

A Networked N -player Trust Game and its Evolutionary Dynamics

Manuel Chica, Raymond Chiong, *Senior Member, IEEE*, Michael Kirley, Hisao Ishibuchi, *Fellow, IEEE*

Abstract—Trust and trustworthiness are of great importance in social and human systems, especially when considering managerial and economic decision-making. In this paper, we investigate the emergent dynamics of an evolutionary game-theoretic model – the N -player evolutionary trust game – consisting of three types of players: an investor, a trustee who is trustworthy, and a trustee who is untrustworthy. Here, we limit the interactions between players to local neighborhoods defined by a specific spatial topology or social network. Players are able to adjust their game-playing strategies using an evolutionary update rule based on the payoffs obtained by their neighbors. Through comprehensive simulation experiments, we find that trust can be promoted when players interact on a social network despite a substantial number of untrustworthy individuals in the initial population. These results differ from findings reported for an unstructured population of the same game, where the existence of a single untrustworthy individual would eliminate trust completely. We compare the dynamics of the model with different social network densities and structures (e.g., from regular lattices to scale-free and random networks). We observe that the levels of trust vary under different network structures, and the level is correlated with how ‘difficult’ the game is. When game conditions are easy (i.e., low temptation to defect and/or almost no initial untrustworthy trustees), homogeneous networks with higher densities can promote higher levels of trust. However, when the game becomes harder, heterogeneous social networks with lower densities are able to promote higher levels of trust and global net wealth.

Index Terms—Trust, Evolutionary Game Theory, Social Networks, Scale-free Networks, Agent-based Modeling.

I. INTRODUCTION

The importance and social role of trust is clear [1]. Trust has implications for social and human systems [2]–[5], including ethical considerations [6], cross-cultural dimensions in an organizational context [7], and shaping human relationships [8]. As individuals use trust to manage complexity, relationships emerge. This, in turn, increases complexity and thus creates more reinforcement and opportunities for trust to spread [9]. Researchers recently found that individuals are more trustworthy when having uncalculating cooperation, as this behavior is utilized to signal trustworthiness [10].

Evolutionary game theory (EGT) has already been used for modeling trust (e.g., see [9], [11], [12]). In comparison with

the vast amount of studies on cooperation models such as the prisoners dilemma (PD) [13]–[16] and snowdrift (SD) [17], [18] games, however, the number of related studies on trust models in the context of EGT is limited. Trust games are sequential in nature. The most well-known version of trust games involves interactions between an *investor* (or *truster*) and a *trustee* [19]–[21]. Fundamentally, it represents a two-player game where the investor begins with an initial stake of value and can decide whether to keep or transfer it to the trustee. Later, the trustee must choose how much to return to the investor (or not at all).

This well-known trust game model, unfortunately, does not generalize to multiple players. Bach et al. [22] illustrated the qualitative difference between two-player and multiplayer games and pointed out the limitations to the general conclusions when extending a two-player game to multiplayer ones. Gokhale et al. [23] showed that findings derived from pairwise interactions cannot be generalized to multiple players with more than two strategies. Consequently, Abbas et al. [9] proposed an evolutionary N -player trust game (also modeled as a sequential game), which generalizes the concept of trust to the case of a population of individuals playing the trust game concurrently in a well-mixed and unstructured environment. However, infinitely large, well-mixed populations and deterministic dynamics are idealizations, as real populations have a finite number of individuals and are not well mixed [24]. In addition, the lack of information or bounded rationality influences human decisions and how players choose their strategies during a game [25].

In our work, we propose a networked version of the N -player trust game where a structured population of players is considered. An agent-based modeling framework is employed to represent players as agents, and they have three possible strategies to choose from: to be an investor (I), a trustworthy trustee (T), or an untrustworthy trustee (U). For a finite population size, $k_i, i = \{I, T, U\}$, is the number of players choosing a possible strategy i . Multipliers R_T and R_U control the received fund by trustworthy and untrustworthy trustees, respectively, fulfilling the following inequality relation: $R_T < R_U < 2R_T$. Additionally, we introduce a ratio r_{UT} , which defines the temptation to defect investors’ trust by the trustees and measures the level of ‘difficulty’ of the game.

Game-playing agents in the population are mapped to a social network [26], which restricts their interactions to local neighborhoods. We consider different kinds of social networks, ranging from a simple regular lattice [27] to the high preferential attachment of scale-free (SF) networks [28], [29], as well as an Erdős-Rényi (ER) random network [30].

Manuel Chica is with The University of Newcastle, Callaghan, NSW 2308, AUSTRALIA and with RØD Brand Consultants, 28001 Madrid, SPAIN. e-mail: manuel.chicaserrano@newcastle.edu.au

Raymond Chiong is with The University of Newcastle, Callaghan, NSW 2308, AUSTRALIA e-mail: raymond.chiong@newcastle.edu.au

Michael Kirley is with The University of Melbourne, Parkville, VIC 3010, AUSTRALIA e-mail: mkirley@unimelb.edu.au

Hisao Ishibuchi is with Department of Computer Science and Engineering Southern University of Science and Technology, Shenzhen 518055, CHINA e-mail: hisao@sustc.edu.cn

We also investigate a fully-connected version of the game, thereby allowing us to directly compare our results with those of Abbass et al. [9], in which the evolutionary dynamics of the population would always converge to an untrustworthy state even when the initial population has only a single untrustworthy player. In our model, players are given an opportunity to imitate others' strategies through an evolutionary update rule. Concretely, we use the proportional imitation rule [31] – this update rule has been chosen because it is very similar to the replicator dynamics studied by Abbass et al. [9], making the comparison of our results with theirs more meaningful.

In our simulation experiments, we consider the equilibrium state of players (i.e., surviving strategies) and global net wealth (i.e., total payoffs) of the population. Different scenarios are examined to better understand how trustworthy and untrustworthy behaviors emerge and what is their corresponding global net wealth like. First, we explore the effects of different values of the trustworthy multiplier parameter R_T and the temptation to defect ratio r_{UT} . We then investigate the whole spectrum of initial conditions of the population by considering different numbers of players with strategies I , T , and U , aiming to discover under which initial conditions the presence of untrustworthy individuals in the population would lead to the extinction of investors I and to what extent they influence the global wealth. Our experiments also take into account different network configurations with varying densities, in order to understand how the density of a network and information diffusion speed impact the population dynamics.

The rest of this paper is structured as follows. In Section II, we provide background information and review related work in the areas of spatial evolutionary games and trust models. After that, we describe the details of our proposed N -player trust game model in Section III, including the network structures and update rule used. Experimental setups and results are then presented in Section IV. Finally, concluding remarks and possibilities for future work are given in Section V to wrap up the paper.

II. BACKGROUND AND RELATED WORK

A. Networked and Spatial Evolutionary Games

The spatial distribution of a population makes interactions among neighbors more likely than interactions between distant individuals [24]. Social networks in human populations, for example, enable friends to interact more often than strangers. These realizations led to spatial approaches for evolutionary game dynamics [17], [27], [32] and later evolutionary graph theory [33]. Since the pioneering work of Nowak and May [27], showing how the spatial structure is able to promote cooperation in the PD game, the use of spatial and networked evolutionary games has attracted much interest from many different disciplines [34]–[36].

In a spatial evolutionary game, players are typically mapped to a regular lattice with periodic boundary conditions. Each player is restricted to interactions within a local neighborhood. From one time step to another, players are allowed to change their strategies either deterministically or probabilistically in order to increase their utility. In [14], Chiong and Kirley

presented a multiplayer evolutionary game model in which agents play iterative games in spatial populations. They used N -player versions of the PD and SD games as the basis of their investigation.

There is also relevant work coming from the spatial public goods (PG) game, a game with PD types of interactions generalized to groups of arbitrary sizes. Santos et al. [37] introduced social diversity to the game by means of heterogeneous networks. In their model, individuals interact along the social ties defined by a heterogeneous SF network. It is found that diversity associated with the size of the PG game can promote strong cooperation in the absence of other mechanisms. Generally speaking, the heterogeneity in the number of connections of the individuals (i.e., degrees of the nodes in the network) leads to social diversity. This diversity has important implications for the evolution of cooperation [36].

Motivated by the work of Santos et al. [37], many studies based on SF networks have emerged as a consequence. Qualitatively similar findings were subsequently reported on the impact of diversity. For example, Yang et al. [38] investigated the effects of individual diversity on the emergence and evolution of cooperation using the PG game and proposed a strategy to maximize cooperative behavior. Zhang et al. [39] and Wei et al. [40] also used SF networks for evolutionary PG games. Zhang et al. [39] studied a mechanism with unequal payoff allocation. Wei et al. [40] showed that cooperation in SF networks has no-trivial dependence on individuals' heterogeneous behaviors.

B. Evolutionary Trust Games

The trust game has its root in economics [19]. It is also known as the investment game. In a conventional trust game, the investor has a strategy p_0 , which basically is the probability of the investor making a transfer to the trustee. The trustee has a strategy r , which represents the fraction of the transfer that the trustee would return to the investor. Behavioral experiments from classical economic theory suggest that investors making transfers with high probabilities would lead to trustees returning a substantial amount of the transfers [19]. Adding an evolutionary dimension to the game, researchers found that knowledge sharing (about trustees) and permitting a “partner choice” (for investors) can alter the nature of game interaction [11] and lead to the evolution of fully trusting and marginally trustworthy behavior.

To the best of our knowledge, the only existing N -player trust game can be found in the work of Abbass et al. [9]. In their proposed game model, each player makes two decisions in advance: (1) whether or not to be trustworthy, and (2) whether to be an investor or a trustee. An investor pays tv to the trustee, where tv (> 0) is the trusted value. With k_I investors, the total fund contributed is $k_I \cdot tv$. Each of the k_{TU} trustees receives the same amount of the fund: $k_I \cdot tv / k_{TU}$. A trustworthy trustee (strategy T) returns the received fund multiplied by R_T to the investors: $R_T \cdot k_I \cdot tv / k_{TU}$ (i.e., $R_T \cdot tv / k_{TU}$ to each of the k_I investors). At the same time, the same amount of fund as the returned is kept by each trustworthy trustee. Each investor receives the same return

from the k_T trustworthy trustees: $R_T \cdot tv/k_{TU}$ in total. Each of the untrustworthy trustees (strategy U) returns nothing but keeps for themselves the received fund multiplied by R_U : $R_U \cdot k_I \cdot tv/k_{TU}$. Table I shows the utility matrix of a player with each strategy¹.

TABLE I
PAYOFFS FOR EACH PLAYER IN THE N -PLAYER TRUST GAME.

	I players (investor)	T players (trustworthy)	U players (untrustworthy)
Pay	tv	$R_T \cdot tv \cdot \frac{k_I}{k_{TU}}$	0
Receive	$R_T \cdot tv \cdot \frac{k_T}{k_{TU}}$	$2 \cdot R_T \cdot tv \cdot \frac{k_I}{k_{TU}}$	$R_U \cdot tv \cdot \frac{k_I}{k_{TU}}$
Net wealth	$tv \cdot (R_T \cdot \frac{k_T}{k_{TU}} - 1)$	$R_T \cdot tv \cdot \frac{k_I}{k_{TU}}$	$R_U \cdot tv \cdot \frac{k_I}{k_{TU}}$

Here, the best net wealth from an individual point of view always goes to players who choose to be an untrustworthy trustee (strategy U). Abbass et al. [9] showed that in a well-mixed environment even if there is just a single untrustworthy player in the initial population, there would be a rapid spread of untrustworthiness throughout the entire population, leading to the extinction of investors. Nevertheless, they also found that a fraction of the population would always remain trustworthy, even in the absence of investors. Additionally, the predicted steady-state outcome is independent from the initial distribution of players when untrustworthy players are present in the population (i.e., $k_U > 0$). With no investors surviving, the final global net wealth is inevitably 0. On the other hand, the numbers of trustworthy and untrustworthy trustees in the final population (k_T and k_U) are dependent on the values of multipliers R_T and R_U . When the values of R_T and R_U are small, investors (i.e., players with strategy I) extinct quickly. This consequently limits the spread of untrustworthy players (i.e., players with strategy U).

However, Abbass et al. [9] did not consider the evolution of trust in a structured population. A recent study by Tarnita [12] explored the role of structured populations for the classical evolutionary trust game with non-random interactions. She showed that the population structure biases selection towards strategies that are both trusting and trustworthy. In the next section, we introduce the very first structured N -player trust model proposed by us.

III. THE NETWORKED TRUST MODEL

Our trust model consists of a finite set of agents occupying the nodes of a social network, and the edges denote interactions between them (both for accumulating payoffs and strategy updating [24]). Like Abbass et al.'s model, each agent in our model can choose from three possible strategies: a) to be an investor (strategy I), b) to be a trustworthy trustee (strategy T) and c) to be an untrustworthy trustee (strategy U). Figure 1 shows an example of a social network with 11 players, highlighting the local neighborhoods of focal agents i , j , and h . Each of the focal agents interacts with all of its

directly connected neighbors at the same time in a single game, i.e., group interactions. We chose this approach because games on networks governed by group interactions have gained much popularity, even though seminal work along this line of research typically focuses on pairwise interactions [36]. In the following sections, we will explain the game structure, payoffs, evolutionary update rule, and network topologies.

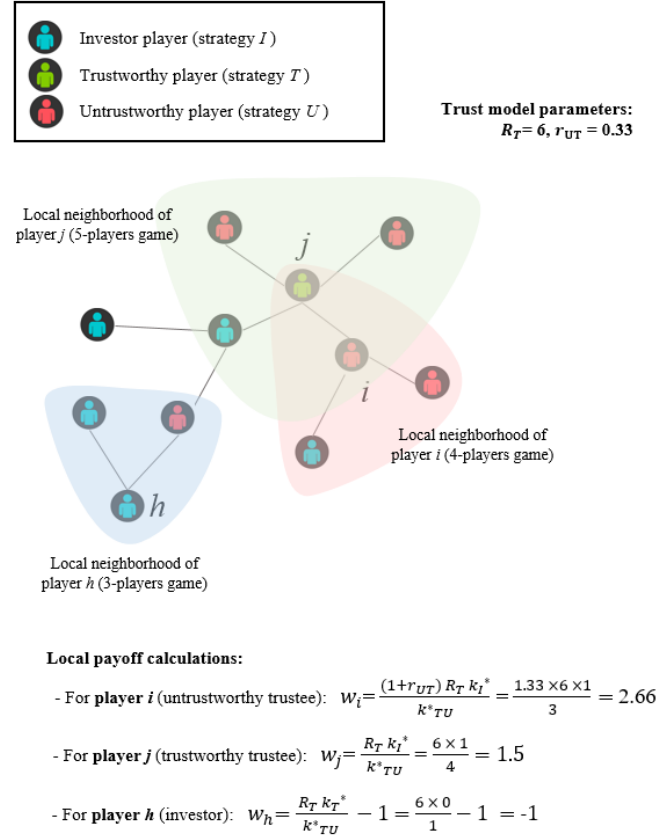


Fig. 1. An example of 11 players mapped to a social network. The three areas shaded in red, green and blue show the local neighborhoods of agents i , j and h , which form 4-, 5- and 3-player trust games, respectively. The figure also shows how the agents' payoffs are calculated according to their strategies and those of their neighbors.

A. Game Payoffs and Net Wealth

All agents in the population play the trust game over a fixed number of time steps with their direct neighbors, defined by the given network structure. The initial population of pop agents is generated at random with three types of players: k_I is the number of investors, k_T the number of trustworthy trustees, and k_U the number of untrustworthy trustees, where $k_I + k_T + k_U = pop$. We set the trusted value tv to 1 because the game dynamics are independent from its specific value [9]. The net wealth of individual agents, calculated based on their payoffs, is determined according to the strategy adopted by themselves and their direct neighbors. For example, the net wealth w_i of focal agent i can be calculated according to the following payoff function:

¹For clarity's sake, we use slightly different notations from that used by Abbass et al. [9]. For example, the number of players k_1, k_2 , and k_3 are substituted by k_I, k_T , and k_U , respectively. Trustee multipliers $R1$ and $R2$ are also substituted by R_T and R_U .

$$w_i = \begin{cases} \frac{R_T \cdot k_T^*}{k_{TU}^*} - 1, & i = 1, \dots, k_I \\ & \text{(investor),} \\ \frac{R_T \cdot k_I^*}{k_{TU}^*}, & i = (k_I + 1), \dots, (k_I + k_T) \\ & \text{(trustworthy),} \\ \frac{(1+r_{UT}) \cdot R_T \cdot k_I^*}{k_{TU}^*}, & i = (k_I + k_T + 1), \dots, pop \\ & \text{(untrustworthy),} \end{cases} \quad (1)$$

where k_I^* , k_T^* , and k_{TU}^* are the numbers of investors, trustworthy trustees, and both trustworthy and untrustworthy trustees in the local neighborhood of i (including focal agent i itself). See Figure 1 for an illustration of how the payoff of agent i can be calculated. The net wealth of agent i is simply 0 when there is no trustee in the neighborhood (i.e., $w_i = 0$ when $k_{TU}^* = 0$).

In our study, we consider $r_{UT} \in (0, 1)$ as the temptation to defect ratio (to be untrustworthy), defined by:

$$r_{UT} = \frac{R_U - R_T}{R_T}, \text{ fulfilling that } 1 < R_T < R_U < 2R_T. \quad (2)$$

We are interested in the global net wealth of the population W , calculated as $W = \sum_{i=1}^{pop} w_i$. We run our model for a maximum number of time steps in a synchronous manner. At each time step t , agents decide on which strategies to choose based on the population state of the previous time step (i.e., $t - 1$). This means that the actions of others at time step t will not affect the focal agent's decision during the same time step [41]. All agents in the population then get an opportunity to update their strategies synchronously according to the payoffs received at the previous time step. This is done by using an evolutionary update rule to be described in the ensuing section.

B. Evolutionary Update

The strategies of agents (i.e., I , T , and U) can change during the trust game, as each agent is given an opportunity to update its strategy through an evolutionary update process. We may perceive this activity of strategy update as information exchange in a social learning process, where agents in the population can imitate the strategies of others [24]. Strategy imitation occurs in all time steps during the game. At time step t , a focal agent (independently from its strategy) evaluates its previous payoff in $t - 1$ and decides whether to imitate a direct neighbor's strategy or not by applying an evolutionary update rule.

Evolutionary update rules of imitative nature represent a situation where bounded rationality or lack of information forces players to copy (imitate) others' strategies [31]. These rules are widely employed in the relevant literature to model evolutionary dynamics. For our model, we use one of the most common update rules: the proportional imitation rule [42]. The main reason for selecting this rule is because it brings, for large well-mixed populations, the evolution of replicator

dynamics [31]. This makes it convenient for us to compare our results to that of Abbass et al. [9].

The proportional imitation rule is a pairwise and stochastic update rule. An agent may adopt one of the three possible strategies for the game (i.e., I , T or U) from another agent within its neighborhood at time step t . For instance, let us assume that agent i in Figure 1 is of strategy U – it may either keep its strategy for the next time step or adopt the strategy of one of its three neighbors (selected at random). If the randomly selected neighbor of agent i is j , the rule will first evaluate if the individual net wealth of j in the previous time step $t - 1$, w_j^{t-1} , is higher than that of agent i (w_i^{t-1}). If it is higher, agent i will adopt the strategy of agent j by a probability that depends on the difference between their payoffs:

$$prob_{ij}^t = \frac{\max\{0, w_j^{t-1} - w_i^{t-1}\}}{\phi} \quad (3)$$

where ϕ is the difference of maximum and minimum possible individual net wealth between two arbitrary agents at time step $t - 1$ to have $prob_{ij}^t \in [0, 1]$. In the example given in Figure 1, $prob_{ij}^t$ will be 0 as agent i has a higher payoff than agent j .

The minimum possible net wealth for a player in the game ϕ is -1 . This -1 value is obtained when the focal agent is an investor (I) and all its neighbors are untrustworthy trustees (U), because there is not any trustworthy trustee to return the investment. On the contrary, the maximum possible net wealth for a focal player in the game is equal to $(1 + r_{UT}) \cdot R_T \cdot k_I^*$. This value is obtained when the focal agent is untrustworthy (U) and all its neighbors are investors (I).

C. Social Network Topologies

A social network is generally defined by a set of actors/agents and the relationships (ties) among them. Social networks affect the word-of-mouth and play a fundamental role in the way information is spread [43]. Several studies have demonstrated that different network topologies impact on information diffusion differently (e.g., see [14], [44]).

In our model, all agents in the population are placed on a SF network by default (as shown in Figure 1). The network is static, meaning that other agents that form a focal agent's group (i.e., the neighbors) do not change over the course of the game. We have chosen SF networks [28], [29] as the default network structure of our model because SF networks have been used extensively in studies related to evolutionary games (e.g., see [34], [37]). Synthetic SF networks with different densities can be generated by the well-known Barabasi-Albert model [29]. This algorithm starts with a small clique (a completely connected graph) of m_0 nodes. At each successive iteration, a new node is added and connected to m different existing nodes, where m is equal or smaller than m_0 . When a new node is connected to an existing node, the probability of choosing an existing node is proportional to the degree. After t iterations, this algorithm results in a network with $n = t + m_0$ nodes and $m \cdot t$ edges. In our simulation experiments, we analyze the model dynamics with varying SF network densities created by means of different values of m .

Apart from the SF network, we also consider another two more homogeneous social networks for our model: the regular lattice and a random graph generated by the ER algorithm [30] (i.e., the ER network). Both of them have varying features in terms of the average path length, centrality, or distribution of the number of connections of the nodes (i.e., nodes' degree $\langle k \rangle$), which affect the way information is spread and players interact.

In a regular lattice, players interact with each other in a totally homogeneous manner [27]. All players have exactly the same number of contacts or degree $\langle k \rangle$ (e.g., $\langle k \rangle$ equals to either 4 or 8). Regular lattices therefore tend to have long average paths and no clustering property. In our work, we use the standard regular lattice with $\langle k \rangle = 4$, i.e., the von Neumann neighborhood, with periodic boundary conditions.

A random ER graph model, on the other hand, assumes that there are n nodes, and these nodes connect with each other based on a probability p . Here, the probability that a vertex has k edges follows a Poisson distribution. ER random networks thus have short average paths and low clustering. This is in contrast to the power-law tail of the SF network, which indicates that highly connected (large $\langle k \rangle$) nodes have a large chance of dominating the connectivity.

IV. EXPERIMENTS AND RESULTS

The main underlying hypothesis tested in this paper is that a networked trust game would promote higher levels of trust and global net wealth in the population than that of an unstructured population [9]. We are also interested to examine the impact of different network structures on the level of trust and global net wealth, and under which initial conditions the population would give rise to higher levels of trust and global net wealth. To do so, we considered network structures ranging from regular to ER random networks as well as SF networks with higher and lower average degrees $\langle k \rangle$, and covered all possibilities of the initial population conditions by running a sensitivity analysis on the initial k_I and k_T values (k_U is given by $pop - (k_I + k_T)$). Through systematic Monte Carlo simulation experiments, we investigated the effects of different parameter settings, including the whole spectrum of temptation to defect values for parameter r_{UT} that determines the level of difficulty of the game (lower r_{UT} values mean less temptation to defect or be untrustworthy, i.e., the game is 'easier' in this case)². Appropriate statistical tests were carried out where claims are made about a particular setting or outcome being significantly better than the other.

In the following subsections, we first describe the experimental setup and values used for the model parameters. Then, we analyze the implications of different R_T and r_{UT} values for the game dynamics. After that, we discuss results based on the default SF network (with an average degree $\langle k \rangle = 4$), results with no network structure in place, and results with different network structures including SF networks with different densities and $\langle k \rangle$ as well as the regular and ER random networks.

²In the work of Abbas et al [9], they considered only a few specific instances of temptation to defect ratios of the game.

A. Experimental Setup and Parameters

Each of our model simulations was run with a population of 1,024 agents (i.e., $pop = 1,024$) and 5,000 time steps. It is worth pointing out that we also investigated the effects of smaller and larger population sizes during our preliminary analysis, and confirmed that a population size of 1,000 or more produces consistent simulation outcomes. Payoffs of the agents were calculated based on Equation 1, with R_T set to 6 (as in [9]). We show the results for three different values of the temptation to defect ratio r_{UT} , i.e., 0.11, 0.33 and 0.66, which cover three different trust scenarios ranging from the game being *easier* to *moderate* and *harder*. The majority of our experiments include a sensitivity analysis for the whole range of initial population conditions: the values of k_I , k_T and k_U , which determine the number of initial players that have strategies I , T and U , respectively.

The source code and runnable application is publicly available at www.manuchise.com/evolutionary-games-trust for readers to access the code, use, and reproduce results. All simulation experiments reported in this paper were repeated for 50 independent Monte Carlo runs, which changed the initial strategies of the agents on the same social network topology. Averaged results were calculated after all the simulations converged to an equilibrium state after 5,000 time steps. Also, results were computed over the 50 independent runs for each model setting.

We used five different values of m ($m = \{2, 3, 4, 6, 8\}$) for the Barabasi-Albert algorithm to generate SF networks with different densities and $\langle k \rangle$. When reporting the results, we refer to SF with $m = 3$ for comparisons between different network structures, since it has similar $\langle k \rangle$ (around 4) and density (0.004) values as the regular and ER random networks. The constructive random generator algorithm creates the ER network with the required $\langle k \rangle$ value and a binomial degree distribution. For each step of the generator, a new node is added and randomly connected with some of the old nodes. After the required steps of the generator, an ER random graph with a $density = \frac{\langle k \rangle}{pop-1}$ can be built. Table II shows details of the social networks used in this paper.

TABLE II
DETAILS OF THE USED SOCIAL NETWORKS.

Social Network	Density	$\langle k \rangle$	Clustering coefficient
SF2 ($m = 2$)	0.003	2.994	0.015
SF3 ($m = 3$)	0.004	3.978	0.021
SF4 ($m = 4$)	0.005	5.019	0.031
SF6 ($m = 6$)	0.007	6.986	0.040
SF8 ($m = 8$)	0.009	9.056	0.060
REGULAR	0.004	4	0
ER4	0.004	4.021	0.005

B. Analysis of the Trustworthy Multiplier and Temptation to Defect Ratio

In this section, we report on the implications of different values for the trustworthy multiplier R_T and temptation to defect ratio r_{UT} . Figure 2 shows the final state of the population in terms of surviving strategies and global net wealth for

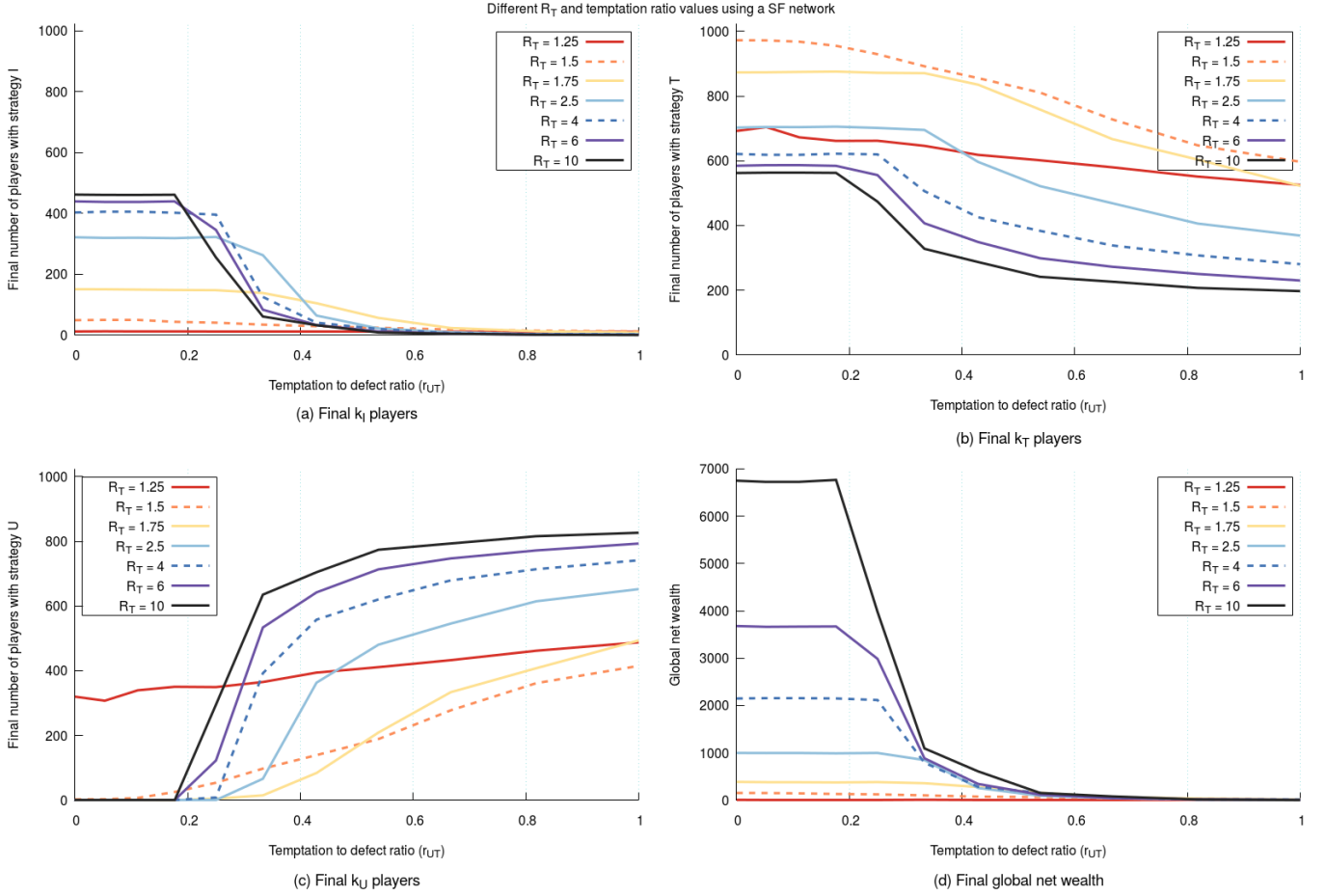


Fig. 2. Averaged numbers of final k_I (top left), k_T (top right) and k_U (bottom left), as well as the averaged global net wealth W (bottom right) for different specifications of R_T and temptation to defect ratio r_{UT} .

different specifications of R_T , from 1.25 to 10. For each of the R_T values, we performed sensitivity analysis with respect to the temptation to defect ratio r_{UT} by changing its value over the range of 0 to 1 using a step size of 0.01³. We used SF with $\langle k \rangle \approx 4$ and an initial population of 30% of strategy I (i.e., $k_I = 307$), 25% of strategy T (i.e., $k_T = 256$) and 45% of strategy U (i.e., $k_U = 461$) in this analysis.

We see in Figure 2 that there are three regions of interest. When r_{UT} is low ($\lesssim 0.2$), the game is easy resulting in high global net wealth. When r_{UT} is high ($\gtrsim 0.5$), the game becomes harder and almost all the players adopt strategy U with the global net wealth $W = 0$. In between, there is an interesting drift from $r_{UT} \gtrsim 0.18$ to $r_{UT} \lesssim 0.5$ where the population dynamics and global net wealth change dramatically because the initial I and T players shift their strategies to untrustworthy strategy U , decreasing the final global net wealth W of the population.

The findings here show that the model is not totally independent from R_T values, as differences in the population dynamics appear. As observed by Abbass et al. [9], the lower the values, the more trustworthy players there are in the final

population. Lower R_T values also lead to less investors I and reinforce the existence of untrustworthy players U easily. This causes investors to go extinct quicker ($k_I = 0$), which then limits the spread of untrustworthy players U . In this region, i.e., low R_T values, T players are therefore dominant.

It is worth noting that when having very low R_T values (e.g., 1.25 or 1.5 in the plots of Figure 2), the dynamics of the game differ from the rest of the R_T values. This is because of the payoff calculation for players with strategy I (first term of Equation 1). Due to space constraints, we report only results with R_T equal to 6 in the rest of this paper, as it represents a value for which we have a ‘normal’ trust game. Nevertheless, we still present results with three different values for r_{UT} , i.e., $r_{UT} = 0.11, 0.33$ and 0.66 , since the population dynamics are different when the difficulty levels of the game vary (as can be seen in Figure 2).

C. Results with SF networks

In this section, we analyze the dynamics of our model with default SF settings (i.e., a SF3 topology with $\langle k \rangle = 4$, and $R_T = 6$). In order to study these dynamics we ran a sensitivity analysis of the initial conditions of the population (i.e., initial number of players k_I, k_T , and k_U) for three different temptation to defect ratio values $r_{UT} = \{0.11, 0.33, 0.66\}$. We show

³Note that both values $r_{UT} = 0$ and $r_{UT} = 1$ do not represent a valid trust game according to the game restriction $R_T < R_U < 2R_T$, but they are still included in the plots for completeness’ sake.

in Figure 3, 12 heatmaps of the distribution of the existing strategies of the population and its global net wealth after reaching equilibrium in 5,000 steps. Each pixel of a heatmap is the averaged Monte Carlo simulation value for the number of players or the global net wealth value. The three rows of the figure correspond to the dynamics using the three different r_{UT} values.

We can see that when $r_{UT} = 0.11$ (first row of the four plots in Figure 3), the networked trust model is able to promote high levels of trust. Investors and trustworthy players (I and T strategies, respectively) are eliminated only when the initial population is clearly dominated by untrustworthy players ($k_U \gtrsim 700$). For the majority of the initial population conditions, however, trust is clearly promoted with almost no U players at the final state of the population. Obviously, high global net wealth (the light area of the fourth heatmap in the first row of Figure 3) is attributed to the absence of untrustworthy players U (the light area of the third heatmap of the first row) and the increase of investors I and trustworthy players T (first and second heatmaps of the first row).

When an intermediate value for the temptation to defect ratio is in place (i.e., $r_{UT} = 0.33$, shown in the second row of heatmaps of Figure 3), the initial conditions to promote trust have to be more favorable compared to when r_{UT} is low. Nevertheless, the model is still able to promote trust when the initial population contains a high number of investors and trustworthy individuals ($(k_I + k_T) > 900$). Under these circumstances, the population obtains a high level of global net wealth W and untrustworthy players are almost fully eliminated at the end of the simulation (see the third heatmap of the second row of Figure 3).

Finally, when setting r_{UT} to 0.66, cooperating is hard for trustees. In this case, non-zero net wealth W can be achieved only when there is no untrustworthy player at all in the initial population. See the top right pixel $(x, y) = 512, 512$ of all the heatmaps in the third row of Figure 3 but take into account the zoomed X and Y axis for these four heatmaps for better readability. For all other distributions of the initial population, investors are extincted and W is simply 0. The second heatmap of the third row shows that trustworthy trustees T still exist, although they are less than untrustworthy trustees U (third heatmap of the row).

D. Results with No Network Structure

In this section, we present results with no specific network structure in place, i.e., an unstructured population equivalent to the original trust model proposed by Abbass et al. [9], for a direct comparison. Figure 4 shows, when setting r_{UT} to 0.11, the heatmaps with the numbers of investors k_I (the first heatmap), trustworthy trustees k_T (the second heatmap), untrustworthy trustees k_U (the third heatmap) of the final population as well as its global net wealth (the last fourth plot).

As expected, the results here are equivalent to those of Abbass et al. [9]. Trust cannot be promoted in the population except in the specific case of not having untrustworthy players in the initial population ($k_U = 0$, see the top right pixel in

the heatmaps and take into account the zoomed X and Y axis for better readability). Even with just a single U player in the initial population, individuals adopting the other two strategies would rapidly change to the U strategy. These results are in stark contrast compared to our networked trust model, where even when $r_{UT} = 0.33$ a higher level of trust can be obtained easily (see second row of heatmaps of Figure 3). In an unstructured population, the information diffusion speed is very high [44]. The benefit of being selfish or untrustworthy is immediately spread throughout the population, and therefore the players rampantly adopt the untrustworthy strategy U .

E. Results with Different Network Structures

In this section, we analyze two scenarios: 1) the influence of having SF networks with different densities and 2) using regular and ER random networks for the game. Figures 5, 6 and 7 show the direct comparisons of the latter networks for 12 different and independent initial conditions of the population configurations, for the three r_{UT} values 0.11, 0.33, and 0.66, respectively. Additionally, we also performed the same sensitivity analysis reported in the previous two sections for the least dense SF network (generated with $m = 2$ and having $\langle k \rangle \approx 3$, see first row of heatmaps in Figure 8), a SF network with $\langle k \rangle \approx 4$ to compare it with other network topologies having the same density (generated with $m = 3$, see second row of heatmaps in Figure 8), the most dense SF network (generated with $m = 8$ and having $\langle k \rangle \approx 9$, see third row of heatmaps in Figure 8), the regular lattice (see fourth row of heatmaps of Figure 8), and the ER random graph with $\langle k \rangle = 4$ (see the last row of heatmaps in Figure 8).

First, let us look at the results of different network structures for temptation to defect $r_{UT} = 0.11$. Figure 5 shows the population state after 5,000 steps (averaged number of k_I , k_T , and k_U players) as well as its averaged global net wealth W for the seven considered network structures in the 12 independent initial population settings. In this case, the regular lattice and most connected SF networks are generally those that promote higher levels of trust. This fact is consistent for the majority of the 12 different initial population configurations shown in Figure 5. SF with $\langle k \rangle \approx 4$ is not capable of promoting trust at the same level as other network structures with equivalent average degrees, i.e., regular and ER random networks. Therefore, under easy conditions of the game, the use of more homogeneous network structures such as the regular lattice can obtain significantly better global net wealth than heterogeneous network structures such as SF networks of similar average degrees under most of the initial population conditions. Bonferroni-Dunn's tests conducted with a significance level of $\alpha = 0.05$ confirmed the statistical significance of differences between the results obtained (e.g., REGULAR versus SF2 or SF3).

Figure 6 shows the results of the same network structure settings as in Figure 5 but this time with $r_{UT} = 0.33$, which makes the game harder than when $r_{UT} = 0.11$. To help with the analysis, we also include heatmaps for two additional SF structures (in addition to those already shown in Section IV-C), the regular lattice and ER random network (see the five rows

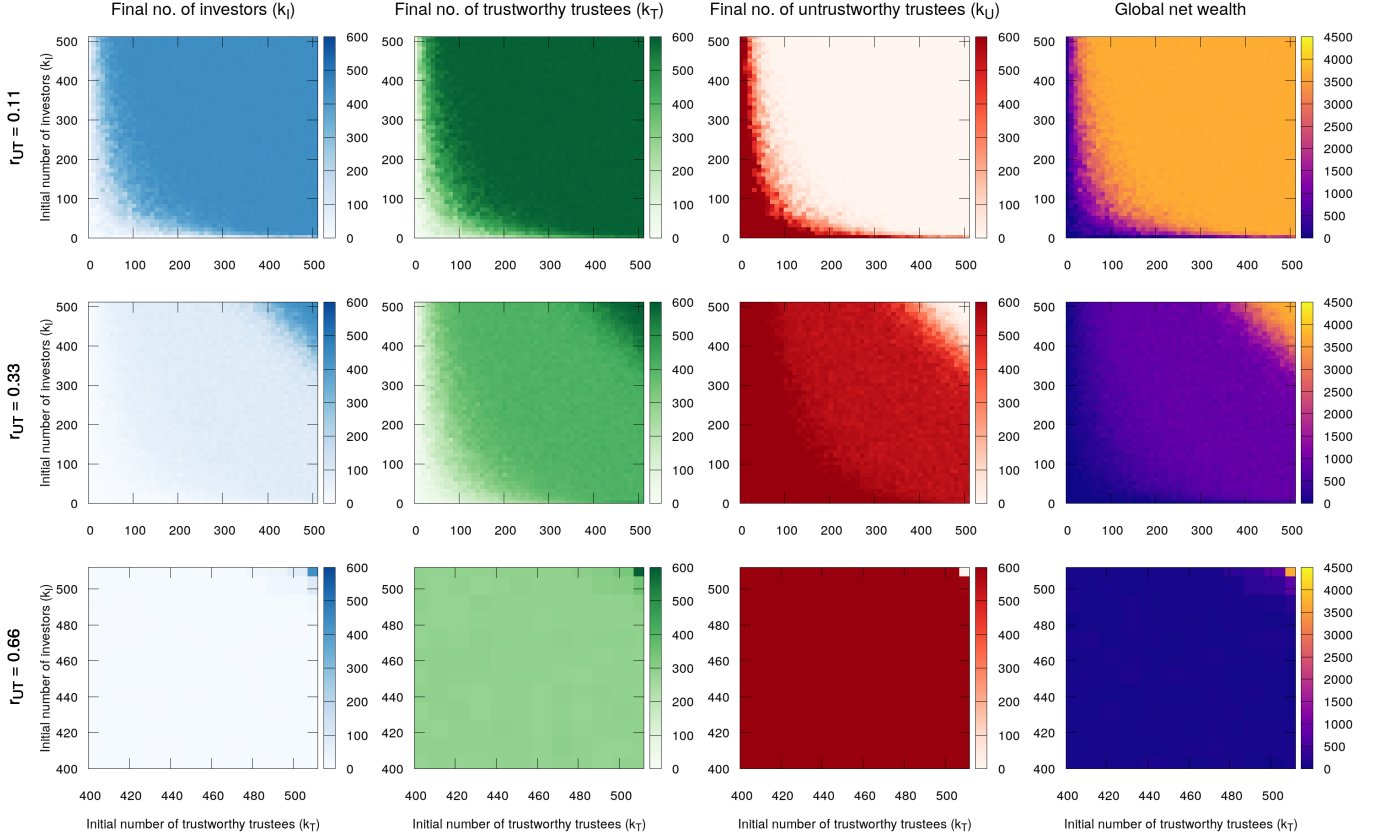
Final population heatmaps for SF3 ($R_T=6$, $r_{UT}=0.33$)

Fig. 3. Sensitivity analysis of the final numbers of players (k_I , k_T , and k_U) and global net wealth of the population to the initial population setting (i.e., number of initial k_I , k_T , and k_U) for a SF network generated by $m = 3$. First row of four plots was obtained with a temptation to defect r_{UT} equals to 0.11, second row with r_{UT} equals to 0.33, and third row with r_{UT} equals to 0.66.

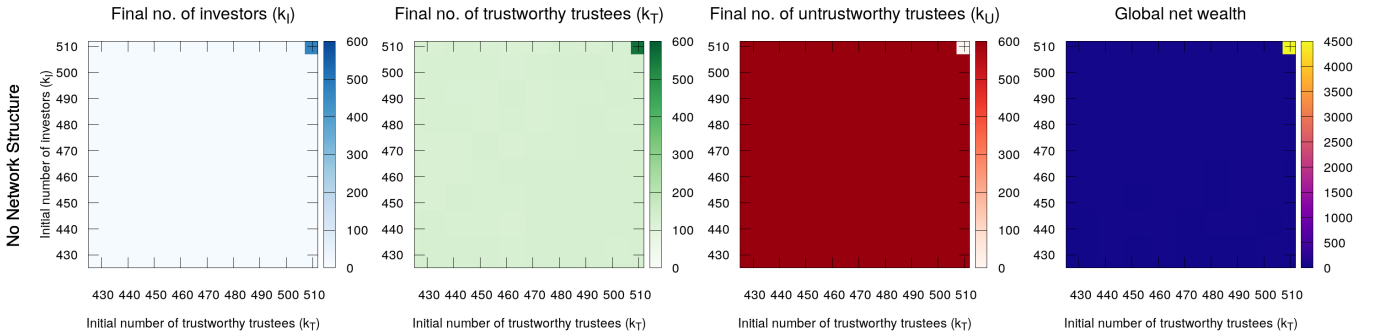
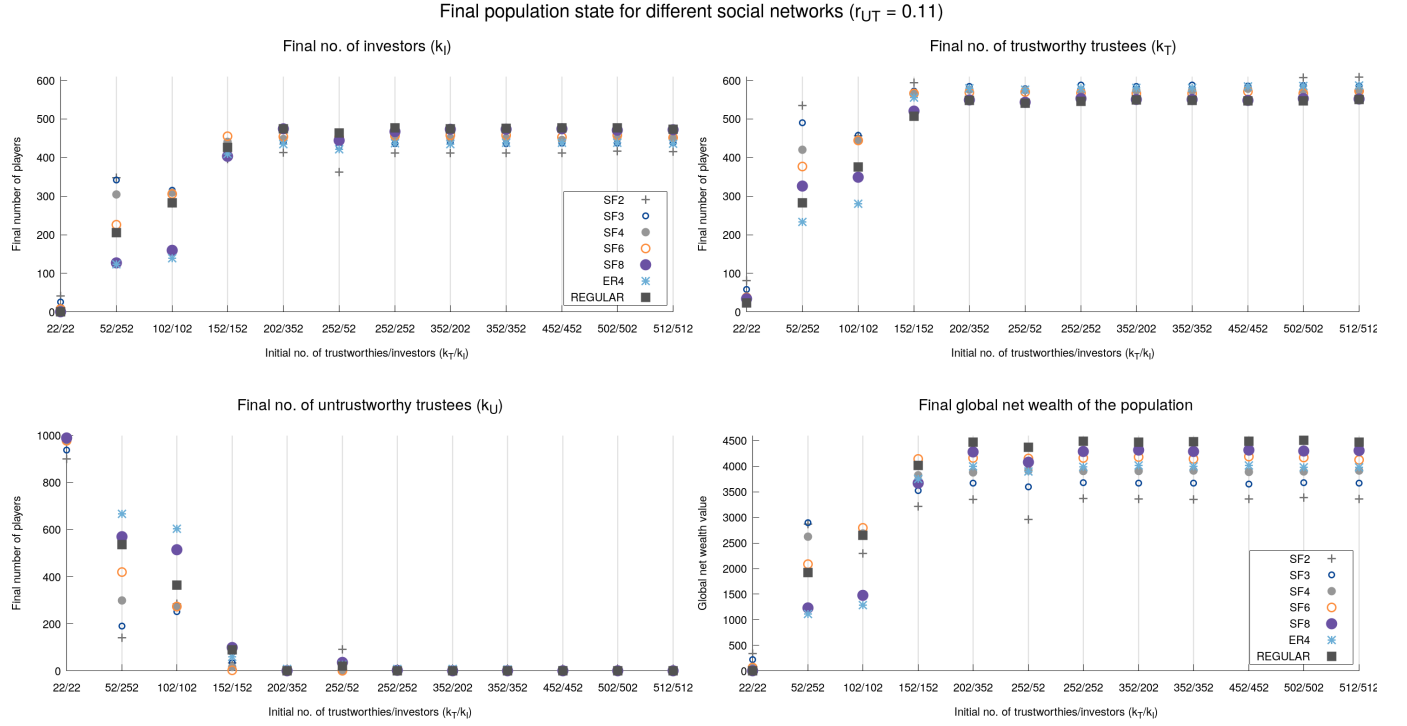
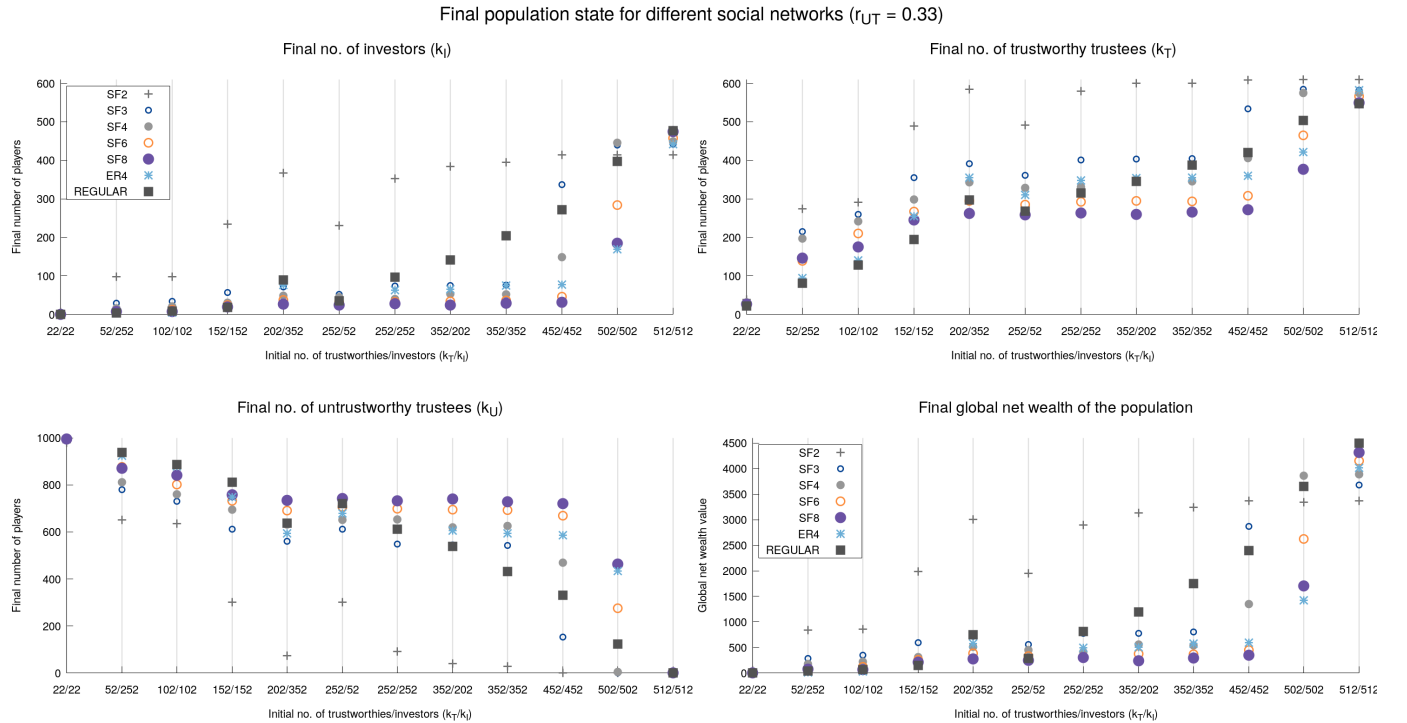
Population state after 5000 simulation steps (No Network Structure and $R_T = 6$)

Fig. 4. Sensitivity analysis of the final numbers of players (k_I , k_T , and k_U) and global net wealth of the population to the initial population setting (i.e., number of initial k_I , k_T , and k_U) with no network structure and temptation to defect r_{UT} equals to 0.11.

of heatmaps in Figure 8). As we can clearly see in these plots, the results are quite different than when $r_{UT} = 0.11$. When $r_{UT} = 0.33$, heterogeneous and low density social networks, i.e., SF networks with $\langle k \rangle \approx 3$ and 4, facilitate the highest levels of trust and global net wealth.

Generally speaking, here the least connected social networks obtain the best results in terms of trust and global net wealth.

The only exception being when the initial population is very favorable for the game, e.g., almost all the initial players have either I or T strategies. When this is the case, the regular lattice is able to obtain a high level of trust. With the most extreme initial population configuration ($k_I = k_T = 512$ and $k_U = 0$), highly connected network structures such as the regular lattice and SF networks with $m \geq 6$ again obtain

Fig. 5. Comparison of SF networks with different densities as well as regular and ER random networks with temptation to defect $r_{UT} = 0.11$.Fig. 6. Comparison of SF networks with different densities as well as regular and ER random networks with temptation to defect $r_{UT} = 0.33$.

significantly better global net wealth than SF networks with $m \leq 3$. This is because it is the structure that promotes higher k_I when $k_U = 0$. For the majority of the specifications, when the SF network is sparsely connected, more investors I shift to strategy T than that in a homogeneous structure (e.g., regular

and ER networks). It is worth pointing out that given all the possible initial conditions of the population, regular and ER networks have much more restricted conditions in promoting trust. See, for example, the heatmaps of the third row (SF8), fourth row (regular) and last row (ER4) in Figure 8 with

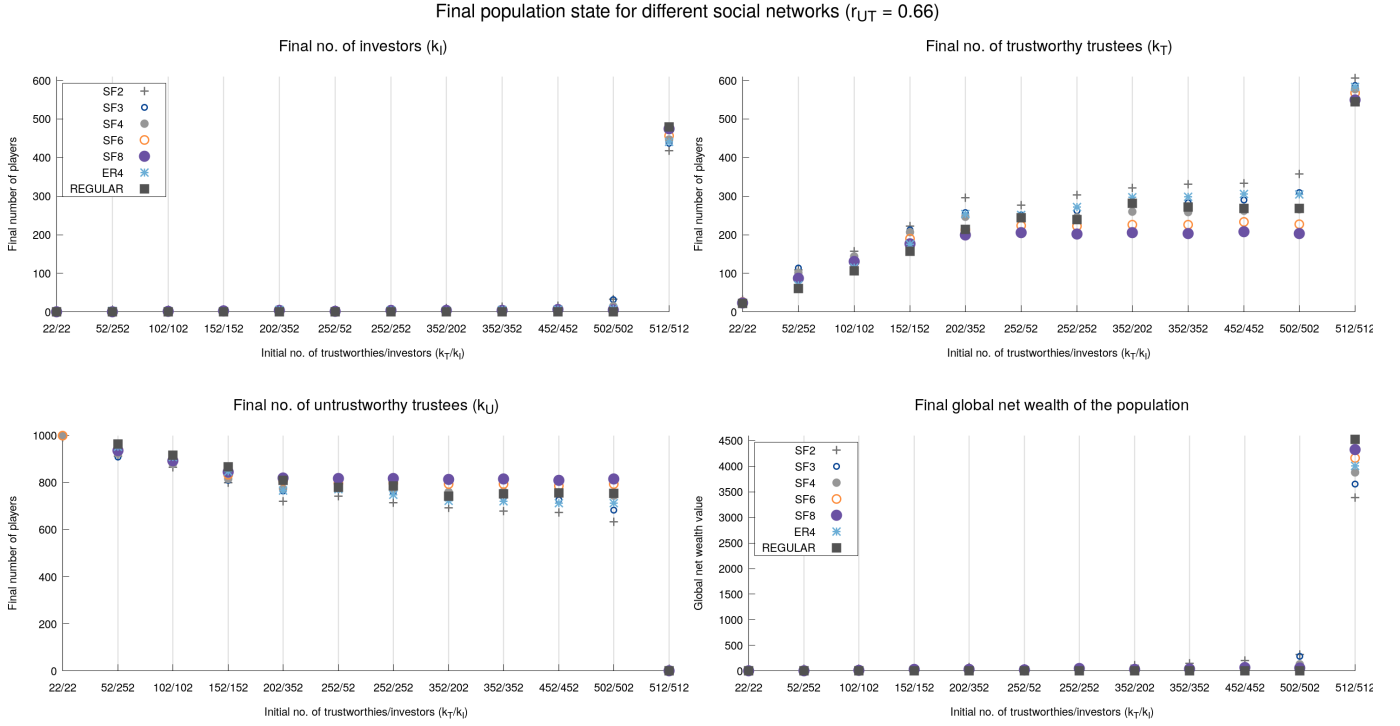


Fig. 7. Comparison of SF networks with different densities as well as regular and ER random networks with temptation to defect $r_{UT} = 0.66$.

respect to the wide ‘cooperation’ area obtained by the least connected SF structure (first row of heatmaps in Figure 8).

More links (i.e., higher densities in the social network) lead to a higher information diffusion speed through the network. This means that untrustworthy players U can quickly spread in the population, since they would have the highest individual payoffs. However, if we compare the heatmaps of the second row (SF3), fourth row (regular), and the last row (ER4) in Figure 8, under the same setting of $\langle k \rangle = 4$, we see that when conditions for the game are not clearly beneficial (high k_U or high r_{UT}), the heterogeneity of the population (i.e., SF structures) can actually help in promoting higher levels of trust.

Finally, Figure 7 shows the results of different network structures when $r_{UT} = 0.66$. In this case, we see similar trends for all the social network structures considered. A substantial amount of global net wealth can be observed only when there is no untrustworthy trustee U in the initial population. As trust is easily promoted in this situation because the population is totally trustworthy, the level of global net wealth obtained is the same as in previous cases when $r_{UT} = 0.11$ and $r_{UT} = 0.33$, with the regular lattice ranked top followed by the most dense SF networks (i.e., SF structures generated with $m = 8, 6, 4$). These results are also consistent with findings from previous analyses: when the population has no untrustworthy trustee U , homogeneous network structures such as the regular lattice would give rise to a higher number of investors I than trustworthy trustees T compared to a SF network with the same density, thereby ending up with higher global net wealth. Again, Bonferroni-Dunn’s tests conducted with $\alpha = 0.05$ confirmed the statistical significance of differences

between the results obtained (for scenarios with $r_{UT} = 0.33$ and $r_{UT} = 0.66$).

V. CONCLUDING REMARKS AND FUTURE WORK

In this paper, we investigated the evolution of trust and global net wealth through a networked N -player trust model. Three strategies are possible for the players, i.e., being an investor, a trustworthy trustee and an untrustworthy trustee. The network structure in place confines players in the population to interact locally with their neighbors.

An important aspect of this work is that we have introduced a temptation to defect ratio r_{UT} to study the transition of the game from easy ($r_{UT} = 0.11$) to a more difficult level ($r_{UT} = 0.66$). For different values of this ratio, we explored the dynamics of the model using different network structures, including the regular, SF and ER random networks. We observed that trust can be promoted with our networked N -player model when the r_{UT} values range from low (e.g., 0.11) to medium (e.g., 0.33). In fact, high levels of trust and global net wealth can be achieved when $r_{UT} = 0.11$. Even with $r_{UT} = 0.33$, positive levels of global net wealth can be obtained for a wide range of initial population conditions. When the game is harder (e.g., $r_{UT} = 0.66$), however, trust is only promoted when no untrustworthy players are present in the initial population.

We also studied the dynamics of our model with no network structure in place. In this case, even one untrustworthy player can fully eliminate investors and lead to zero global net wealth. These results are in line with those obtained by Abbass et al. [9] based on a well-mixed, unstructured population.

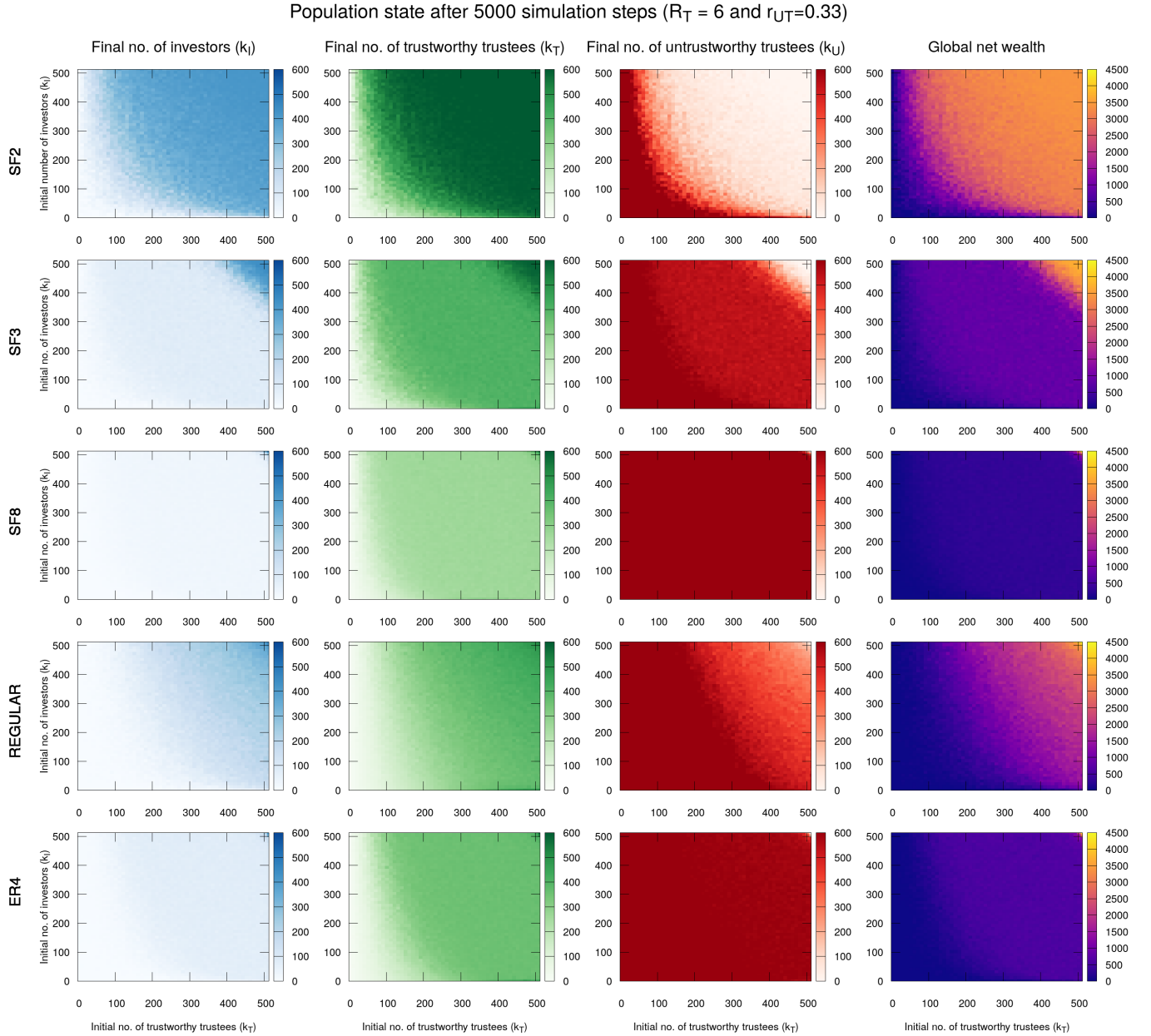


Fig. 8. Sensitivity analysis of the final numbers of players (k_I , k_T , and k_U) and global net wealth of the population to the initial population setting (i.e., number of initial k_I , k_T , and k_U) for SF2 network ($\langle k \rangle \approx 3$), SF3 network ($\langle k \rangle \approx 4$), SF8 network ($\langle k \rangle \approx 9$), regular lattice, and ER when temptation to defect r_{UT} is set to 0.33.

A major contribution of this piece is that we discovered the importance of network densities for promoting trust and global net wealth. When game conditions are easy, SF structures with high densities can obtain higher levels of trust and global net wealth. The homogeneous regular lattice is even better in this regard. Interestingly, however, when the game conditions turn harder (e.g., $r_{UT} = 0.33$), heterogeneous but sparsely connected SF structures are better for promoting trust and global net wealth. SF networks are known for creating social diversities in the population [37], which facilitate trust when conditions are not totally favorable. However, when no trustworthy players are present in the initial population, homogeneous network structures with less randomness (e.g.,

the regular lattice) can obtain the best net wealth results (especially when the temptation to defect ratio is low).

There are several possibilities for future research. One is to study game dynamics using a mathematical approach. We are also interested in examining the impact of having a historical account of players' past strategies [14] and different update rules (e.g., unconditional imitation [27] or Moran's rules [13]) on the evolutionary outcomes. Last but not least, we are keen to apply our trust model to real-world problems where trust is crucial, such as the sharing economy [45].

ACKNOWLEDGMENTS

This work was supported by the Australian Government's 2016 Endeavor Research Fellowship and Spanish Ministerio de Economía y Competitividad under the NEWSOCO project (ref. TIN2015-67661-P), including European Regional Development Funds (ERDF).

REFERENCES

- [1] E. Petraki and H. A. Abbass, "On trust and influence: A computational red teaming game theoretic perspective," in *the 2014 Seventh IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA)*. IEEE, 2014, pp. 1–7.
- [2] E. Fehr and U. Fischbacher, "The nature of human altruism," *Nature*, vol. 425, no. 6960, pp. 785–791, 2003.
- [3] D. Gambetta *et al.*, "Can we trust trust," *Trust: Making and breaking cooperative relations*, vol. 13, pp. 213–237, 2000.
- [4] F. S. Grodzinsky, K. W. Miller, and M. J. Wolf, "Developing automated deceptions and the impact on trust," *Philosophy & Technology*, vol. 28, no. 1, pp. 91–105, 2015.
- [5] D. S. Reina and M. L. Reina, *Trust & betrayal in the workplace: Building effective relationships in your organization*. Berrett-Koehler Publishers, 2006.
- [6] A. Baier, "Trust and antitrust," *Ethics*, vol. 96, no. 2, pp. 231–260, 1986.
- [7] F. D. Schoorman, R. C. Mayer, and J. H. Davis, "An integrative model of organizational trust: Past, present, and future," *Academy of Management review*, vol. 32, no. 2, pp. 344–354, 2007.
- [8] A. W. Brooks, H. Dai, and M. E. Schweitzer, "Im sorry about the rain! superfluous apologies demonstrate empathic concern and increase trust," *Social Psychological and Personality Science*, vol. 5, no. 4, pp. 467–474, 2014.
- [9] H. Abbass, G. Greenwood, and E. Petraki, "The n-player trust game and its replicator dynamics," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 3, pp. 470–474, 2016.
- [10] J. J. Jordan, M. H. Hoffman, M. A. Nowak, and D. G. Rand, "Uncalculating cooperation is used to signal trustworthiness," *Proceedings of the National Academy of Sciences*, vol. 113, no. 31, pp. 8658–8663, 2016.
- [11] M. L. Manapat, M. A. Nowak, and D. G. Rand, "Information, irrationality, and the evolution of trust," *Journal of Economic Behavior & Organization*, vol. 90, pp. S57–S75, 2013.
- [12] C. E. Tarnita, "Fairness and trust in structured populations," *Games*, vol. 6, no. 3, pp. 214–230, 2015.
- [13] D. Vilone, J. J. Ramasco, A. Sánchez, and M. San Miguel, "Social imitation versus strategic choice, or consensus versus cooperation, in the networked prisoner's dilemma," *Physical Review E*, vol. 90, no. 2, p. 022810, 2014.
- [14] R. Chiong and M. Kirley, "Effects of iterated interactions in multiplayer spatial evolutionary games," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 4, pp. 537–555, 2012.
- [15] H. Ishibuchi and N. Namikawa, "Evolution of iterated prisoner's dilemma game strategies in structured demes under random pairing in game playing," *IEEE Transactions on evolutionary computation*, vol. 9, no. 6, pp. 552–561, 2005.
- [16] S. Mittal and K. Deb, "Optimal strategies of the iterated prisoner's dilemma problem for multiple conflicting objectives," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 3, pp. 554–565, 2009.
- [17] C. Hauert and M. Doebeli, "Spatial structure often inhibits the evolution of cooperation in the snowdrift game," *Nature*, vol. 428, no. 6983, pp. 643–646, 2004.
- [18] W.-X. Wang, J. Ren, G. Chen, and B.-H. Wang, "Memory-based snowdrift game on networks," *Physical Review E*, vol. 74, no. 5, p. 056113, 2006.
- [19] J. Berg, J. Dickhaut, and K. McCabe, "Trust, reciprocity, and social history," *Games and Economic Behavior*, vol. 10, no. 1, pp. 122–142, 1995.
- [20] J. C. Cox, "How to identify trust and reciprocity," *Games and economic behavior*, vol. 46, no. 2, pp. 260–281, 2004.
- [21] B. King-Casas, D. Tomlin, C. Anen, C. F. Camerer, S. R. Quartz, and P. R. Montague, "Getting to know you: reputation and trust in a two-person economic exchange," *Science*, vol. 308, no. 5718, pp. 78–83, 2005.
- [22] L. A. Bach, T. Helvik, and F. B. Christiansen, "The evolution of n-player cooperation threshold games and its bifurcations," *Journal of Theoretical Biology*, vol. 238, no. 2, pp. 426–434, 2006.
- [23] C. S. Gokhale and A. Traulsen, "Evolutionary games in the multiverse," *Proceedings of the National Academy of Sciences*, vol. 107, no. 12, pp. 5500–5504, 2010.
- [24] M. A. Nowak, C. E. Tarnita, and T. Antal, "Evolutionary dynamics in structured populations," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 365, no. 1537, pp. 19–30, 2010.
- [25] F. Gargiulo and J. J. Ramasco, "Influence of opinion dynamics on the evolution of games," *PloS one*, vol. 7, no. 11, p. e48916, 2012.
- [26] M. Newman, A.-L. Barabási, and D. J. Watts, *The structure and dynamics of networks*. Princeton University Press, 2006.
- [27] M. A. Nowak and R. M. May, "Evolutionary games and spatial chaos," *Nature*, vol. 359, no. 6398, pp. 826–829, 1992.
- [28] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.
- [29] A. L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [30] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, pp. 17–61, 1960.
- [31] K. H. Schlag, "Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits," *Journal of Economic Theory*, vol. 78, no. 1, pp. 130–156, 1998.
- [32] D. Helbing and W. Yu, "The future of social experimenting," *Proceedings of the National Academy of Sciences*, vol. 107, no. 12, pp. 5265–5266, 2010.
- [33] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, "A simple rule for the evolution of cooperation on graphs and social networks," *Nature*, vol. 441, no. 7092, pp. 502–505, 2006.
- [34] F. C. Santos, J. M. Pacheco, and T. Lenaerts, "Evolutionary dynamics of social dilemmas in structured heterogeneous populations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 9, pp. 3490–3494, 2006.
- [35] G. Szabó and G. Fath, "Evolutionary games on graphs," *Physics reports*, vol. 446, no. 4, pp. 97–216, 2007.
- [36] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floría, and Y. Moreno, "Evolutionary dynamics of group interactions on structured populations: a review," *Journal of The Royal Society Interface*, vol. 10, no. 80, p. 20120997, 2013.
- [37] F. C. Santos, M. D. Santos, and J. M. Pacheco, "Social diversity promotes the emergence of cooperation in public goods games," *Nature*, vol. 454, no. 7201, pp. 213–216, 2008.
- [38] H.-X. Yang, W.-X. Wang, Z.-X. Wu, Y.-C. Lai, and B.-H. Wang, "Diversity-optimized cooperation on complex networks," *Physical Review E*, vol. 79, no. 5, p. 056107, 2009.
- [39] H. Zhang, H. Yang, W. Du, B. Wang, and X. Cao, "Evolutionary public goods games on scale-free networks with unequal payoff allocation mechanism," *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 5, pp. 1099–1104, 2010.
- [40] G. Wei, P. Zhu, A. V. Vasilakos, Y. Mao, J. Luo, and Y. Ling, "Cooperation dynamics on collaborative social networks of heterogeneous population," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, pp. 1135–1146, 2013.
- [41] M. Chica and W. Rand, "Building agent-based decision support systems for word-of-mouth programs. a freemium application," *Journal of Marketing Research*, vol. 54, pp. 752–767, October 2017.
- [42] D. Helbing, "Interrelations between stochastic equations for systems with pair interactions," *Physica A*, vol. 181, no. 1-2, pp. 29–52, 1992.
- [43] B. Libai, R. Bolton, M. S. Bügel, K. De Ruyter, O. Götz, H. Risselada, and A. T. Stephen, "Customer-to-customer interactions: broadening the scope of word of mouth research," *Journal of Service Research*, vol. 13, no. 3, pp. 267–282, 2010.
- [44] C.-H. Yeh and C.-Y. Yang, "Social networks and asset price dynamics," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 3, pp. 387–399, 2015.
- [45] M. Chica, R. Chiong, M. T. Adam, S. Damas, and T. Teubner, "An evolutionary trust game for the sharing economy," in *Evolutionary Computation (CEC), 2017 IEEE Congress on*. IEEE, 2017, pp. 2510–2517.