

RESEARCH

Open Access



Homophily in networked agent-based models: a method to generate homophilic attribute distributions to improve upon random distribution approaches

Marie Lisa Kapeller* , Georg Jäger and Manfred Füllsack

*Correspondence:
marie.kapeller@uni-graz.at
University of Graz,
Merangasse 18/I A,
8010 Graz, Austria

Abstract

In the standard situation of networked populations, link neighbours represent one of the main influences leading to social diffusion of behaviour. When distinct attributes coexist, not only the network structure, but also the distribution of these traits shape the typical neighbourhood of each individual. While assortativity refers to the formation of links between similar individuals inducing the network structure, here, we separate the formation of links from the actual distribution of an attribute on the topology. This is achieved by first generating different network types (e.g., lattice, scale free, and small world), followed by the procedure of distributing attributes. With this separation, we try to isolate the effect that attribute distribution has on network diffusion from the effect of the network structure itself. We compare random distributions, where behaviour types are highly mixed, and homophilic distributions, where similar individuals are very likely to be linked, and examine the effects on social contagion in a population of mainly reciprocal behaviour types. In addition, we gradually mix homophilic distribution, by random rewiring, adding links and relocating individuals. Our main results is that attribute distribution strongly influences collective behaviour and the actual effect depends on the network type. Under homophilic distribution the equilibrium collective behaviour of a population tends to be more diverse, implying that random distributions are limited for illustration of collective behaviour. We find that our results are robust when we use different gradual mixing methods on homophilic distribution.

Keywords: Network diffusion, Social contagion, Small-world networks, Scale-free networks, Agent-based models, Homophily

Introduction

The influence of social context on individual behaviour is a notable topic in various fields. In its most basic form, the problem can be reduced to the question of how a certain property (attribute, behaviour, decision, and trait) spreads throughout a social system. To analyse such processes in greater detail, modelling is a viable way of gaining new insights. Agent-based models (ABM) can cover heterogeneity among individuals and provide a useful tool to study social contagion dynamics. Applications of ABM capture phenomena as diverse as rumour spreading [28], memory transmission [27], attitude

polarization [39], and social norm contagion in the case study of protection rackets of the Sicilian Mafia [30].

The structure of the interaction network is known to be deeply connected to the overall spreading pattern [7, 10, 18, 36, 43, 44]. Many studies have been dedicated to deepen our understanding of the role of network features, such as central nodes [2], clustering [15], and weak ties [13] and structural characteristics such as centrality and bridging ties have been identified to foster diffusion processes [44].

Diffusion patterns are highly contextual. They do not only depend on the general network structure, but also on who tends to be connected with whom. The tendency to associate with beings similar to oneself is known as homophily or assortative mixing [32]. Preferences in ‘who to interact with’ generate social patterns known as bonding and bridging [35]. While bonding between homogeneous groups can be valuable for marginalized members of society, bridging of heterogeneous groups allows different individuals to share and exchange information and ideas and build consensus among groups representing diverse interests [34].

A common way to include the homophily effect in ABMs is to utilize a network generator that uses a higher chance to generate links between similar individuals [6, 16, 19, 24]. However, since the proportions of population shares have a direct effect on link formation and thus topology, the resulting network may differ from a network that is generated by an unbiased generator. On the other hand, many studies on networked ABM generation simplify the distribution of traits, assuming that they are random [9, 33, 36].

The structural proximity (network structure) and the attribute proximity (homophily effect) have been tackled by several lines of research. Node embedding refers to techniques that try to find ‘similar’ nodes in the graph, being a valuable way for classification, link prediction, and graph visualization. In this context, most work focuses on structural proximity [8, 47], while the homophily effect of attributes was incorporated in [25]. Community embedding [4] has been shown to be beneficial for community detection and node classification. Community detection is especially beneficial in the analysis of real-world data sets, but to our knowledge, less applicable to improve ABM network generation.

How homophily affects diffusion and contagion in ABM networks with heterogeneous agents has been studied in [11] using a probability model of homophily. This study merges the network generation and the attribute distribution and is limited to Erdős–Rényi random network models. Conditions under which a behaviour diffuses and becomes persistent in the population have been investigated in [17]. This study captures many important aspects of diffusion such as the level of homophily, but comes short in investigating the homophily effect on different network types due to their focus on adoption types and the type of interaction mechanisms.

With this study, we want to present an attribute distribution mechanism for heterogeneous agents, usable for various network types. The mechanism generates homophilic attribute distributions, where individuals with identical traits are highly clustered, while the network structure is unaffected by the attribute distribution. We refer to different distribution of attributes on the network topology as allocation. To decouple the network generation and the allocation of individuals, we first generate a topology with an

unbiased network generator, and after the network generation is complete, position the heterogeneous individuals.

The analysis of the effect of the homophilic attribute distribution mechanism focusses on a comparison of diffusion in homophilic and random allocations. For this, we use a general model of diffusion that promotes a simple social contagion process through the population. To facilitate the discussion on the contagious property, we interpret the property as a continuous variable that governs a decision making process. Various contexts of the contagious process are suitable, e.g., environmental awareness or competitiveness, provided that the decision can be related to a continuous scale.

We can compare diffusion in different allocations on various network types: lattice topologies, scale-free topologies, cave-people topologies, spatial-proximity topologies, and small-world topologies. In addition, we generate and explore intermediate attribute distributions, which combine features of both the mixed and homophilic allocations. For this, we introduce random alterations via rewiring and constant changes of the topology, and target-oriented alterations via additional long ties and repositioning of individuals.

Method

We use a software modelling approach to create a population of agents, who make a certain decision at every time step. The attribute that governs this decision is on a continuous scale from 0 to 1. Individuals are connected via links resulting in an underlying interaction network topology. The influence of link neighbours creates diffusion dynamics within the modelled population, leading to social contagion processes of the property in question. The model is generic in the sense that the actual decision that is made does not need to be specified and can be interpreted, e.g., as environmental awareness [22] or an investment in game theory [3].

Agent behaviour types

In addition to the structure of a population, social contagion is also deeply connected to the response of individuals to their surroundings. Not all individuals or groups react in the same way to their environment. While some are easier influenced by their peers and reciprocate observed behaviour, others might be less flexible and do not change their actions based on the behaviour of others. A simple example of such non-reciprocators is the strategy of so-called continuous cooperators in the public goods game [3]. These players do not deviate from their decision to always invest in the creation of a public good, even when their peers do not contribute.

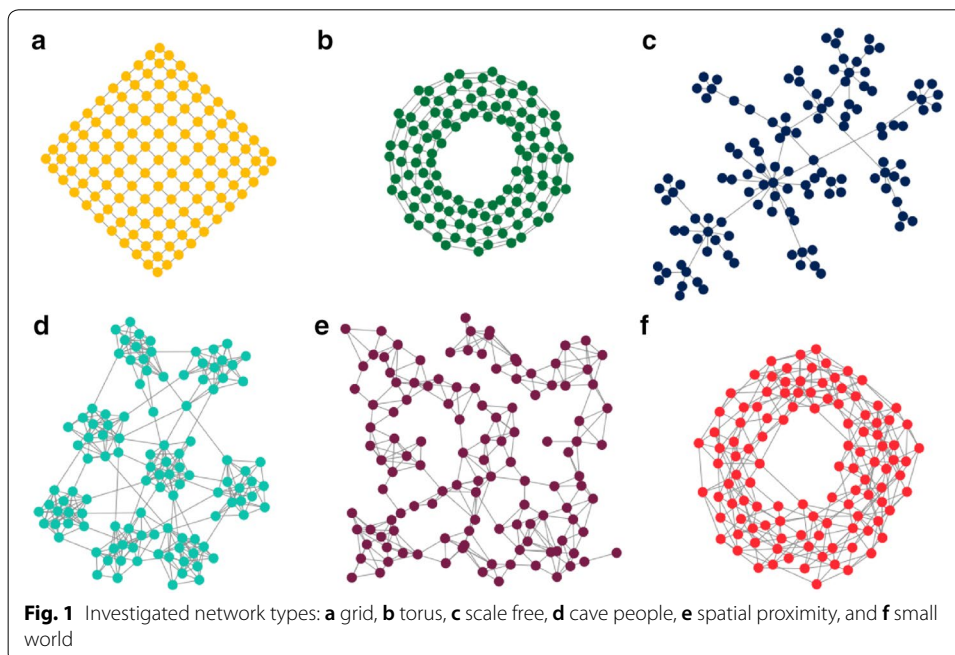
In the presented model, the population is divided in three types of agents: non-reciprocative type A, non-reciprocative type B, and reciprocative type S. Their attribute is given by the decision mechanism, which differs for each type. This behavioural type, not to be confused with the contagious property, is constant over time for each individual. Type A and type B individuals abide to the same decision and cannot be influenced by their neighbourhood. Their decision attribute is constantly 0 and 1, respectively, and thus not directly affected by social contagion. Type S individuals make their decision using a best-response mechanism, reflecting the mean decision of its direct neighbourhood (link neighbours).

The focus of our investigation is on spreading dynamics and local pattern formation through the population share of reciprocal individuals S , induced by the decisions of non-reciprocal types A and B .

Network topologies

We use six different topology types, as shown in Fig. 1. Each nodes represents an individual. Link neighbourhood is shown via gray lines.

- Both the grid topology (Fig. 1a) and the torus topology (Fig. 1b) consist of a regular distribution of link neighbours on a lattice (“large world”). The torus topology includes periodic boundary conditions.
- The *scale-free network* (Fig. 1c) exhibits a distribution of degrees (i.e., number of links for each node) that follows a power law and is generated using the preferential-attachment algorithm by [1].
- The *cave-people topology* (Fig. 1d) is a version of the caveman networks [31] but with less symmetry. The algorithm uses a parameter to define the cluster size c_1 and the number of clusters c_2 . The probability to have a link between individuals of the same cluster is 50%, and the probability of ties between clusters is $4/c_2$.
- The spatial-proximity topology (Fig. 1e) depicts networks with a high clustering based on spatial proximity and was introduced in [41] to model spreading dynamics of epidemics (SIR model).
- The *small-world topology* (Fig. 1f) is based on the Kleinberg model [20], using a lattice topology and a number of long-range links, added to the network, leading to a shorter average path-length on the network. When adding long-range links, the probability of connecting two random nodes is proportional to $1/d^q$ with q being the clustering coefficient and d the distance of the nodes.

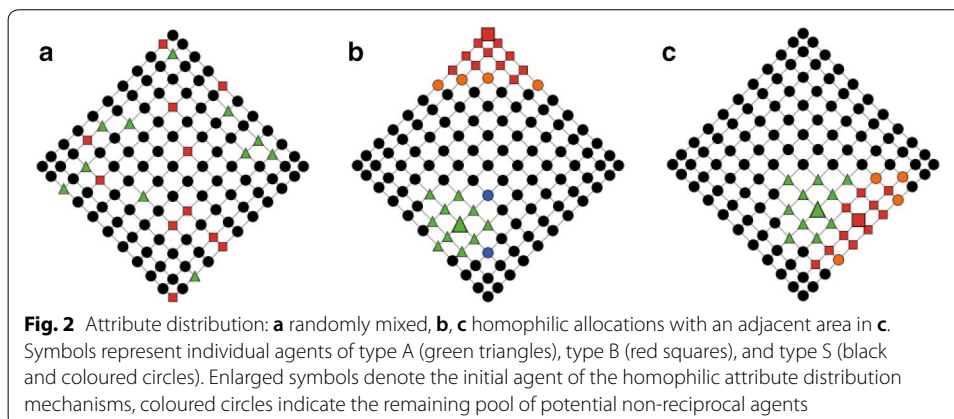


Attribute distributions

The population consists of three different types of agents. The minority of the individuals in the system (10%) are of the non-reciprocal type A or type B, with the identical population shares $N_A = N_B = 5\%$. The majority of the population consists of reciprocal individuals type S with a population share $N_S = 90\%$. The allocation gives the proportion of bonding (links between similar individuals) and bridging (links between different individuals), ultimately shaping the contagion dynamics. In random distributions, the average neighbourhood of all individuals is only influenced by the population shares. Homophilic distributions result in highly self-similar link neighbourhoods of each individual. Figure 2 shows the allocation mechanisms for different distributions of type A (green rectangles) and type B (red squares) and type S (black and coloured circles) on the network. Random positioning on the network results in mixed attribute distribution (Fig. 2a) which generally leads to low bonding and high bridging in the network. Ordered distributions (Fig. 2b, c) are given by homophilic allocations, generally leading to high bonding and low bridging.

Homophilic attribute distribution mechanism

To create homophilic attribute distributions, we use a mechanism to generate highly ordered allocations, which operates as follows: first, two random nodes are chosen. One of them is transformed into a type A node, and the other one into a type B node. All the other nodes do not have any type at this stage. Second, all link neighbours of type A and type B that do not have a type assigned to them yet are selected. The selected nodes form a 'pool' of potential type A and type B nodes, respectively. In each step, a random node from the pool of potential type As is transformed into type A. Simultaneously, this process is done for the pool of potential type Bs, so that one type A and one type B are added in each step. This process is repeated until the desired number of types A and B is reached. When a pool becomes empty and further transformations are required, a new pool is created consisting of the type-less link neighbours of all already transformed nodes of type A or type B. All nodes which are not transformed into type A and type B are considered as reciprocal type S nodes. For the rare cases in which one type hinders the growth of the other type completely, such that the final number of N_A or N_B cannot



be reached, the procedure is cancelled and the initial nodes of type A and type B are re-selected.

Figure 2b shows this process on the grid topology for $N_A = 11$ (green rectangles) and $N_B = 11$ (red squares). The algorithm starts with the enlarged nodes and progresses to include the next proximate nodes of the neighbourhoods. Blue circles and orange circles mark nodes of the pools of potential candidates, which have not been selected to transform. Figure 2c shows the mechanism when type A and type B are in close proximity. Here, the pools of potential As and Bs intersect, such that the trait distribution evolves into deformed regions of type A and type B. However, in large populations, these cases are rarely observed for most of the used topology types, with the exception of scale-free networks.

Intermediate homophilic distributions

Social ties of populations are not necessarily static, but often dynamic. Who interacts with whom can change over time, leading to constant updates of the network structure [12, 46]. It has been shown that dynamic social networks can promote cooperation [37] and that adaptive networks have important consequences for the spreading of diseases [14].

Attribute distribution in real-world examples with social contagion typically displays intermediate states of mixed and homophilic allocations. Diffusion is known to be amplified by bridging ties, which link two otherwise unconnected network clusters [26, 45], weak ties [13], referring to less frequent interactions, and long ties [5], connecting socially distant locations. These notations are interchangeable to a certain degree. Structural changes associated with bridging can dramatically accelerate the spread of disease, the diffusion of job information, the adoption of new technologies, and the coordination of collective action [5].

Another aspect influencing diffusion in a societal context is relocation, such as student exchange and university enrolment [40], and migration, influencing the evolution of norms [29]. Leaving a familiar environment to replace it with a new neighbourhood introduces drastic changes to the network, both at the point of origin, as well as at the destination point and is thus of great interest when investigating attribute distribution effects in populations.

To capture variations in bonding and bridging, we introduce three gradual alterations mechanisms of the structural proximity between individuals. These mechanisms relate to the phenomena of adjusting of social ties, long-range interactions, and exchange of the societal environment. These random and target-oriented changes in the network have been implemented to test the robustness of diffusion effects under the homophilic attribute treatment.

Dynamic rewiring

To perform a dynamic analysis of the network, we adjust the social ties between individuals by a similar approach as presented in [38], but replacing the need for a satisfaction level and fitness with a random choice of individuals, keeping the rewiring dynamics as generic as possible. The adjustment of ties between an individual i and an individual j is done by removal of their link followed by rewiring of i with a random chosen link

neighbour of j . The number of adjusted ties is given by R . An illustration of the rewiring of $R = 50$ is shown in Fig. 3a. Correlated with the adjustments, measuring the level of homophily can be done by calculating the mean of $k_{i(x)}/k_i$ for each individual i with the link number $k_{i(x)}$ of its neighbours of identical type X and link degree k .

In real social networks, individuals are able to leave the system and new ones are able to join. However, in our investigation, the number of individuals of certain types needs to be kept constant, so that different simulation runs can be compared. This means every time a individual of a certain type leaves the system, a new one needs to enter it. Since it does not enter at the same position, the new node might have different links, but the same type. Therefore, this process can also be approximated by dynamic rewiring.

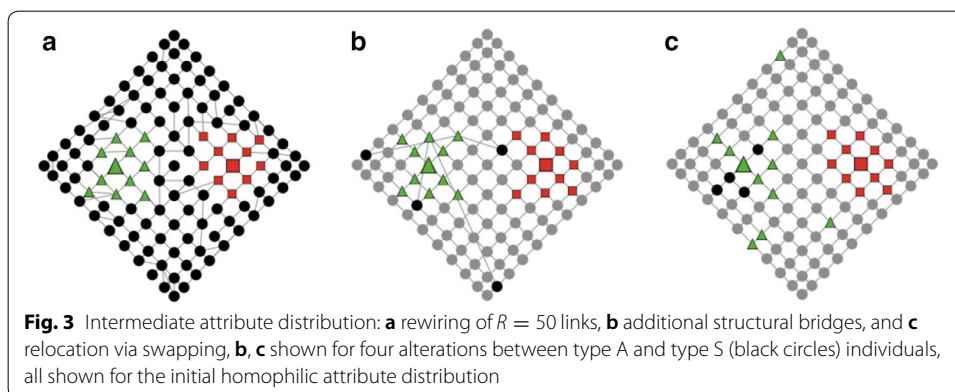
Structural bridges and relocation

We increase bridging by adding long ties between non-reciprocative and reciprocal individuals. For illustration, Fig. 3b shows the homophilic attribute distribution with four long-distant links between randomly chosen type A and type S individuals (bridged type S nodes are highlighted as black circles). Furthermore, we perform positional swaps between two randomly chosen individuals, one non-reciprocative and the other of reciprocal type, to capture reallocations. Figure 3c shows the allocation of four swaps of type A with type S individuals when initially having the homophilic allocation (swapped type S nodes are highlighted as black circles).

For each long tie or swap, both individuals A and S are randomly chosen, with the additional condition that each individual is only allowed to swap once. In general, both mechanisms can be applied to the two non-reciprocative type A or type B, leading to a smooth transition from homophilic to mixed allocations when increasing the number of alterations. Since we are particularly interested in changes which foster the promotion of a single behaviour in the population, we limit additional long ties and swapping to type A and type S individuals (target-oriented alterations) while leaving type B unaltered.

Model

The model consists of N software agents of three populations shares N_A, N_B, N_S . Each agent has an internal behaviour state, which is reflected by the decision variable d . Non-reciprocative agents of type A have a fixed behaviour given by $d = 1$, and type B's behaviour is given by $d = 0$. Since we consider repeated decision making, type S individuals i



decide on their behaviour $d_i \in [0, 1]$ in each round t . Their decision is based on a best-response mechanisms: Their decision variable d is the mean value of all link-neighbours' decision variables:

$$d_i = 1/k \sum_{n=1}^k d_n, \quad (1)$$

with k being the node's degree and n referring to individual link neighbours of the node i .

The resulting mean behavioural state of the population share of type S $\hat{d}(t) = 1/N_S \sum_j d_j(t)$ gives an overall measurement of the tendency of susceptible individuals. The standard deviation of the behavioural state

$$h(t) = \sqrt{\sum_j (d_j - \hat{d}(t))^2 / (N_S - 1)} \quad (2)$$

holds information on the heterogeneity of decisions in the population. After a certain time $t = T$, an equilibrium state is reached leading to the overall behavioural state $\bar{d} = \hat{d}(T)$ and $\bar{h} = h(T)$.

Simulations

We investigate the diffusion process of the decision variable d_j in the population in all networks detailed above for both random and homophilic allocations. We then use rewiring, structural bridges (long ties) and relocation (swapping) to statistically investigate the effect of topological changes on the overall population state. The numerical details of our simulations are as follows: we use a population of $N = 400$ with $N_A = N_B = 20$ and $N_S = 360$. All results were obtained for time steps $T = 2000$. Lattice-based topologies are sized 20×20 . The cave-people networks use 20 clusters consisting of 20 individuals each. The average node degree of the spatial-proximity topology is $\hat{k} = 6$. The small-world topology uses the optimal clustering exponent $q = 2$ of the Kleinberg model [21]. Statistical analysis is based on $s = 500$ simulation runs for each set of parameters. Additional investigations on larger populations have been performed for $N = 800, 1600, 2500$.

We use Netlogo 6.0.4 and the network extension package for simulations, and Python 3.6 to run and evaluate NetLogo-based data using the pyNetLogo library [23].

Results

Our results focus on the collective state of the population, given by the average decision \bar{d} of type S individuals when an equilibrium state is reached. In addition, we observe the heterogeneity \bar{h} of the collective decision in the reciprocative population. In addition to statistically averaged observations, we explore single simulation examples to highlight some relevant cases and to provide deeper insights into differences in attribute distribution depending on the topology type. First, we show results on the dependency of allocations in regard to different networks “[Effects of attribute distribution on various network types](#)”. This is followed by our investigation on intermediate attribute distribution by

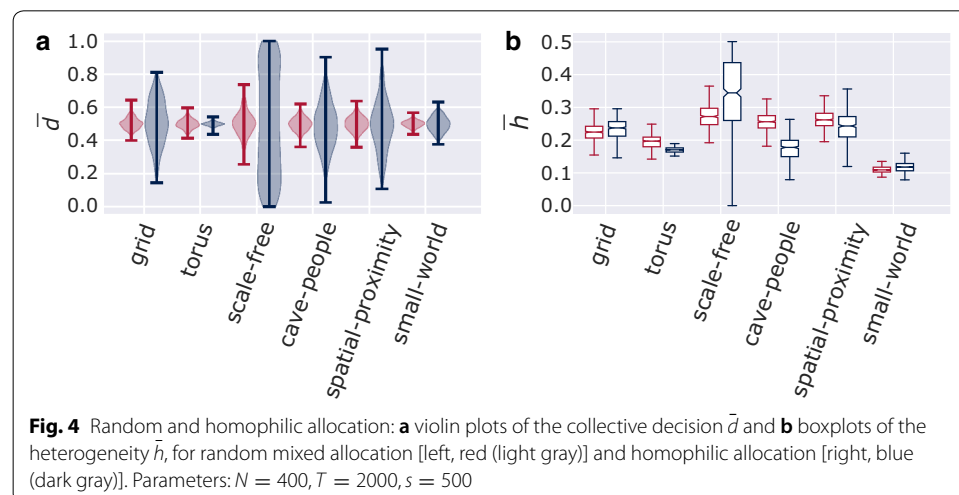
presenting effects of rewiring “[Rewiring](#)” of effects of additional long ties “[Additional long ties](#)” and positional swapping “[Relocation](#)”.

Effects of attribute distribution on various network types

Performing a statistical analysis of $s = 500$ simulations for each of the six topologies, the collective decision \bar{d} is distributed around $D = 1/s \sum \bar{d} \simeq 0.5$ for both, the randomly mixed and homophilic allocations, as shown in Fig. 4a as violin plots. Since $N_A = N_B$, this result is naturally occurring, however, the probability density depends strongly on the network type and allocation. Figure 4a shows different topologies on the x -axis and the colour code refers to randomly mixed allocations [red (light gray)] and homophilic allocations [dark blue (dark gray)].

Random allocation exhibits similar distribution ranges of \bar{d} on all topologies, with the narrowest distribution for small-world topologies and the widest for scale-free networks. In contrast, homophilic allocations lead to a great variation in distribution ranges with a strong dependence on the network structure. Here, torus topologies show the smallest statistical variations due to the periodic boundary conditions. Collective decision on scale-free networks shows a broad distribution, covering the complete range between type A or type B decisions. For most network types, an increase of the probability range of \bar{d} has been observed for the homophilic allocation compared to randomly mixed, being especially pronounced for the scale-free and cave-people topologies. The only exception, showing a decrease, is the torus network.

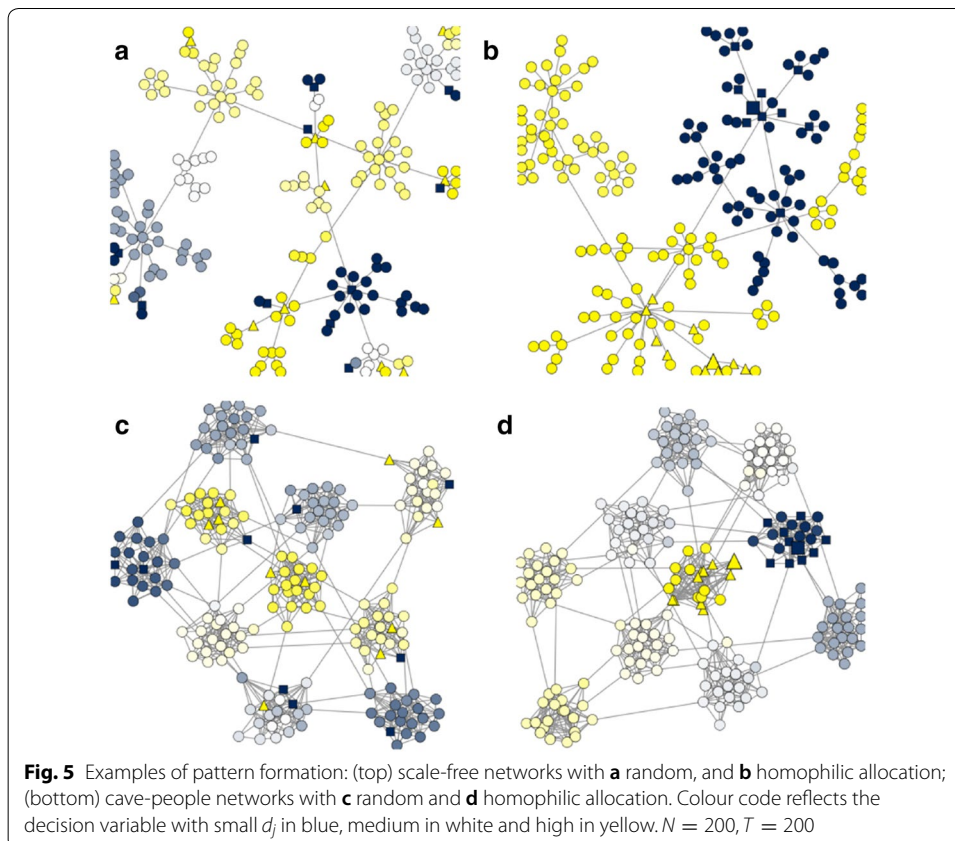
Figure 4b shows the heterogeneity \bar{h} of $s = 500$ simulations as boxplots. For both allocation types, the network structure affects the heterogeneity of decisions significantly, with the mean heterogeneity $H = 1/s \sum h$ being highest for scale-free networks (random: $H \simeq 0.27$, and homophilic: $H \simeq 0.33$) and lowest for small-world topologies (random: $H \simeq 0.11$, and homophilic: $H \simeq 0.11$). Effects of random and homophilic allocations are diverse: Two topologies show a decrease in mean heterogeneity (torus, cave people), three network types show minor changes (grid, spatial proximity, and small world) and one network shows an increase in H (scale free) as well a strong decrease in correlated statistical distribution under homophilic allocation.



The results of Fig. 4 have been tested for larger populations $N = 800, 1600, 2500$ and matching number of non-reciprocative agents $N_A = N_B = 5\%$ and prolonging the simulation time $T = 5000, 1000, 1500$ accordingly. For all topology types, the statistical results on collective behaviour \bar{d} and \bar{h} have been replicable.

Pattern formation

To highlight different effects of allocation on scale-free and cave-people networks, Fig. 5 provides four examples networks with $N = 200$. The colour code from blue over white to yellow represents the decision $d_j \in [0, 1]$ at $T = 200$, with small values of d_j in blue and high values in yellow. Comparing the random allocation Fig. 5a, c with the homophilic allocation Fig. 5b, d on both network types, several characteristics can be noted. Branches of a single scale-free network may hold different tendencies in d_j , resulting in a rather heterogeneous behaviour of the overall populations. While both allocations show this feature, using homophilic allocations increases this phenomenon since the formation of encapsulated branches, only influenced by one type of non-reciprocal type A or B, is more likely. Figure 5c shows a random allocation on the cave-people topology, leading to a dispersion of type A, B agents through the majority of clusters. Figure 5d shows the homophilic allocation, where the influence of non-reciprocators is rather localized due to the accumulation in a single cluster. Here, the influence on clusters of solely type S individuals is weak, such that most clusters show intermediate behaviour of small



heterogeneity. Further examples of the grid, torus, spatial-proximity, and small-world topologies are provided in “Appendix” Figs. 11, 12, 13, and 14.

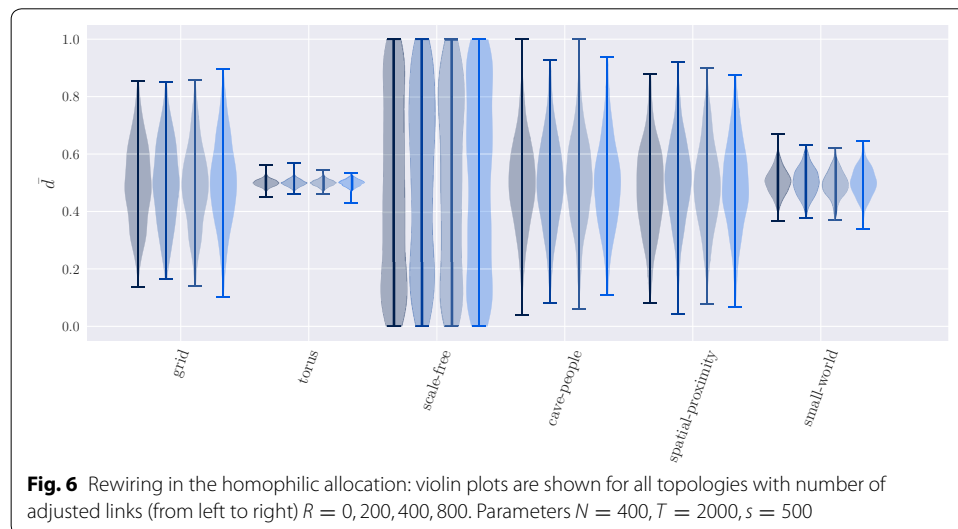
Rewiring

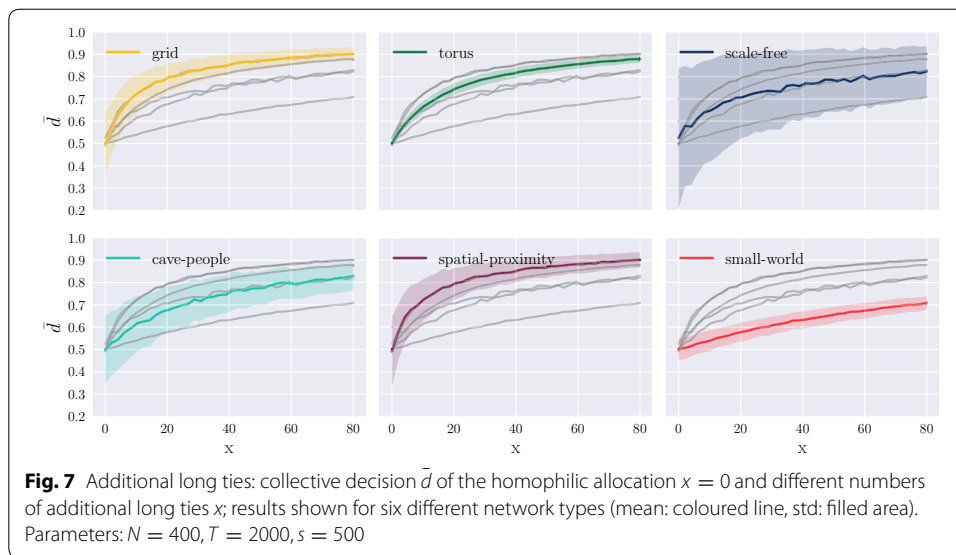
To test the results on the homophilic attribute distribution of Fig 4 [left, blue (dark gray) violin plots], three levels of rewiring have been evaluated, as shown in Fig. 6. Dynamic rewiring of the topologies has been investigated for $R = 0, 200, 400, 800$ with $R = 0$ coinciding with the results of Fig. 4. No significant changes in the statistical dispersion are observed. Moreover, we tested the level of homophily over different time lines of the evolution of the collective behaviour. We observed no significant changes over time for the collective level of homophily. The level of homophily of the reciprocative type C showed a very slight decrease which appeared to be negligible. We conclude that the effects of homophilic allocation on statistical properties of collective behaviour are robust towards a reasonable number of adjustments of social ties.

Additional long ties

To test the effect of long-tie interaction and promotion of the behaviour of type A, alterations of the interaction options are introduced via additional links. First, the basic network topology as introduced in “Attribute distributions” and the homophilic allocation is generated. Afterwards, additional long ties are added as depicted in “Structural bridges and relocation” and permanently placed prior to the start of the simulation run. The number of alterations x , corresponding to the number of additional links, is varied for different simulations runs.

To analyse the impact of structural bridging, $s = 500$ simulations have been evaluated for every $x = 0, 2, 4, \dots, 78, 80$. The results for the six different network types are shown in Fig. 7, where each plot shows one highlighted result (coloured, including the standard deviation as filled area) to the corresponding topology noted in the legend. The reinforcement of type A strategies ($d = 1$) via A–S long ties is clearly visible for all topologies by a decrease of \bar{d} when raising x . While all topologies show





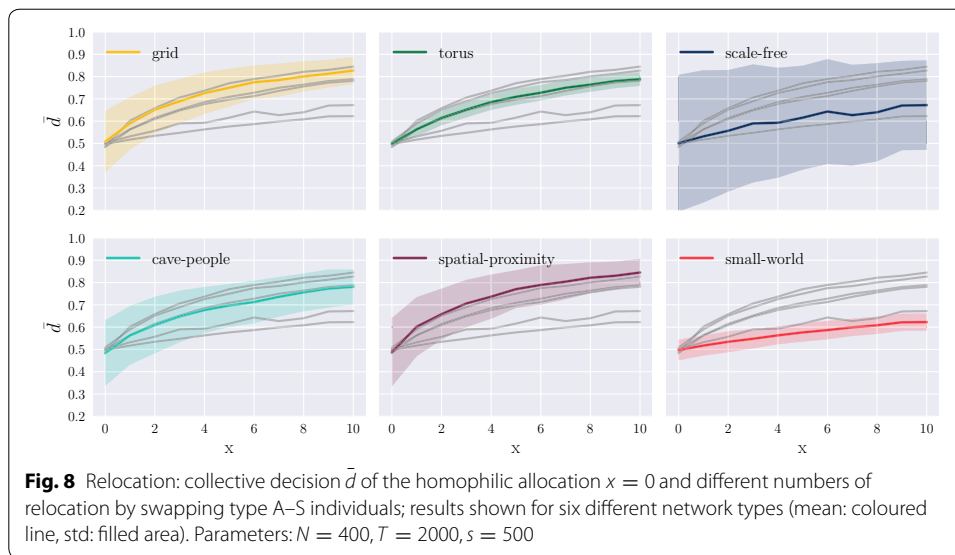
a logarithmic increase, small-world networks show a near linear correlation and the overall increase is the smallest. In case of scale-free networks, the standard deviation (filled area) is the largest compared to other network types for all x and decreases the strongest the more links are added.

Regarding the heterogeneity under variation of x , we observe no significant effects on \bar{h} for three topologies (torus, cave people, and small world) and linear decreases for three topologies (grid, scale free, and spatial proximity) when using different numbers additional ties (see “Appendix”: Fig. 9).

Relocation

An alternative way to analyse intermediate allocations is to use relocation of single individuals. This method allows for a stepwise alterations by varying the number of swaps but of lesser number than additional links. The number of maximally possible alterations is half of the type A population share. The procedure, as introduced in “Structural bridges and relocation”, is applied after generating the network and homophilic allocation, and before the start of the simulation run.

Using individual swaps to blend the ordering of the homophilic allocation, we vary the number of relocations $x = 0, 1, \dots, 9, 10$ and generate $s = 500$ simulations for each network type. Result for the six different topologies is shown in Fig. 8, with highlighting one result (mean value coloured line, standard deviation filled area) as marked in the legend (top left). The tendency towards type A decisions is clearly enhanced for all topologies, being logarithmic for most cases, except scale-free and small-world networks, which show a linear correlation. Similar to our observations of long ties, scale-free networks exhibit the largest standard deviation and no significant effects on the heterogeneity \bar{h} have been observed for variation in x (see “Appendix”: Fig. 10).



Discussion

Diffusion processes in populations are governed by various factors, shaping the spreading of behaviour, traits, or decisions. One such factor is given by the network structure, encompassing network types such as scale-free and small-world topologies. Regarding heterogeneous populations, another important aspect is the allocation of individuals, which strongly influences the direct and unique neighbourhood of each individual. Many approaches that explore ordered allocation on networks combine the network generation with the probability to connect similar individuals, such that the resulting network structure emerges as a function of the behaviour-type proportions.

In this study, we detach the network generation and allocation of individuals, using separate mechanisms: one to generate a specific network type, followed by the procedure to position heterogeneous attributes, referred to as attribute distribution or allocation. To identify effects of allocations, we compare collective behaviour and pattern formation on two contrasting scenarios: random and homophilic allocations. Generally speaking, random allocations lead to a normal distribution of traits in the individual neighbourhoods, and homophilic allocations lead to a higher separation of different types and higher bonding between similar individuals. In addition, the influence of allocations is examined on various network types (lattice, scale-free, cave-people, spatial-proximity, and small-world topologies).

We observe a strong effect of allocation on social contagion. Moreover, the actual impact depends on the network type. A general comparison of random and homophilic allocations shows a tendency towards an increase of possible collective behaviour states of the population for the homophilic allocation on all considered network types, except the torus network. For the latter, the lattice structure with periodic boundary conditions favours the spreading towards uniform behaviour. Since nearest neighbour interactions are local, homophilic allocations are more likely to foster extreme behaviour, often resulting in two population shares exhibiting opposite behaviour. The most prominent example for this observation are scale-free networks. Homophilic allocation on

scale-free networks, which consists of a few highly connected hubs and several branches of lower connected nodes, leads to a strong increase in the distribution range of the collective behaviour as well as the heterogeneity of behaviour within a population.

Extending our investigation to scenarios that emit mixed traits of random and homophilic allocations, stepwise alteration of the population structure is introduced. Here, we compare two options: additional links (long ties) between heterogeneous individuals and mixing of heterogeneous individuals (relocation), both for the initially homophilic population. For both alterations, we find similar responses of the collective behavioural state on all network types. Here, similar progressions of the collective state towards the behaviour, which is stepwise more spread throughout the network, are clearly visible. Moreover, due to the negligible impact of these alterations on the statistical distribution of possible collective states, we conclude that restructuring via additional ties and mixing via relocation have a weaker influence on the pattern formation than the overall allocation.

The presented approach outlines network diffusion in a simple population, categorised in three distinct behaviour types. While this construction serves sufficiently for a first assessment, widening the discussion to more realistic behavioural observations is possible, but exceeds the scope of this primary investigation. Moreover, the allocation mechanism for homophilic allocations can be adapted by incorporating a chance to slightly increase mixing in each group, allowing for softer distributions. In addition, setting the attribute distribution in relation to real data [42] could support possible enhancements to the allocation mechanism. We assume that differences in average degrees and betweenness centrality are crucial for the effects observed on different network types and these specific influences can be further explored.

Conclusion

In summary, we separated network generation and attribute distribution to highlight isolated effects of structural proximity and attribute proximity. We presented a homophilic attribute distribution mechanism and compared the results of random and homophilic attribute distributions using a basic diffusion mechanism on the network. Our main finding is that the effect of attribute distribution is diverse and depends strongly on the network type (structural proximity). A general observation on attribute distribution was that random allocations tend to limit possible collective states in the majority of observed network types. These results indicate that random distribution of attributes used in networked ABM might be limited and of less accuracy for statistical analysis of collective behaviour than expected. We conclude that the homophilic distribution is a substantial feature for improving agent-based modelling and can be easily implemented on various network topologies with the here presented homophilic attribute distribution mechanism.

Acknowledgements

MLK acknowledges financial support of the University of Graz, Austria, and the Steiermärkischen Sparkassen.

Authors' contributions

MLK performed all computational simulations related tasks and is the major contributor in writing the manuscript. MLK and GJ conceptualized the model and interpreted the results. GJ contributed to the introduction, and GJ and MF revised the manuscript. All authors read and approved the final manuscript.

Funding

MLK receives a student Grant of the University of Graz, Austria, and the Steiermärkischen Sparkassen.

Availability of data and materials

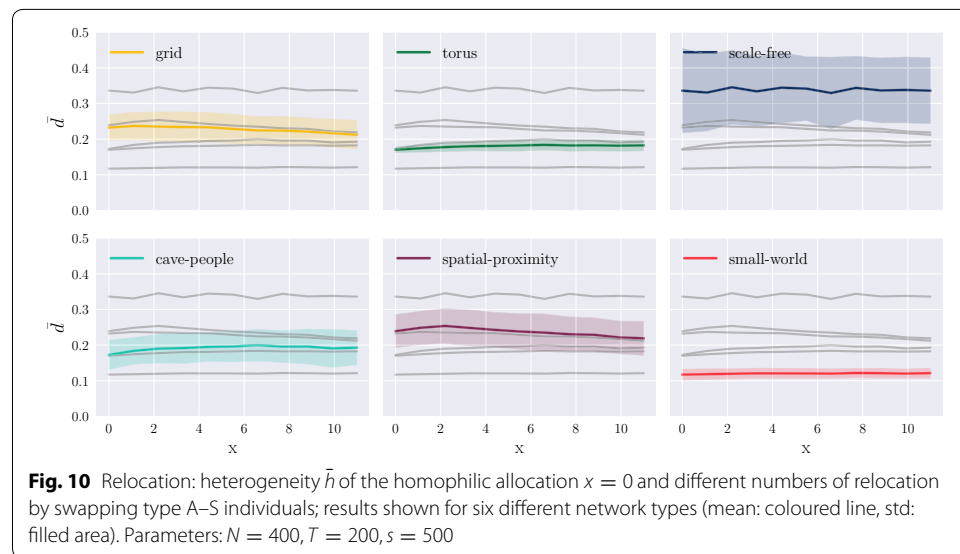
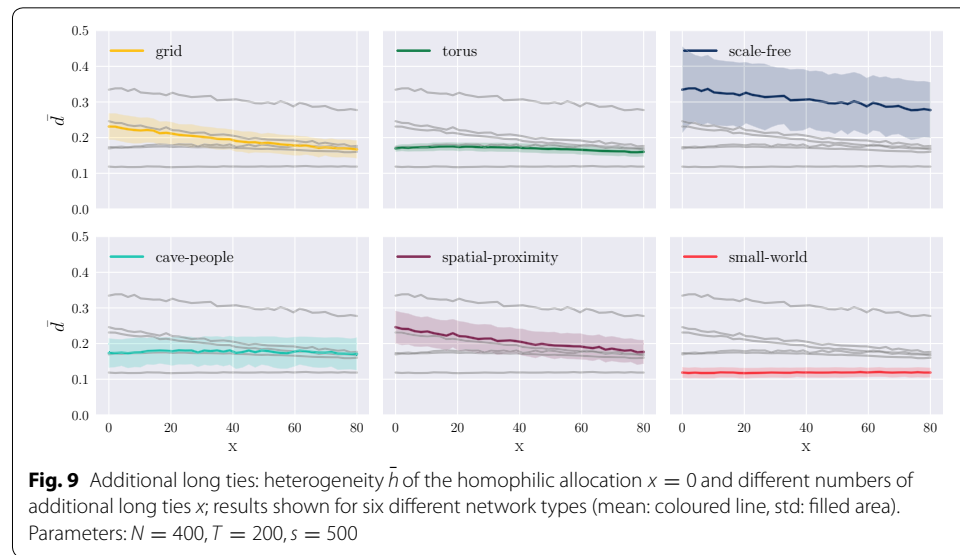
The data sets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Appendix

See Figs. 9, 10, 11, 12, 13 and 14.



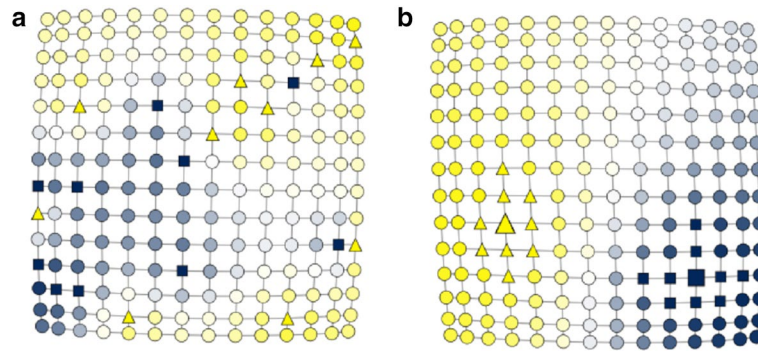


Fig. 11 Examples of pattern formation on grid networks: **a** random allocation, **b** homophilic allocation. Parameters $N = 14 \times 14$, $T = 200$

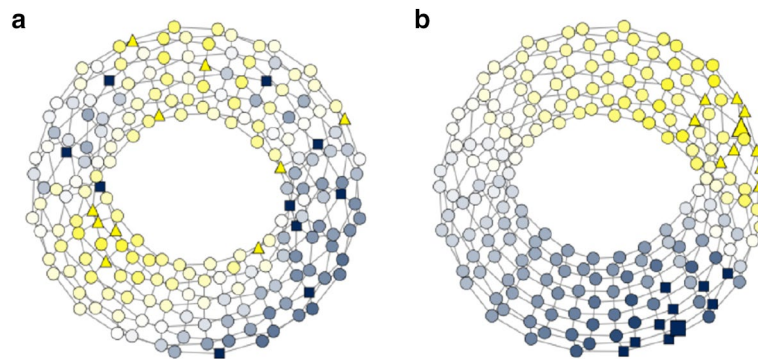


Fig. 12 Examples of pattern formation on torus networks: **a** random allocation, **b** homophilic allocation. Parameters $N = 10 \times 20$, $T = 200$

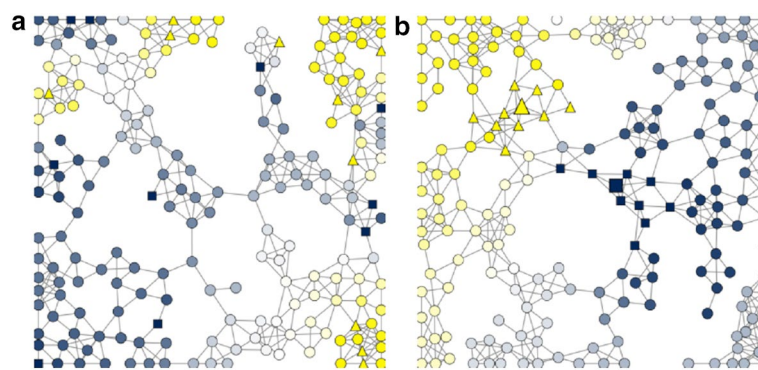
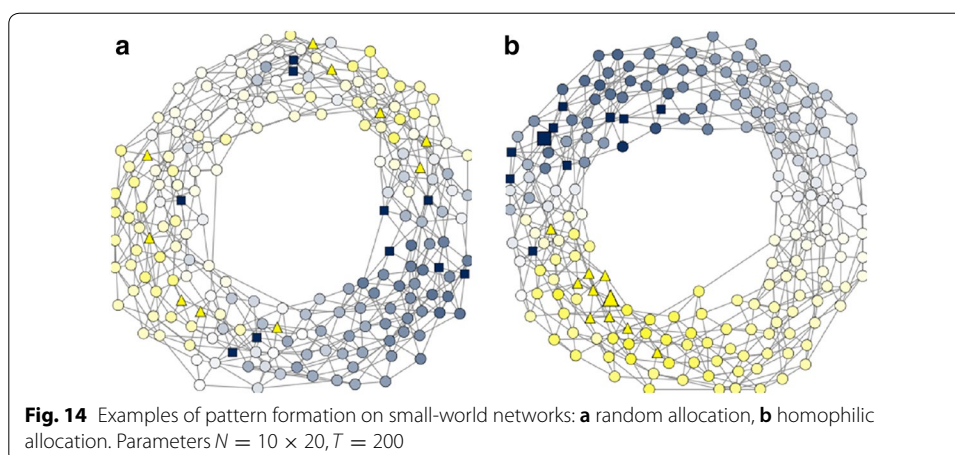


Fig. 13 Examples of pattern formation on spatial-proximity networks: **a** random allocation, **b** homophilic allocation. Parameters $N = 200$, $T = 200$



Received: 8 February 2019 Accepted: 10 October 2019

Published online: 01 November 2019

References

- Barabási AL, Albert R. Emergence of scaling in random networks. *Science*. 1999;286(5439):509–12.
- Borgatti SP, Everett MG. A graph-theoretic perspective on centrality. *Soc Netw*. 2006;28(4):466–84.
- Burlando RM, Guala F. Heterogeneous agents in public goods experiments. *Exp Econ*. 2005;8(1):35–54.
- Cavallari S, Zheng VW, Cai H, Chang KCC, Cambria E. Learning community embedding with community detection and node embedding on graphs. In: *Proceedings of the 2017 ACM on conference on information and knowledge management*. New York: ACM; 2017. p. 377–86.
- Centola D, Macy M. Complex contagions and the weakness of long ties. *Am J Sociol*. 2007;113(3):702–34.
- Chiang YS, Takahashi N. Network homophily and the evolution of the pay-it-forward reciprocity. *PLoS ONE*. 2011;6(12):e29188.
- Choi H, Kim SH, Lee J. Role of network structure and network effects in diffusion of innovations. *Ind Market Manag*. 2010;39(1):170–7.
- Chowdhury NMK, Rahman MR, Boutaba R. Virtual network embedding with coordinated node and link mapping. In: *IEEE INFOCOM*; 2009. p. 783–91.
- Cowan R, Jonard N. Network structure and the diffusion of knowledge. *J Econ Dynam Control*. 2004;28(8):1557–75.
- Delre SA, Jager W, Bijmolt TH, Janssen MA. Will it spread or not? the effects of social influences and network topology on innovation diffusion. *J Prod Innov Manag*. 2010;27(2):267–82.
- Golub B, Jackson MO. How homophily affects learning and diffusion in networks. In: *Technical reports*. 2009.
- Grabowska-Zhang AM, Hinde CA, Garroay CJ, Sheldon BC. Wherever i may roam: social viscosity and kin affiliation in a wild population despite natal dispersal. *Behav Ecol*. 2016;27(4):1263–8.
- Granovetter M. The strength of weak ties: a network theory revisited. *Sociol Theory*. 1983;1:201–33.
- Gross T, D'Lima CJD, Blasius B. Epidemic dynamics on an adaptive network. *Phys Rev Lett*. 2006;96(20):208701.
- Handcock MS, Raftery AE, Tantrum JM. Model-based clustering for social networks. *J R Stat Soc*. 2007;170(2):301–54.
- Holzhauser S, Krebs F, Ernst A. Considering baseline homophily when generating spatial social networks for agent-based modelling. *Comput Math Org Theory*. 2013;19(2):128–50.
- Jackson MO, López-Pintado D. Diffusion and contagion in networks with heterogeneous agents and homophily. *Netw Sci*. 2013;1(1):49–67.
- Karsai M, Kivela M, Pan RK, Kaski K, Kertész J, Barabási AL, Saramäki J. Small but slow world: how network topology and burstiness slow down spreading. *Phys Rev E*. 2011;83(2):025102.
- Kim M, Leskovec J. Multiplicative attribute graph model of real-world networks. *Internet Math*. 2012;8(1–2):113–60.
- Kleinberg J. The small-world phenomenon: an algorithmic perspective. *Technical reports*. Ithaca: Cornell University; 1999.
- Kleinberg JM. Navigation in a small world. *Nature*. 2000;406(6798):845.
- Krause D. Environmental consciousness: an empirical study. *Environ Behav*. 1993;25(1):126–42.
- Kwakkel JH, Jaxa-Rozen M. pynetlogo documentation. 2017. <https://pynetlogo.readthedocs.io/en/latest/>. Accessed 4 Apr 2019.
- Largerion C, Mougél PN, Rabbany R, Zaiane OR. Generating attributed networks with communities. *PLoS ONE*. 2015;10(4):e0122777.

25. Liao L, He X, Zhang H, Chua TS. Attributed social network embedding. *IEEE Trans Knowl Data Eng*. 2018;30(12):2257–70.
26. Liu W, Sidhu A, Beacom AM, Valente TW. Social network theory. In: *The international encyclopedia of media effects*. 2017. p. 1–12.
27. Luhmann CC, Rajaram S. Memory transmission in small groups and large networks: an agent-based model. *Psychol Sci*. 2015;26(12):1909–17.
28. Mazzoli M, Re T, Bertilone R, Maggiora M, Pellegrino J. Agent based rumor spreading in a scale-free network. *arXiv preprint arXiv:180505999*. 2018.
29. McElreath R, Boyd R, Richerson P. Shared norms and the evolution of ethnic markers. *Curr Anthropol*. 2003;44(1):122–30.
30. Nardin LG, Andrighetto G, Conte R, Székely Á, Anzola D, Elsenbroich C, Lotzmann U, Neumann M, Punzo V, Troitzsch KG. Simulating protection rackets: a case study of the sicilian mafia. *Autonom Agents Multi-Agent Syst*. 2016;30(6):1117–47.
31. Neuman EJ, Mizruchi MS. Structure and bias in the network autocorrelation model. *Soc Netw*. 2010;32(4):290–300.
32. Newman M. *Networks*. Oxford: Oxford University Press; 2018.
33. Nishikawa T, Motter AE, Lai YC, Hoppensteadt FC. Heterogeneity in oscillator networks: are smaller worlds easier to synchronize? *Phys Rev Lett*. 2003;91(1):014,101.
34. Norris P. *The bridging and bonding role of online communities*. 2002.
35. Putnam RD. *Bowling alone: America's declining social capital*. Culture and politics. New York: Springer; 2000. p. 223–34.
36. Rahmandad H, Sterman J. Heterogeneity and network structure in the dynamics of diffusion: comparing agent-based and differential equation models. *Manag Sci*. 2008;54(5):998–1014.
37. Rand DG, Arbesman S, Christakis NA. Dynamic social networks promote cooperation in experiments with humans. *Proc Natl Acad Sci*. 2011;108(48):19,193–8.
38. Santos FC, Pacheco JM, Lenaerts T. Cooperation prevails when individuals adjust their social ties. *PLoS Comput Biol*. 2006;2(10):e140.
39. Song H, Boomgaarden HG. Dynamic spirals put to test: an agent-based model of reinforcing spirals between selective exposure, interpersonal networks, and attitude polarization. *J Commun*. 2017;67(2):256–81.
40. Stangor C, Jonas K, Stroebe W, Hewstone M. Influence of student exchange on national stereotypes, attitudes and perceived group variability. *Eur J Soc Psychol*. 1996;26(4):663–75.
41. Stonedahl F, Wilensky U. Netlogo virus on a network model. 2008. <http://ccl.northwestern.edu/netlogo/models/VirusonaNetwork>. Accessed 4 Apr 2019.
42. Thiriot S, Kant JD. Generate country-scale networks of interaction from scattered statistics. In: *The fifth conference of the European social simulation association, Brescia, Italy*. 2008. p. 240.
43. Ugander J, Backstrom L, Marlow C, Kleinberg J. Structural diversity in social contagion. *Proc Natl Acad Sci*. 2012;2011:16502.
44. Valente TW. Network models of the diffusion of innovations. *Comput Math Org Theory*. 1996;2(2):163–4.
45. Valente TW, Fujimoto K. Bridging: locating critical connectors in a network. *Soc Netw*. 2010;32(3):212–20.
46. Van Rooy D. A connectionist abm of social categorization processes. *Adv Complex Syst*. 2012;15(06):1250,077.
47. Wang X, Cui P, Wang J, Pei J, Zhu W, Yang S. Community preserving network embedding. In: *Thirty-first AAAI conference on artificial intelligence*. 2017.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)