the system. As b increases, the frequency of defectors tends to increase slowly and then remains almost stable at a relatively low level which is less than the frequency of cooperators. We deduce that the TFT strategy weakens defection player's exploitation among the population since defectors cannot gain benefits from the TFT players. Moreover, the TFT players can support each other with cooperators. Thus, TFT players can spread among defectors and it is difficult for defectors to expand their region. For extortioners, they can still do favors for cooperators to survive, especially when the strength of the social dilemma is not too large. As b increases, the frequency of extortioners tends to be the lowest level among other competing strategies, so their sheltering effect on cooperators is no longer prominent. Generally speaking, the sheltered role played by extortioners and the support role played by the TFT players jointly favor the level of cooperation. In the region of low b, extortioners have not lost their roles in favoring cooperators, but TFT is

more profitable than extortion. However, in the region of large b, extortioners can timely switch the strategy to TFT, and the TFT player's role is more prominent in restricting the expansion of defectors. Moreover, cooperators have the chance to invade the TFT players. Notably, the new evidential reasoning-based combination rule highlights the effect of successful strategy and makes them promote each other in the emergence of cooperation.

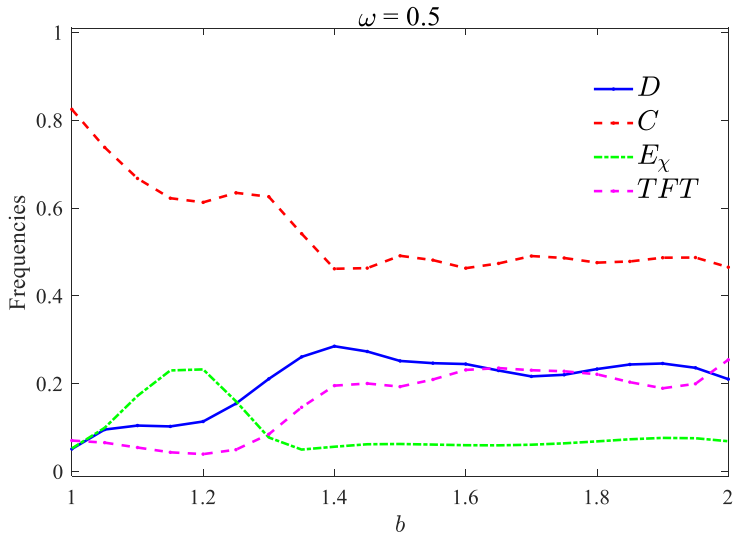

Fig. 14 With $\chi = 1.5$ and $\omega = 0.5$, average frequencies of defector (D), cooperator (C), extortioner (E_χ), and tit-for-tat players (TFT) vary as b increases.

4. Conclusion

In this paper, by introducing the different attitudes of a player towards a stronger or weaker opponent, we extend the emotion-imitating rule into the three-strategy spatial PDG model with extortion. Further, a new strategy learning rule has been constructed via evidential reasoning by integrating the emotion-imitating profile and original myopic best response learning profile. The fusion weight tunes the role of two sources to provide comprehensive information for strategy updating. Simulation results reveal that the new rule can significantly improve cooperation when the fusion weight is close to 0.5. With the shelter provided by extortioners, cooperators can survive and persist even at extremely high temptation to defect. Imitating emotion, but not the strategies, enables cooperators to avoid direct exploitation of defectors. Thus, cooperators can prevail in the population while defectors are suppressed. Besides, the myopic best response rule can also enhance the performance of the emotion-imitating rule, especially in the region of high temptation to defect. The performance of the new rule based on evidential reasoning can be superior to any single rule for the evolution of cooperation in a large area of parameter space if the fusion weight is tuned well. Moreover, the extortion player's catalyst role for the evolution of cooperation has also been verified when action errors are present. Furthermore, we find that the cognition cost of extortion players, namely, the extra burden from monitoring other strategies weakens the chances of extortioners emerging among defectors. Without the help of extortioners, cooperators become infeasible under a wide range of parameter values, which indirectly highlights the role of extortioners in the evolution of cooperation. Our model has also been extended to the situation that four strategies including TFT are competing in the system, and the results show that extortioners can act as catalysts effectively in the region of low tempt values; however, with the increase of the dilemma strength, the TFT

player's role is more prominent in promoting cooperation.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (No.71871173, 72031009, 71871171), Chinese National Funding of Social Sciences (No.20&ZD058), and the Fundamental Research Funds for the Central Universities (WUT: 2021III036JC). The final author is grateful for the support by NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Informatization (U1709215).

References

[1] M.A. Nowak, Evolutionary Dynamics: Exploring the Equations of Life, Harvard University Press, Cambridge, MA, 2006.

[2] D.D.P. Johnson, P. Stopka, S. Knights, Sociology: The puzzle of human cooperation, Nature, 421 (2003) 911-912.

[3] E. Pennisi, How did cooperative behavior evolve, Science, 309 (2005) 93.

[4] D.G. Rand, M.A. Nowak, Human cooperation, Trends Cogn. Sci., 17 (2013) 413-425.

[5] S.J. Maynard, G.R. Price, The logic of animal conflict, Nature, 246 (1973) 15-18.

[6] R. Axelrod, W.D. Hamilton, The evolution of cooperation, Science, 211 (1981) 1390.

[7] M. Perc, J.J. Jordan, D.G. Rand, Z. Wang, S. Boccaletti, A. Szolnoki, Statistical physics of human cooperation, Phys. Rep., 687 (2017) 1-51.

[8] W.H. Press, F.J. Dyson, Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent, Proc. Natl. Acad. Sci. USA, 109 (2012) 10409-10413.

[9] A.J. Stewart, J.B. Plotkin, From extortion to generosity, evolution in the Iterated Prisoner's Dilemma, Proc. Natl. Acad. Sci. USA, 110 (2013) 15348-15353.

[10] H. Ding, G.-S. Zhang, S.-H. Wang, J. Li, Z. Wang, Q-learning boosts the evolution of cooperation in structured population by involving extortion, Physica A, 536 (2019) 122551.

[11] A. Szolnoki, M. Perc, Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations, Sci. Rep., 4 (2014) 05496.

[12] J. Wang, J. Guo, A synergy of punishment and extortion in cooperation dilemmas driven by the leader, Chaos Soliton Fract., 119 (2019) 263-268.

[13] Z.J. Wang, Y.R. Zhou, J.M.W. Lien, J. Zheng, B. Xu, Extortion can outperform generosity in the iterated prisoner's dilemma, Nat. Commun., 7 (2016) 11125.

[14] C. Hilbe, M.A. Nowak, K. Sigmund, Evolution of extortion in Iterated Prisoner's Dilemma games, Proc. Natl. Acad. Sci. USA, 110 (2013) 6913-6918.

[15] A. Szolnoki, M. Perc, Evolution of extortion in structured populations, Phys. Rev. E, 89 (2014) 022804.

[16] G. Szabo, G. Fath, Evolutionary games on graphs, Phys. Rep., 446 (2007) 97-216.

[17] X. Chen, F. Fu, L. Wang, Promoting cooperation by local contribution under stochastic win-stay-lose-shift mechanism, Physica A, 387 (2008) 5609-5615.

[18] X. Deng, Z. Zhang, Y. Deng, Q. Liu, S. Chang, Self-adaptive win-stay-lose-shift reference selection mechanism promotes cooperation on a square lattice, Appl. Math. Comput., 284 (2016) 322-331.

[19] Y. Liu, X. Chen, L. Zhang, L. Wang, M. Perc, Win-stay-lose-learn promotes cooperation in the

spatial prisoner's dilemma game, PLoS ONE, 7 (2012) e30689.
[20] M. Perc, Phase transitions in models of human cooperation, Phys. Lett. A, 380 (2016) 2803-2808.
[21] A. Szolnoki, M. Perc, G. Szabó, Topology-independent impact of noise on cooperation in spatial public goods games, Phys. Rev. E, 80 (2009) 056109.
[22] H.-X. Yang, Z. Rong, P.-M. Lu, Y.-Z. Zeng, Effects of aspiration on public cooperation in structured populations, Physica A, 391 (2012) 4043-4049.
[23] A. Matsui, Best response dynamics and socially stable strategies, Journal of Economic Theory, 57 (1992) 343-362.
[24] M.A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, Emergence of cooperation and evolutionary stability in finite populations, Nature, 428 (2004) 646-650.
[25] J. Quan, W. Liu, Y.Q. Chu, X.J. Wang, Stochastic evolutionary voluntary public goods game with punishment in a Quasi-birth-and-death process, Sci. Rep., 7 (2017) 16110.
[26] J. Quan, W. Liu, Y.Q. Chu, X.J. Wang, Stochastic dynamics and stable equilibrium of evolutionary optional public goods game in finite populations, Physica A, 502 (2018) 123-134.
[27] J. Quan, X.J. Wang, Evolutionary games in a generalized Moran process with arbitrary selection strength and mutation, Chinese Phys. B, 20 (2011) 030203.
[28] L. Messerschmidt, A.P. Engelbrecht, Learning to play games using a PSO-based competitive learning approach, IEEE Evolut. Comput., 8 (2004) 280-288.
[29] J. Quan, X.K. Yang, X.J. Wang, Continuous spatial public goods game with self and peer punishment based on particle swarm optimization, Phys. Lett. A, 382 (2018) 1721-1730.
[30] J. Quan, X.K. Yang, X.J. Wang, Spatial public goods game with continuous contributions based on Particle Swarm Optimization learning and the evolution of cooperation, Physica A, 505 (2018) 973-983.
[31] L.M. Pereira, T. Lenaerts, L.A. Martinez-Vaquero, T.A. Han, Social Manifestation of Guilt Leads to Stable Cooperation in Multi-Agent Systems, in: Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems, International Foundation for Autonomous Agents and Multiagent Systems, 2017, pp. 1422-1430.
[32] C.M. de Melo, K. Terada, F.C. Santos, Emotion expressions shape human social norms and reputations, Iscience, 24 (2021) 102141.
[33] A. Szolnoki, N.-G. Xie, C. Wang, M. Perc, Imitating emotions instead of strategies in spatial games elevates social welfare, EPL, 96 (2011) 38002.
[34] N.-G.X. Attila Szolnoki, Ye Ye and Matjaˇz Perc, Evolution of emotions on networks leads to the evolution of cooperation in social dilemmas, Phys. Rev. E, 87 (2013) 042805.
[35] Y. Wang, T. Chen, Q. Chen, G. Si, Emotional decisions in structured populations for the evolution of public cooperation, Physica A, 468 (2017) 475-481.
[36] L.M. Pereira, F.C. Santos, Counterfactual thinking in cooperation dynamics, in: International conference on Model-Based Reasoning, Springer, 2018, pp. 69-82.
[37] M. Devaine, G. Hollard, J. Daunizeau, Theory of mind: Did evolution fool us?, PLoS ONE, 9 (2014) e87619.
[38] T.A. Han, F.C. Santos, T. Lenaerts, L.M. Pereira, Synergy between intention recognition and commitments in cooperation dilemmas, Sci. Rep., 5 (2015) 9312.
[39] L. McNally, S.P. Brown, A.L. Jackson, Cooperation and the evolution of intelligence, P. Roy. Soc. B-Biol. Sci., 279 (2012) 3027-3034.
[40] H.T. Anh, L. Moniz Pereira, F.C. Santos, Intention recognition promotes the emergence of

cooperation, Adapt. Behav., 3 (2011) 264-279.
[41] A.P. Dempster, A Generalization of Bayesian Inference, J. R. Stat. Soc., 30 (1968) 205-247.
[42] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.
[43] G. Shu, X. Du, Y. Li, Surrounding information consideration promotes cooperation in Prisoner's dilemma game, Chaos Soliton Fract., 91 (2016) 689-694.
[44] Y. Li, X. Lan, X.Y. Deng, R. Sadiq, Y. Deng, Comprehensive consideration of strategy updating promotes cooperation in the prisoner's dilemma game, Physica A, 403 (2014) 284-292.
[45] Y.F. Zhang, G. Shu, Y. Li, Strategy-updating depending on local environment enhances cooperation in prisoner's dilemma game, Appl. Math. Comput., 301 (2017) 224-232.
[46] X.Y. Deng, X. Zheng, X.Y. Su, F.T.S. Chan, Y. Hu, R. Sadiq, Y. Deng, An evidential game theory framework in multi-criteria decision making process, Appl. Math. Comput., 244 (2014) 783-793.
[47] X.Y. Deng, Q. Liu, Y. Deng, Matrix games with payoffs of belief structures, Appl. Math. Comput., 273 (2016) 868-879.
[48] X.Y. Deng, D.Q. Han, J. Dezert, Y. Deng, Y. Shyr, Evidence Combination From an Evolutionary Game Theory Perspective, IEEE Trans. Cybern., 46 (2016) 2070-2082.
[49] J. Quan, Y. Zhou, X. Wang, J.-B. Yang, Information fusion based on reputation and payoff promotes cooperation in spatial public goods game, Appl. Math. Comput., 368 (2020) 124805.
[50] J. Quan, Y. Zhou, X. Wang, J.-B. Yang, Evidential reasoning based on imitation and aspiration information in strategy learning promotes cooperation in optional spatial public goods game, Chaos Soliton Fract, 133 (2020) 109634.
[51] M. Zhou, X.-B. Liu, J.-B. Yang, Y.-W. Chen, J. Wu, Evidential reasoning approach with multiple kinds of attributes and entropy-based weight assignment, Knowl.-Based Syst., 163 (2019) 358-375.
[52] M. Zhou, X.-B. Liu, J.-B. Yang, Evidential reasoning-based nonlinear programming model for MCDA under fuzzy weights and utilities, Int. J. Intell. Syst., 25 (2010) 31-58.
[53] C. Fu, J.-B. Yang, S.-L. Yang, A group evidential reasoning approach based on expert reliability, Eur. J. Oper. Res., 246 (2015) 886-893.
[54] C. Fu, W. Chang, D. Xu, S. Yang, An evidential reasoning approach based on criterion reliability and solution reliability, Comput. Ind. Eng., 128 (2019) 401-417.
[55] J.-B. Yang, D.-L. Xu, Evidential reasoning rule for evidence combination, Artif. Intell., 205 (2013) 1-29.
[56] N.A.H. Haldar, J. Li, M. Reynolds, T. Sellis, J.X. Yu, Location prediction in large-scale social networks: an in-depth benchmarking study, VLDB J., 28 (2019) 623-648.
[57] T. Cai, J. Li, A.S. Mian, R. Li, T. Sellis, J.X. Yu, Target-aware Holistic Influence Maximization in Spatial Social Networks, IEEE Trans. Knowl. Data. Eng., (2020).
[58] Z.-X. Wu, Z. Rong, Boosting cooperation by involving extortion in spatial prisoner's dilemma games, Phys. Rev. E, 90 (2014) 062102.
[59] J. Tanimoto, H. Sagara, Relationship between dilemma occurrence and the existence of a weakly dominant strategy in a two-player symmetric game, BioSystems, 90 (2007) 105-114.
[60] Z. Wang, S. Kokubo, M. Jusup, J. Tanimoto, Dilemma strength as a framework for advancing evolutionary game theory: Reply to comments on “Universal scaling for the dilemma strength in evolutionary games”, Phys. Life. Rev., 14 (2015) 56-58.
[61] H. Ito, J. Tanimoto, Scaling the phase-planes of social dilemma strengths shows game-class changes in the five rules governing the evolution of cooperation, Roy. Soc. Open. Sci., 5 (2018) 181085.
[62] M.W. Macy, A. Flache, Learning dynamics in social dilemmas, Proc. Natl. Acad. Sci. USA, 99 (2002)

7229-7236.
[63] L.A. Zadeh, Review of Shafer's a mathematical theory of evidence, AI Mag., 5 (1984) 81-83.
[64] M. Perc, Stability of subsystem solutions in agent-based models, Eur. J. Phys., 39 (2018) 014001.
[65] M.R. Arefin, K.M.A. Kabir, M. Jusup, H. Ito, J. Tanimoto, Social efficiency deficit deciphers social dilemmas, Sci. Rep., 10 (2020) 16092.
[66] M.R. Arefin, T. Masaki, K. Kabir, J. Tanimoto, Interplay between cost and effectiveness in influenza vaccine uptake: A vaccination game approach, P. Roy. Soc. A, 475 (2019) 1-19.
[67] J. Tanimoto, Sociophysics Approach to Epidemics, Springer Singapore, 2021.
[68] A. Szolnoki, M. Perc, Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations, Sci. Rep., 4 (2014) 5496.
[69] G. Szabó, T. Antal, P. Szabó, M. Droz, Spatial evolutionary prisoner's dilemma game with three strategies and external constraints, Phys. Rev. E, 62 (2000) 1095-1103.