

# Enhancing social cohesion with cooperative bots in societies of greedy, mobile individuals

Lei Shi <sup>a,b,1</sup>, Zhixue He <sup>a,d,1</sup>, Chen Shen <sup>c,\*</sup> and Jun Tanimoto <sup>c,d</sup>

<sup>a</sup>School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming 650221, China

<sup>b</sup>Interdisciplinary Research Institute of data science, Shanghai Lixin University of Accounting and Finance, Shanghai 201209, China

<sup>c</sup>Faculty of Engineering Sciences, Kyushu University, Fukuoka 816-8580, Japan

<sup>d</sup>Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Fukuoka 816-8580, Japan

\*To whom correspondence should be addressed: Email: [steven\\_shen91@hotmail.com](mailto:steven_shen91@hotmail.com)

<sup>1</sup>L.S. and Z.H. contributed equally to this work.

Edited By: David Rand

## Abstract

Addressing collective issues in social development requires a high level of social cohesion, characterized by cooperation and close social connections. However, social cohesion is challenged by selfish, greedy individuals. With the advancement of artificial intelligence (AI), the dynamics of human–machine hybrid interactions introduce new complexities in fostering social cohesion. This study explores the impact of simple bots on social cohesion from the perspective of human–machine hybrid populations within network. By investigating collective self-organizing movement during migration, results indicate that cooperative bots can promote cooperation, facilitate individual aggregation, and thereby enhance social cohesion. The random exploration movement of bots can break the frozen state of greedy population, help to separate defectors in cooperative clusters, and promote the establishment of cooperative clusters. However, the presence of defective bots can weaken social cohesion, underscoring the importance of carefully designing bot behavior. Our research reveals the potential of bots in guiding social self-organization and provides insights for enhancing social cohesion in the era of human–machine interaction within social networks.

**Keywords:** social cohesion, migration, prisoner’s dilemma, human–machine hybrid population, self-organization movement

## Significance Statement

Migration behavior involves individuals transitioning to different environments by changing their position. In today’s digital age, online social platforms and collaborative networks play crucial roles in daily interactions and work. This expands the scope of migration behavior beyond physical relocation to include movement within online social networks. With the advancement of AI technology, bot agents are becoming integral to online environments, expected to address social issues. Understanding their influence on individuals is crucial. Our research reveals that cooperative bots have the potential to disrupt the stagnant state of mobile populations, initiating self-organizing movements and fostering social cohesion. This contributes to a deeper understanding of the impact of bots on the collective behavior within networks.

## Introduction

Social cohesion represents a crucial collective consciousness in contemporary societies facing collective issues such as epidemics (1), economic crises (2), social inequality (3), and climate change (4). However, the establishment of a highly cohesive social system is often hindered by individual selfish behaviors (5, 6). So far, some studies employing evolutionary game theory (7, 8) have revealed the generation and maintenance of social cohesion through self-organizing processes in interpersonal interactions within mobile population (6, 9–13), aiming to explore ways to enhance social cohesion (14). With the advent and integration of artificial intelligence (AI) technology in social settings, AI-driven agents,

or bots, are becoming a part of social fabric (15–17), shifting traditional human-to-human interactions to a new paradigm of human–machine hybrid interactions (18–24). The impact of this shift in interaction on social cohesion and collective behavior remains unclear. This study aims to deepen the understanding of the influence of bots on social cohesion by investigating self-organized movements within human–machine hybrid populations, and to explore potential avenues for enhancing social cohesion within such hybrid population contexts.

Social cohesion is fundamentally composed of two elements: *orientation towards the common goods* and *nurturing of social relationships* (6). The *orientation towards the common goods* involves

**Competing Interest:** The authors declare no competing interest.

**Received:** March 11, 2024. **Accepted:** May 24, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).

cooperative behavior that individuals prioritizing collective benefits over personal gains (7, 25). It is often challenged by selfish free-riding behavior, leading to the “tragedy of the commons” (26). Over the last few decades, various reciprocity mechanisms that support evolution of cooperation have been revealed. These include direct reciprocity, which arises from repeated interactions (27), indirect reciprocity that relies on the behavioral information transmission (28), and network reciprocity, which is influenced by the structure of interactions (29).

Migration is a fundamental behavior among individuals, providing a perspective for exploring the maintenance and the formation of *social relationships*. Traditionally, migration refers to physical movement, such as residential relocation. However, in the information technology era, research now includes digital network migration behavior (30–32). In online networks like *Reddit* and *GitHub*, users can freely switch between working groups and communities, altering their interactional relationships as desired. These migration promotes the formation of self-organized movements (33, 34). Different social interaction, including segregation (35), interweaving (36), and aggregation (14, 37), arise in self-organizing movements driven by various migration preferences. In particular, individuals driven by payoffs tend to form close social bonds through movement (14). This spatial clustering during migration can enhance cooperation, but depends on specific conditions, such as moderate population density or low mobility (33, 37). Conversely, excessive greed for personal interests impedes the development of cooperation, disrupts the establishment of social connections, and ultimately erodes social cohesion (14).

Recent research has sparked interest in using AI-driven agents or bots to study cooperation issues (38, 39). They have revealed bots’ ability to address coordination dilemmas (21) and scaffold cooperation (23) by integrating bots into network engineering and game interactions (22, 24). Here, our focus extends beyond the influence of bots on individual cooperation to their role in the collective self-organization movements within populations. To achieve this, we introduce bots into a mobile populations comprised of selfish, greedy individuals whose behavioral decisions aim to maximize personal gains. Our model does not assume bots possess complete knowledge of individual behavior or engage in coordinated actions towards normal individuals (21, 22). Instead, we enable bots to participate autonomously in migration, using a simple behavioral design characterized by consistent adoption of cooperative action and mobile exploration with some randomness. As we will see, these cooperative bots can facilitate cooperation and fostering spatial clustering within mobile populations, thereby promoting highly social cohesion. Interestingly, cooperative bots can break the population out of its frozen state, stimulating the self-organization movement among selfish, greedy individuals. The emergence and maintenance of social cohesion in sparse mobile populations have often been a challenge in human-human interactions, but the introduction of cooperative bots can solve this challenge situation in such hybrid populations. Therefore, our study reveal the potential of simple cooperative bots in guiding individual behavior to address collective issues.

## Model

### Hybrid population

We investigate a hybrid population consisting of both bots and greedy normal players, with proportions  $\phi \in [0, 0.5]$  and  $1 - \phi$  respectively. This population is placed on a grid lattice network of

size  $N = L \times L$  with periodic boundary and  $K$ -nearest neighboring sites (specifically focusing on the von Neumann neighborhood where  $K = 4$ ). Each site in the network can be occupied by either a bot/normal player, or it may remain unoccupied, so we define the population density as  $\rho = n/N$  ( $0 < \rho < 1$ ). Consequently, the network contains  $N\rho\phi$  bots and  $N\rho(1 - \phi)$  normal players.

Our model is conducted using Monte Carlo (MC) asynchronous simulations. In each MC time step, both of bots and normal players undergo three stages: game interaction, strategic updating, and migration. The model architecture is depicted in Fig. 1.

### Game interaction

Game interactions are implemented through a one-shot PD game. Both of bots and normal players participate in the paired game with their neighbors, making decisions to adopt either unconditional cooperation strategy (C) or unconditional defection strategy (D). Cooperation means incurring a cost  $c$  to benefit others by an amount  $b$ , while defection does nothing. Mutual cooperation yields a reward of  $R = b - c$ , while mutual defection leads to a punishment of  $P = 0$  for both agents. A cooperator receives a sucker’s payoff of  $S = -c$ , whereas a defector gains a temptation to defect payoff of  $T = b$  upon meeting each other. To simplify the model without loss generality, we define the dilemma strength as  $r = c/(b - c)$  and set  $b - c = 1$  following the method outlined in ref. (40). The payoff matrix is then re-scaled as:

$$\begin{array}{c} C \\ D \end{array} \begin{array}{cc} C & D \\ \left( \begin{array}{cc} 1 & -r \\ 1+r & 0 \end{array} \right) \end{array} \quad (1)$$

We focus on the influence of cooperative bots which are programmed to consistently choose unconditional cooperation without altering their behavior in interaction. We also explore defective bots which consistently choose unconditional defection, and the corresponding outcomes are presented in the “Supplementary Material”.

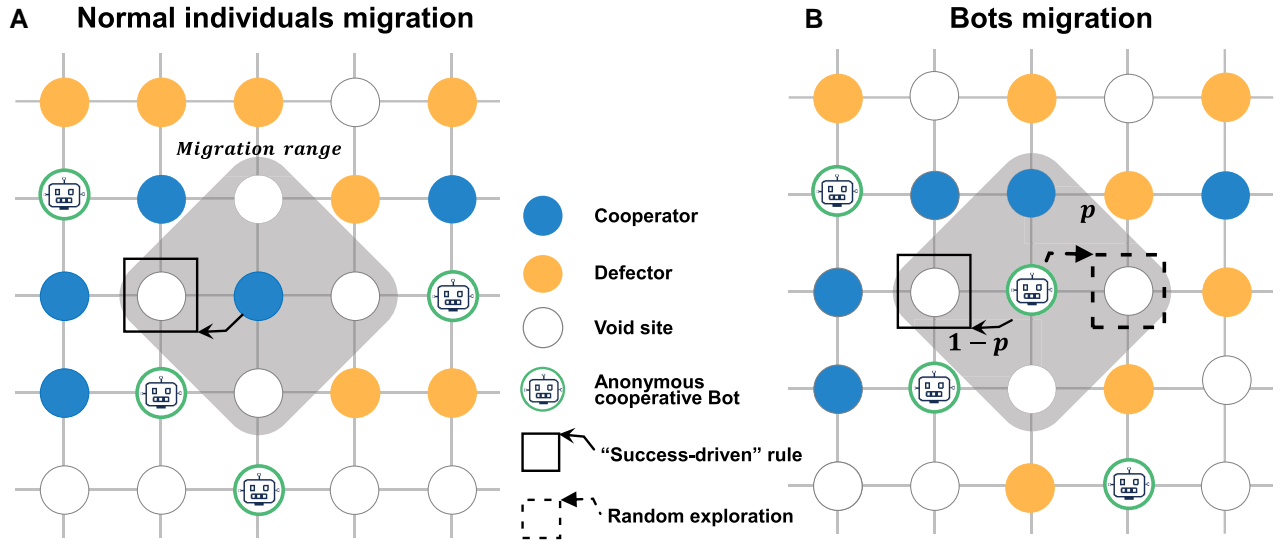
### Strategic updating

We utilize an anonymous setup where normal players remain unaware of the presence of bots, thereby exclude potential bias against the bots among them (41–43). The decision-making of greedy normal players is driven by the maximization of their own profits. They employ the “best-take-over” rule (25), imitating the strategy of their neighbors that yields the highest payoff when their neighbors’ payoff exceeds their own.

### Migration

Normal players strategically also migrate to maximize their payoffs by adhering to the “success-driven” rule (34, 37). They move to adjacent vacant sites (including their current position) that offer the highest payoffs in fictitious play. To prevent bots from being consistently exploited by their counterparts and to enhance adaptation to the migration environment, bots are programmed to engage in exploratory migration. Bots take random movements with a probability of  $p \in [0, 1]$  and follow the “success-driven” rule with a probability of  $1 - p$ .

To analyze the impact of bots on social cohesion, we utilize the fraction of cooperation among normal players ( $F_c$ ) as a metric to evaluate how bots promote individuals to consider collective interests. On the other hand, we assess the overall spatial aggregation of normal players, denoted as *Agg*, to gain insights into the influence of bots on social cohesion from a spatial perspective.



**Fig. 1.** Schematic diagram of model setting. A hybrid population of size  $n$ , consisting of proportional  $\phi$  mobile bots and  $1 - \phi$  normal individuals deployed on a lattice network of size  $N$  where some nodes on the network are empty ( $n < N$ ) and available for agents to migrate. Panel A: normal individuals mimic the strategy of most successful neighbors (including themselves) and migrate based on the “success-driven” rule, moving to adjacent unoccupied location (within the shaded area) that generate the highest payoffs in fictitious play. Panel B: anonymous bots are programmed to engage in unconditional cooperation, and migrate randomly with a probability of  $p$ , while follow “success-driven” rule with a probability of  $1 - p$ .

$Agg$  is the weighted average of the aggregation levels of normal cooperators ( $Agg_C$ ) and defectors ( $Agg_D$ ), expressed as:

$$Agg_i = \frac{L_{i-C} + L_{i-D}}{K} \quad i \in \{C, D\}, \quad (2)$$

$$Agg = F_C \times Agg_C + (1 - F_C) \times Agg_D, \quad (3)$$

where  $L_{x-C}$  ( $L_{x-D}$ ) is the average number of neighboring cooperators (defectors) for normal players who adopt strategy  $x$ .

To investigate how bots influence the self-organization dynamics during migration, inspired by the cluster shape analysis detailed in ref. (44), we introduce  $\lambda_C \in [-1, 1]$  and  $\lambda_D \in [-1, 1]$  as indices to assess the clustering level of normal cooperators and the separation level of normal defectors relative to the cooperative clusters, respectively:

$$\lambda_C = \frac{1}{|\Omega_C|} \sum_{i \in \Omega_C} \frac{m_C^i - m_D^i}{K}, \quad (4)$$

$$\lambda_D = \frac{1}{|\Omega_D|} \sum_{i \in \Omega_D} \frac{m_{void}^i - m_C^i}{K}, \quad (5)$$

where  $\Omega_C$  and  $\Omega_D$  denote the sets of normal cooperators and normal defectors, respectively, with their respective quantities denoted by  $|\Omega_C|$  and  $|\Omega_D|$ .  $m_C^i$ ,  $m_D^i$ , and  $m_{void}^i$  are the numbers of cooperators, defectors, and void sites within the neighborhood of player  $i$ , respectively. A  $\lambda_C$  approaching 1 suggests cooperators can form tightly clusters. Conversely, a greater number of links between cooperators and defectors result in  $\lambda_C < 0$ , indicating a negative assortment among cooperators (44). For  $\lambda_D > 0$ , defectors are adjacent to cooperative clusters, and as  $\lambda_D$  approaches 1, it indicates that the degree of separation between defectors and cooperative clusters increases. Conversely, a negative  $\lambda_D$  implies defectors are embedded within the cooperative clusters.

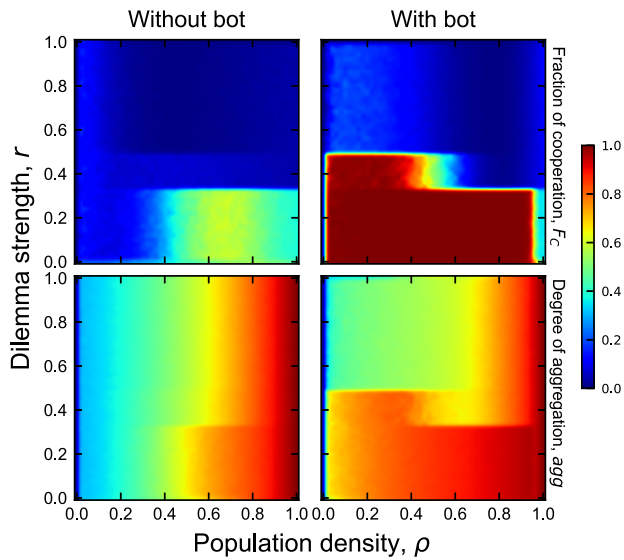
For our computer simulation, we maintained a fixed grid size of  $N = 100 \times 100$ . To ensure the reliability and stability of our results,

we averaged the final outcomes over 50 independent runs. Each run involved averaging the last 5,000 time steps out of more than  $10^6$  Monte Carlo (MC) time steps. To confirm the robustness of our model and the obtained results, we extensively explored various scenarios, detailed in [Supplementary Material](#). This exploration encompassed different proportions of bots, varied migration and interaction strategies for bots, limited mobility of normal players, and the effects of lattice network size and its topological structure. We investigate the influence of bots across different levels of complete information acquisition, taking into account the strategic decisions of normal players. Furthermore, we examine the scenarios where the decision-making processes of normal individuals involve behavioral noise within the [Supplementary Material](#), aiming to validate the robustness of the results obtained by relaxing the assumption of absolute rationality among normal individuals presented in the main text.

## Results

### Simple bot promote social cohesion

We begin by examining the influence of cooperative bots. Previous studies have shown that migration can facilitate the clustering of cooperators in populations of moderate density, thereby enhancing network reciprocity when the dilemma strength is low (i.e.  $r < 1/3$ ) (14, 45), as depicted in the left panels of Fig. 2. However, in sparse populations (i.e.  $\rho < 0.4$ ), the abundance of empty sites hampers the formation of cooperative clusters, resulting in the decline of cooperation. Under high levels of dilemma strength (i.e.  $r > 1/3$ ), cooperation cannot be sustained even with available empty nodes and individual migration. Interestingly, the introduction of cooperative bots significantly promote cooperation among normal individuals, see the right panels of Fig. 2. Compared to scenarios without bots, the presence of bots greatly enhances cooperation across a wider range of population density parameters when  $r < 1/3$ , ensuring high levels of cooperation in both sparse and dense populations. Even under high dilemma



**Fig. 2.** Simple cooperative bots facilitate cooperation and population aggregation, promoting the establishment of highly cohesive collective behavior. The color code indicates the fraction of cooperation (top panels) and the degree of aggregation of normal individuals (bottom panels) as a function of population density  $\rho$  and temptation  $r$ . Parameters are set to  $\phi = 0.5$  and  $p = 0.01$  for the scenario with bots.

strength, bots prove effective in maintaining cooperation. Furthermore, only a few bots are needed to have a significant impact on individual behavior, see Fig. S3 in the [Supplementary Material](#). At a moderate population density of  $\rho = 0.6$ , introducing a minority of mobile bots is sufficient to sustain a high level of cooperation (i.e. approximately  $\phi \approx 0.02$  for random mobile bots and  $\phi \approx 0.11$  for low-exploration bots with  $p = 0.01$ ). Even under high dilemma strength, as low as  $\phi \approx 0.2$  proportion of low-exploration bots can maintain cooperation.

In sparse populations, the abundance of empty nodes separating individuals leads to high isolation and low aggregation. As population density increases, a corresponding increase in the degree of aggregation as expected. Bots also can promote population clustering. Even at low population densities (i.e.  $\rho < 0.4$ ), bots can induce normal individuals to aggregate, resulting in a high degree of aggregation, as shown in the bottom panels of Fig. 2. Similarly, a minority of bots can significantly enhance the degree of aggregation, as depicted in Fig. S4 in the [Supplementary Material](#). These findings demonstrate that cooperative bots not only facilitate cooperation but also aggregate individuals, thereby promoting the emergence of high levels of social cohesion.

## Self-organizing movement

To understand the influence of mobile bots on social cohesion, Fig. 3 depicts the temporal evolution and spatial distribution of the population. The dynamic visualization of this process is available online at <https://osf.io/tu6cq>. In the absence of bots, a typical evolutionary process unfolds (shown in the top panels of Fig. 3). Initially, the random spatial distribution impedes the survival of isolated cooperators, leading to a decline in the fraction of cooperation ( $F_c$ ) in the early stages of evolution. When migration is feasible, cooperative migration further drives the aggregation into clusters, compared to the cluster formation process without migration as discussed in (44). While empty nodes may partition some defectors and cooperators, they also constrain further expansion of cooperators. When no more profitable positions exist,

normal individuals cease movement (with the fraction of normal individuals who moved  $F_m$  remaining at zero), leading to a frozen state in self-organizing movement of greedy population (37).

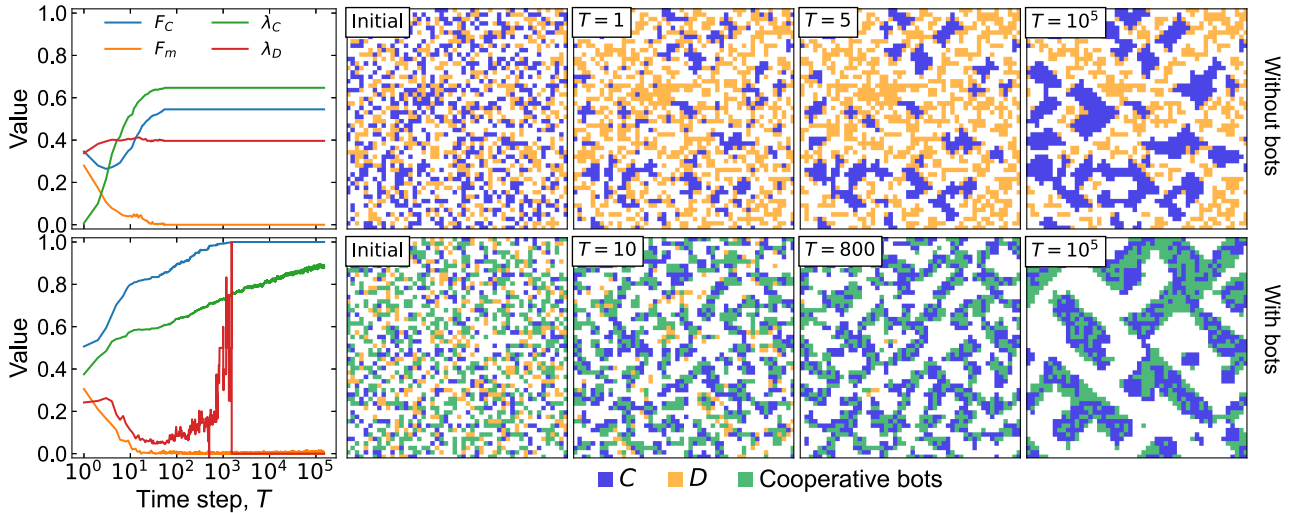
Interestingly, the introduction of cooperative bots disrupts this frozen state (although the value of  $F_m$  is low, it is not zero), as depicted in the bottom panels of Fig. 3. It indicates that the presence of bots indirectly creates profitable positions, driving normal individuals to move, and thus break the frozen state. More importantly, cooperative bots can facilitate tight cooperator cluster formation (evidenced by continuous  $\lambda_c$  increase) while driving defector separation from cooperative clusters (as shown by rising  $\lambda_D$ ). This self-organizing movement, catalyzed by cooperative bots, renders defectors defeated by tightly knit cooperative clusters. Even after defectors vanish, bots can further facilitate the aggregation of cooperator (The final value of  $\lambda_c$  stabilizes at a high level). In the presence of behavioral noise, individuals randomly reset their strategies and migration with a certain probability. These stochastic behaviors can prevent the system from reaching a complete freeze (37). Intriguingly, cooperative bots also can foster the levels of cooperation and aggregation among the population, compare to the scenarios in the absence of bots, as illustrated in Figs. S8 and S9 in the [Supplementary Material](#).

However, in extremely dense populations (i.e.  $\rho = 0.97$ ), cooperative bots fail to eliminate defectors as they cannot efficiently drive defector separation from cooperative clusters, see top panels of Fig. S2 in [Supplementary Material](#). On the other hand, in extremely sparse populations (i.e.  $\rho < 0.02$ ), cooperative clusters cannot be established. Under a high dilemma strength (i.e.  $r = 0.4$ ), the self-organization movement promoted by the bot cannot eliminate defection due to the payoffs advantage of the defector, see bottom panels of Fig. S2 in [Supplementary Material](#). It is worth noting that the normal individuals' strategy updating and "success-driven" migration depends on access to complete information about others' behaviors. When avenues for acquiring such comprehensive information are restricted, individuals rely solely on their own judgment for strategy updating (employing the myopic principle (25), wherein normal individuals tend to adopt a better response strategy in the current situation) and resort to random migration. Results show that cooperative bots still can to promote cooperation under conditions of limited information, as depicted in Fig. S7 of [Supplementary Material](#). Nevertheless, cooperative bots no longer effectively aggregate normal individuals, as illustrated in Fig S5 of [Supplementary Material](#).

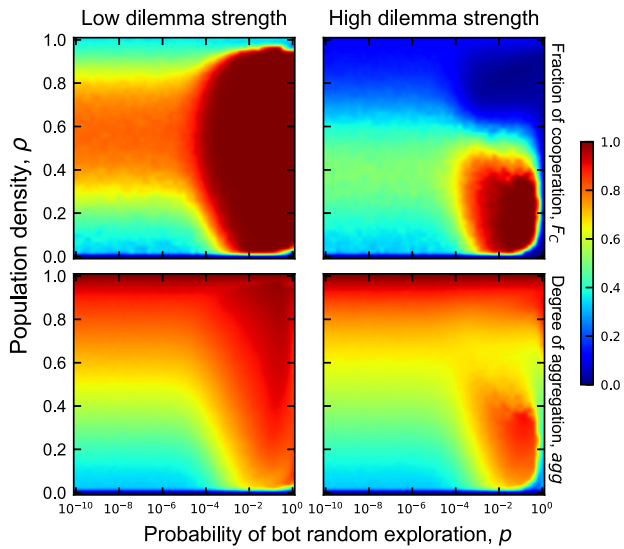
## The effects of bot behavior

The mobility of bots is instrumental in driving the self-organizing dynamics of individuals. In Fig. 4 and 5, we examine the impact of bots on social cohesion across varying migration. Our findings indicate that random exploration behavior enhances bots' ability to promote cohesion within the population. Bots that remain static or lack random exploration (i.e.  $p = 0$ ) are unable to disrupt the frozen state (with  $F_m$  remaining at 0), and fails to facilitate the segregation of defectors from cooperative clusters (with  $\lambda_D$  remains around 0.45), thus demonstrating limited effectiveness in promoting cooperation, as depicted by the two leftmost columns in Fig. 5. Conversely, when bots can randomly migrate exploration, even with a low level of exploration (i.e.  $p = 0.01$ ), they stimulate self-organizing movement and maintain a high level of aggregation among normal individuals, as shown in Fig. 4. Notably, bots with random migration (i.e.  $p = 1$ ) can further elevate the migration level of normal individuals, resulting in a higher value of  $F_m$





**Fig. 3.** Cooperative bots can drive self-organized movement of normal individuals, preventing the population from entering a frozen state that would typically occur in their absence. Leftmost panels show the fraction of cooperation  $F_C$ , the fraction of normal individuals who have moved  $F_m$ , clustering level of normal cooperators  $\lambda_C$  and the isolation level of normal defectors  $\lambda_D$  as a function of time steps  $T$ . The right panels showcase evolutionary snapshots in scenarios with and without bot, respectively. Results are obtained by setting  $\rho = 0.5$ ,  $r = 0.2$ , and  $p = 0.01$ .



**Fig. 4.** Bots with a moderate level of random exploration can contribute to the promotion of social cohesion. The color code indicates the fraction of cooperation (top panels) and the degree of aggregation of normal individuals (bottom panels) as a function of population density  $\rho$  and the probability of bot random exploration  $p$  under low temptation  $r = 0.2$  and high temptation  $r = 0.4$ . The fraction of bot is set to  $\phi = 0.5$ .

compared to  $p = 0.01$ . This can drive the establishment of an exceptionally large cooperative cluster, as depicted in the rightmost column of Fig. 5. However, under a high dilemma strength, bots with high-level exploration (i.e.  $p > 0.7$ ) fail to promote cooperation. In contrast, bots with low-level exploration can still increase cooperation, as demonstrated in Fig. 4 and Fig. S1 of [Supplementary Material](#).

Interactive action of bots is another critical factor influencing formation of social cohesion. When a population includes defective bots that consistently opt for unconditional defection, despite their potential to facilitate individual migration (as illustrated in Fig. S6 of [Supplementary Material](#) where  $F_m$  is non-zero), it does

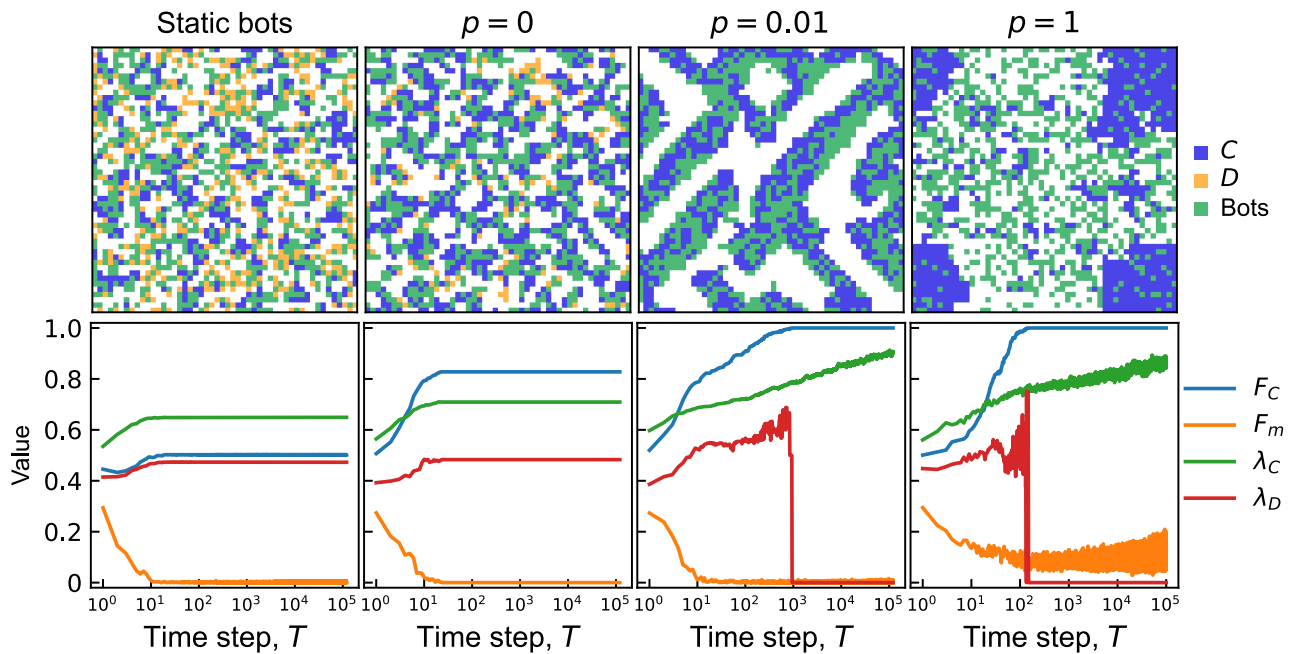
not lead to the separation of defectors from cooperative clusters. Moreover, the presence of defective bots diminishes the ability of cooperative bots to promote social cohesion. Enhanced social cohesion only occurs when the proportion of cooperative bots within the bot subgroup exceeds a certain threshold, while both defective and cooperative bots coexist, as depicted in Fig. S5 of [Supplementary Material](#).

### Robustness of model

To evaluate the impact of action sequence in our model, we varied the sequence such that all individuals migrate before engaging in interactions and updating strategies (33). Results illustrate that altering the action sequence can enhance social cohesion when cooperative bots are involved, as shown in Fig. S10 of [Supplementary Material](#). However, under a high dilemma strength, bots without migration exploration outperform those with such exploration. When the action is limited, where individuals can only migrate or update strategies within a Monte Carlo time step, a minority of cooperative bots can still foster social cohesion, as depicted in Fig. S11 of [Supplementary Material](#). Moreover, we explore the influence of different lattice network topologies, where individuals have broader interaction and migration ranges, as well as larger population sizes. Results from Figs. S12 and S13 in the [Supplementary Material](#) demonstrate the robustness of cooperative bots in promoting social cohesion across diverse lattice network topologies and population sizes.

## Conclusions and discussions

This work employs evolutionary game theory to analyze how bot affect collective behavior of mobile population within networks. Results shown that cooperative bots can enhance the social cohesion among selfish, greedy individuals. When these individuals cannot find favorable migration positions, the lack of migration motivations leads to the emergence of frozen state in population (37). Interestingly, introduction of cooperative bots can break this frozen state. The random exploratory movements of bots create favorable positions, facilitating the clustering of cooperators



**Fig. 5.** The random exploration of mobile bots drives the separation of defectors from cooperative clusters and facilitates the aggregation of cooperators. Top panels show stable spatial distribution in scenarios with static bots (i.e. unable to migrate), success-driven bots (i.e.  $p = 0$ ), low-exploration bots (i.e.  $p = 0.01$ ) and randomly migrating bots (i.e.  $p = 1$ ). Bottom panels show that the fraction of cooperation  $F_C$ , the fraction of normal individuals who have moved  $F_m$ , clustering level of normal cooperators  $\lambda_C$  and the isolation level of normal defectors  $\lambda_D$  as a function of time steps  $T$ , respectively. All stable spatial distribution are obtained at  $T = 10^5$ . These outcomes were obtained with parameter settings of  $\rho = 0.6$ ,  $r = 0.2$ , and  $\phi = 0.5$ .

and isolating defectors from cooperative clusters, thus leading to the defeat of defectors by tightly knit cooperative clusters. Even a minority of cooperative bots can promote social cohesion within the population, see Fig. S3 of Supplementary Material. This suggests that cooperative bots act as internal forces social self-organization. The effect of cooperative bot is not limited by specific migration-imitation sequences or individual limited mobility, nor does it rely on the population size and the topology of normal network (refer to Fig. S10, Fig. S11 and Fig. S13 in the Supplementary Material).

However, in extremely sparse populations, cooperative bots cannot help the establishment of cooperative clusters. In extremely densely populations, they fail to facilitate the separation of defectors from cooperative clusters during the self-organization process, thus diminishing their ability to eliminate defection. In our model, normal individual decisions rely on complete information regarding neighbors' behavior. When individuals lack information and resort to random movement, they rely on self-judgment to update strategies, cooperative bots still contribute to maintaining cooperation but cannot drive the aggregation of individuals. These conditions weaken the capacity of cooperative bots to promote social cohesion. Furthermore, our findings indicate that the presence of defective bots also hampers the establishment of social cohesion and diminishes the efficiency of cooperative bots in facilitating cooperation. These suggest the need for careful consideration in the design of bot behaviors.

Our research is conducted within the context of one-shot games, where individuals make decisions without access to information regarding the past behavior of their co-players. While cooperative bots share some similarities with human zealots (46–48)—both consistently opting for cooperation—there are critical distinctions. Human zealots are rare in realistic settings, making it impractical to rely on them for widespread application if a high propensity for cooperation requires a substantial number of

human zealots. In contrast, the behavior and scale of digital bots are controllable, making them effective tools for influencing human beliefs and behaviors in various online aspects, such as elections (49), voting (50), and political issues (51).

Our findings hold broad implications for online social platform, particularly concerning trust and opinion conflicts. For example, users often share opinions and collaborate with others online to accomplish tasks. However, misinformation and hostile communication environments frequently escalate conflicts, leading users to sever social connections and neglect collective interests. Using cooperative bots—designed to maintain friendly communication and provide collaborative assistance—can help create a more congenial communication environment and propagate collective consciousness (52). Particularly, recent advancements in large language models exhibit impressive communication prowess and the potential to shape individuals' beliefs (15). This enables the construction of these cooperative agents to facilitate connectivity and communication among users, potentially enhancing cooperation and trust within online communities. Furthermore, our results highlight the critical role of incorporating random exploratory migration into bot design—allowing bots to roam different online communities—can help bridge connections among disconnected users and isolated communities to shape collective cohesion.

We employ a two-dimensional grid network, the simple network structure, though not a fully reflection of real-life social networks, encapsulates crucial social network features: limited interactions among individuals and engagement with neighbors. We anticipate that our findings remain robust, as they stem from the involvement of bots in individual limited interactions, which is independent of specific topological structures. Real-world network structures may display heterogeneity and time-varying (53), future investigations into these characteristics will enhance understanding of bot impacts.

In real-world, besides one-shot interactions, repeated interactions are also common. Bots with simple behaviors may not suffice to guide collective actions in this scenario. Instead, bots might be susceptible to manipulation and exploitation by humans, resulting in inefficiencies (24). Further consideration of memory-based strategy design may help explore the impact of bots on individuals in repeated games (24, 54, 55). A key assumption in our study is that humans are unaware of interacting with bots, whether in game interactions or migration processes. When individuals become aware that their counterparts are bots, issues of trust in human-machine interactions (42) and biases towards bots (41) emerge, which are critical factors affecting bot efficiency. Unfortunately, the impact of these factors remains unclear. Furthermore, we only focused on selfish and greedy individuals, as this aids in our exploration of whether bots alone can foster cooperation. However, human behavior is motivated by various factors beyond the pursuit of self-interest maximization. It is also influenced by social norms (56) and various value-orientations (57). Future endeavors may benefit from integrating diverse behavioral decisions to comprehensively understand the impact of bots on collective behavior. Addressing these challenges will provide deeper insights to harness bots as effective tools in solving complex social issues.

## Acknowledgments

We appreciate the valuable suggestions provided by two anonymous reviewers.

## Supplementary Material

Supplementary material is available at PNAS Nexus online.

## Funding

We acknowledge the support provided by (i) the National Natural Science Foundation of China (Grant No. 11931015), Major Program of National Fund of Philosophy and Social Science of China (Grants Nos. 22&ZD158 and 22VRCO49) to L.S.; (ii) China Scholarship Council (Grant No. 202308530309) and Yunnan Provincial Department of Education Science Research Fund Project (Grant No. 2024Y503) to Z.H.; (iii) JSPS Postdoctoral Fellowship Program for Foreign Researchers (Grant No. P21374), and JSPS KAKENHI (Grant No. JP 23H03499) to C.S.; and (iv) the Grant-In-Aid for Scientific Research from JSPS, Japan, KAKENHI (Grants Nos. JP 20H02314, JP 23K28189, and JP 23H03499) awarded to J.T.

## Author Contributions

L.S., Z.H., and C.S. designed research; L.S., Z.H., and C.S. performed research; Z.H., C.S., and J.T. analyzed results; L.S. and J.T. supervision; L.S., Z.H., C.S., and J.T. wrote the paper.

## Preprints

A preprint of this article is published at <https://doi.org/10.48550/arXiv.2403.00311>.

## Data Availability

There are no data underlying this work. The results are obtained by computer simulation, the code is available at OSF: <https://osf.io/tu6cq>.

## References

- 1 Ash T, Bento AM, Kaffine D, Rao A, Bento AI. 2022. Disease-economy trade-offs under alternative epidemic control strategies. *Nat Commun.* 13(1):3319.
- 2 Acharya VV, Richardson M. 2009. Causes of the financial crisis. *Crit Rev.* 21(2-3):195–210.
- 3 Hauser OP, Hilbe C, Chatterjee K, Nowak MA. 2019. Social dilemmas among unequals. *Nature.* 572(7770):524–527.
- 4 Steffen W, et al. 2015. Planetary boundaries: guiding human development on a changing planet. *Science.* 347(6223):1259855.
- 5 Tainter J. 1988. *The collapse of complex societies.* Cambridge: Cambridge University Press.
- 6 Schiefer D, Van der Noll J. 2017. The essentials of social cohesion: a literature review. *Soc Indic Res.* 132:579–603.
- 7 Sachs JL, Mueller UG, Wilcox TP, Bull JJ. 2004. The evolution of cooperation. *Q Rev Biol.* 79(2):135–160.
- 8 Nowak MA. 2006. *Evolutionary dynamics: exploring the equations of life.* Cambridge (MA): Harvard University Press.
- 9 Helbing D, Yu W, Rauhut H. 2011. Self-organization and emergence in social systems: modeling the coevolution of social environments and cooperative behavior. *J Math Sociol.* 35(1-3): 177–208.
- 10 Castellano C, Fortunato S, Loreto V. 2009. Statistical physics of social dynamics. *Rev Mod Phys.* 81(2):591.
- 11 Fu F, Nowak MA. 2013. Global migration can lead to stronger spatial selection than local migration. *J Stat Phys.* 151:637–653.
- 12 Wu T, Fu F, Zhang Y, Wang L. 2012. Expectation-driven migration promotes cooperation by group interactions. *Phys Rev E.* 85(6): 066104.
- 13 Wu T, Fu F, Wang L. 2011. Moving away from nasty encounters enhances cooperation in ecological prisoner's dilemma game. *PLoS ONE.* 6(11):e27669.
- 14 Roca CP, Helbing D. 2011. Emergence of social cohesion in a model society of greedy, mobile individuals. *Proc Natl Acad Sci USA.* 108(28):11370–11374.
- 15 Abdullah M, Madain A, Jararweh Y. 2022. Chatgpt: fundamentals, applications and social impacts. In: Ninth International Conference on Social Networks Analysis Management and Security. IEEE. p. 1–8.
- 16 Santos FP. 2024. Prosocial dynamics in multiagent systems. *AI Magazine.* 45(1):131–138. doi:<https://doi.org/10.1002/aaai.12143>
- 17 Chen X, Fu F. 2023. Ensuring the greater good in hybrid Ai-human systems: comment on “reputation and reciprocity” by Xia et al. *Phys Life Rev.* 48:41–43.
- 18 Crandall JW, et al. 2018. Cooperating with machines. *Nat Commun.* 9(1):233.
- 19 Shirado H, Christakis NA. 2020. Interdisciplinary case study: understanding the cooperation of humans and robots through the collaboration of social and computer scientists. *iScience.* 23(12):101680.
- 20 Santos FP, Pacheco JM, Paiva A, Santos FC. 2019. Evolution of collective fairness in hybrid populations of humans and agents. In: Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. AAAI. p. 6146–6153.
- 21 Shirado H, Christakis NA. 2017. Locally noisy autonomous agents improve global human coordination in network experiments. *Nature.* 545(7654):370–374.
- 22 Shirado H, Christakis NA. 2020. Network engineering using autonomous agents increases cooperation in human groups. *iScience.* 23(9):101438.

- 23 McKee KR, et al. 2023. Scaffolding cooperation in human groups with deep reinforcement learning. *Nat Hum Behav.* 7(10):1787–1796.
- 24 Hilbe C, Röhl T, Milinski M. 2014. Extortion subdues human players but is finally punished in the prisoner's dilemma. *Nat Commun.* 5(1):3976.
- 25 Nowak M, Highfield R. 2011. *Supercooperators: altruism, evolution, and why we need each other to succeed.* New York: Simon and Schuster.
- 26 Hardin G. 1968. The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality. *Science.* 162(3859):1243–1248.
- 27 Schmid L, Chatterjee K, Hilbe C, Nowak MA. 2021. A unified framework of direct and indirect reciprocity. *Nat Hum Behav.* 5(10):1292–1302.
- 28 Rockenbach B, Milinski M. 2006. The efficient interaction of indirect reciprocity and costly punishment. *Nature.* 444(7120):718–723.
- 29 Nowak MA. 2006. Five rules for the evolution of cooperation. *Science.* 314(5805):1560–1563.
- 30 Kumar S, Zafarani R, Liu H. 2011. Understanding user migration patterns in social media. In: *Proceedings of the AAAI Conference on Artificial Intelligence.* Vol. 25. AAAI. p. 1204–1209.
- 31 Newell E, et al. 2016. User migration in online social networks: a case study on reddit during a period of community unrest. In: *Proceedings of the International AAAI Conference on Web and Social Media.* Vol. 10. AAAI. p. 279–288.
- 32 Fiesler C, Dym B. 2020. Moving across lands: online platform migration in fandom communities. *Proc ACM on Hum-Comput Interact.* 4(CSCW1):1–25.
- 33 Vainstein MH, Silva ATC, Arenzon JJ. 2007. Does mobility decrease cooperation? *J Theor Biol.* 244(4):722–728.
- 34 Helbing D, Yu W. 2008. Migration as a mechanism to promote cooperation. *Adv Complex Syst.* 11(04):641–652.
- 35 Hamnett C. 2001. Social segregation and social polarization. In: *Handbook of urban studies.* London: Sage. p. 162–176.
- 36 He Z, Geng Y, Du C, Shi L, Wang Z. 2022. Q-learning-based migration leading to spontaneous emergence of segregation. *New J Phys.* 24(12):123038.
- 37 Helbing D, Yu W. 2009. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proc Natl Acad Sci USA.* 106(10):3680–3685.
- 38 Guo H, et al. 2023. Facilitating cooperation in human-agent hybrid populations through autonomous agents. *iScience.* 26(11):108179.
- 39 Sharma G, Guo H, Shen C, Tanimoto J. 2023. Small bots, big impact: solving the conundrum of cooperation in optional prisoner's dilemma game through simple strategies. *J R Soc Interface.* 20(204):20230301.
- 40 Wang Z, Kokubo S, Jusup M, Tanimoto J. 2015. Universal scaling for the dilemma strength in evolutionary games. *Phys Life Rev.* 14:1–30.
- 41 Karpus J, Krüger A, Verba JT, Bahrami B, Deroy O. 2021. Algorithm exploitation: humans are keen to exploit benevolent AI. *iScience.* 24(6):102679.
- 42 Ishowo-Oloko F, et al. 2019. Behavioural evidence for a transparency–efficiency tradeoff in human–machine cooperation. *Nat Mach Intell.* 1(11):517–521.
- 43 Wang Z, et al. 2017. Onymity promotes cooperation in social dilemma experiments. *Sci Adv.* 3(3):e1601444.
- 44 Fu F, Nowak MA, Hauert C. 2010. Invasion and expansion of cooperators in lattice populations: prisoner's dilemma vs. snowdrift games. *J Theor Biol.* 266(3):358–366.
- 45 Meloni S, et al. 2009. Effects of mobility in a population of prisoner's dilemma players. *Phys Rev E.* 79(6):067101.
- 46 Masuda N. 2012. Evolution of cooperation driven by zealots. *Sci Rep.* 2(1):646.
- 47 Cardillo A, Masuda N. 2020. Critical mass effect in evolutionary games triggered by zealots. *Phys Rev Res.* 2(2):023305.
- 48 Shen C, et al. 2023. How committed individuals shape social dynamics: a survey on coordination games and social dilemma games. *Europhys Lett.* 144(1):11002.
- 49 Bessi A, Ferrara E. 2016. Social bots distort the 2016 US presidential election online discussion. *First Mon.* 21(11-7).
- 50 Stella M, Ferrara E, De Domenico M. 2018. Bots increase exposure to negative and inflammatory content in online social systems. *Proc Natl Acad Sci USA.* 115(49):12435–12440.
- 51 Bail CA, et al. 2018. Exposure to opposing views on social media can increase political polarization. *Proc Natl Acad Sci USA.* 115(37):9216–9221.
- 52 Traeger ML, Strohkorb Sebo S, Jung M, Scassellati B, Christakis NA. 2020. Vulnerable robots positively shape human conversational dynamics in a human–robot team. *Proc Natl Acad Sci USA.* 117(12):6370–6375.
- 53 Strogatz SH. 2001. Exploring complex networks. *Nature.* 410(6825):268–276.
- 54 He Z, Shen C, Shi L, Tanimoto J. 2024. Impact of committed minorities: unveiling critical mass of cooperation in the iterated prisoner's dilemma game. *Phys Rev Res.* 6(1):013062.
- 55 Chen X, Fu F. 2023. Outlearning extortioners: unbending strategies can foster reciprocal fairness and cooperation. *PNAS Nexus.* 2(6):pgad176.
- 56 Capraro V, Perc M. 2021. Mathematical foundations of moral preferences. *J R Soc Interface.* 18(175):20200880.
- 57 Schwarting W, Pierson A, Alonso-Mora J, Karaman S, Rus D. 2019. Social behavior for autonomous vehicles. *Proc Natl Acad Sci USA.* 116(50):24972–24978.