Social Networks 76 (2024) 12–21

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Diffusion in small worlds with homophily and social reinforcement: A theoretical model

Elena M. Tur [a,b,*], Paolo Zeppini [c,d], Koen Frenken [e]

[a] School of Innovation Sciences, Eindhoven University of Technology, The Netherlands
[b] Institute of Innovation and Entrepreneurship, School of Business, Law and Economics, University of Gothenburg, Sweden
[c] Université Côte d'Azur, CNRS, GREDEG, France
[d] Department of Economics, University of Bath, United Kingdom
[e] Copernicus Institute of Sustainable Development, Utrecht University, The Netherlands

ARTICLE INFO

Keywords:
Diffusion
Word-of-mouth
Percolation
Homophily
Social reinforcement
Small worlds

ABSTRACT

We introduce homophily in a percolation model of word-of-mouth diffusion in social networks by reorganizing the nodes according to similarity in preferences for adoption of an innovation. Such preferences are described by a "minimum utility requirement" for an agent to adopt. We show that homophily removes the non-linear relation between preferences and diffusion in the standard percolation model with a high diffusion regime ("hit") and a low diffusion regime ("flop"). Instead, in a model with perfect homophily, the final diffusion scales linearly with individual preferences: all agents who are willing to adopt, do adopt the innovation. We also investigate the combined effect of homophily and social reinforcement in diffusion. Results indicate that social reinforcement renders clustered networks more efficient in terms of diffusion size for network with strong homophily, while the opposite is true for networks without homophily. The simple structure of our model allows to disentangle the effect of social influence, homophily and the network structure on diffusion. However, the controllability of the theoretical structure comes at the expenses of the realism of the model. For this, we discuss possible extensions and empirical applications.

1. Introduction

Social networks are generally considered a key channel for the diffusion of innovations – here, new products, services and practices that improve people's lives – to the extent that people's decisions and behaviors are affected by their contacts. Hence, social networks may greatly affect processes of social diffusion that underlie economic development, behavioral change, and human health (Valente, 1996).

Innovation diffusion in social networks can be broken down analytically in two elementary questions. The first question holds: What is the role of one's personal network in deciding about adopting an innovation? If adoption decisions were fully random, innovation diffusion would be a simple probabilistic process of spreading. Instead, one would expect that in many social contexts individuals are influenced by the adoption decisions of their network contacts. In particular, when adoption is costly, risky, or controversial, the willingness to adopt may require "social reinforcement" resulting from the observations that multiple network contacts adopted before. This would explain why some innovations tend to spread better in clustered networks rather than in random networks (Centola and Macy, 2007; Centola, 2010).

This leaves us with the second question: what drives individuals to adopt an innovation beyond influences from their social network contacts? To explain adoption only from social reinforcement would ignore the individual characteristics that affect adoption decisions. Economists traditionally subsume individual characteristics under individual preferences towards an innovation (a product's "demand curve"), which are generally assumed to be uniformly distributed among individuals in a social network (Campbell, 2013). From a sociological view, however, it is well known that individuals with similar individual characteristics tend to be connected much more than dissimilar individuals, a feature that goes under the name of "homophily" (McPherson et al., 2001). Similarity here refers to the personal characteristics that correlate with individual preferences for some innovation.

Homophily in social networks leads to a puzzle in the understanding of social reinforcement: if one observes diffusion only in certain clusters in a network, this may simply be a reflection of homophily in personal characteristics that affect individuals in adopting the innovation regardless of whether their network contacts adopted before (Shalizi and Thomas, 2011). Indeed, in observational studies of innovation

* Corresponding author at: School of Innovation Sciences, Eindhoven University of Technology, The Netherlands.
E-mail address: e.m.mas.tur@tue.nl (E.M. Tur).

https://doi.org/10.1016/j.socnet.2023.05.004

Available online 12 June 2023

diffusion, it has proven quite hard to disentangle the effect of social reinforcement from the effect of homophily (Aral et al., 2009). Against this background, it remains hard to probe the potential of any innovation as its diffusion will depend on the complex interplay between homophily, social reinforcement and network structure, because both homophily in individual preferences and the social reinforcement dynamics result from the structure of the social network. This motivated a turn to experimental studies where homophily and network structure can be manipulated. Such studies are powerful in isolating the effect of either social reinforcement (Centola, 2010) or homophily (Centola, 2011).

We propose a simulation modeling framework that allows to unpack the complexities underlying innovation diffusion in social networks. We extend the standard percolation model of word-of-mouth with three parameters so as to regulate the levels of homophily, social reinforcement and network structure. Note that in percolation models the links between agents are fixed and do not change during the diffusion process, which means that we do not address the question of 'selection versus influence', where selection refers to agents changing their links to other agents depending on homophily. While our model is theoretical rather than empirical, the theoretical nature of the framework allows us to investigate the full three-dimensional parameter space. Where we build on past models to implement social reinforcement (Tur et al., 2018) and network structure (Watts and Strogatz, 1998), the implementation of a homophily parameter posits a new challenge. One possible approach is to moderate the influence of network contacts depending on how similar they are, reflecting a confirmation bias (Konc and Savin, 2019). In such a model, an agent is more likely to adopt if the adopting social contact is similar rather than dissimilar to the agent in question. For such an effect of homophily to take effect, however, one would have to assume that agents can observe not only the adoption decision of their network contact, but also her underlying preference regarding the specific innovation that is diffusing. Instead, one can take homophily into account as a property of the network and its structure, defining homophily as the extent to which agents with similar preferences are clustered together. This means that we do not have to assume that agents can observe the preferences of their network contacts, but model a network in which agents with similar preferences are more often in contact than agents with dissimilar preferences.

After introducing the standard percolation model as our benchmark model In Section 2, we present our extended framework of analysis in Section 3. Here the main focus is on understanding how diffusion depends on the three main factors of the model: homophily, social reinforcement and network structure. We analyze the extent to which an innovation diffuses through a network in different scenarios, using as reference the baseline scenario of a complete network of individuals who adopt only based on their personal characteristics (preferences). This approach lends itself to two main assessments of results in Sections 3 and 4. First, from the perspective of the innovator, the interest is to understand – given a level of homophily, social reinforcement and network structure – how an effort to increase the utility of the innovation will result in larger diffusion. Second, from the perspective of society as a whole, we can evaluate the efficiency of the diffusion process, that is the welfare loss associated with incomplete diffusion. This is caused by information not reaching individuals who would otherwise have adopted the innovation, and amount to an incomplete use of social capital in the network. In Section 5 we conclude with some reflections on how the model can be extended and applied empirically.

2. The benchmark model

2.1. Network structure

We consider social networks, where nodes represent agents and links represent contacts. Henceforth, agents that are linked to a given agent are identified as her 'neighbors'. Social networks may exhibit different link structures with varying degrees of randomness. Social networks may exhibit different structures with varying degrees of randomness. We follow Watts and Strogatz (1998) by focusing on diffusion in small-world networks (see also, Cowan and Jonard, 2004; Delre et al., 2007). We make this choice for two main reasons. First, small-world structures are a common, albeit not a universal, feature of social networks (Dorogovtsev and Mendes, 2013). Second, the small-world model provides a simple yet powerful framework to study the role of important network structures in terms of clustering and average path length while relying on one parameter only.

Scale-free networks constitute another relevant network structure to consider for the study of diffusion in social systems (Albert and Barabasi, 2002). However, the low degree of clustering in scale-free networks render such networks less suitable for the study of how homophily and social reinforcement affect diffusion.

The small-world network algorithm generates a family of graphs, which are an interpolation of regular lattices and completely random networks. The algorithm starts with a regular one-dimensional lattice (a circle where only agents that are geographically next to each other are connected, possibly with two or more links on each side) and rewires every link with some probability.[1] The rewiring probability is expressed by a parameter $\mu \in [0, 1]$ which allows to fine tune the randomness of the network.

Different values of the rewiring probability μ lead to different average path length and different clustering coefficient (Fig. 1). The clustering coefficient of a network is the relative number of triads in a network. For clustered networks, the probability of two agents being connected increases if they have a neighbor in common. The case with $\mu = 0$ is the one-dimensional regular lattice, and the case with $\mu = 1$ is the fully random network. For intermediate values of μ, a small-world network presents at the same time a relatively high clustering together with relatively low path lengths. Around $\mu = 0.01$ this feature is most evident, as the average path length is nearly as low as the one of a random network, while the clustering coefficient is comparable with the one-dimensional regular lattice (Watts and Strogatz, 1998).

2.2. Percolation model of diffusion

We use a percolation model to describe word-of-mouth diffusion in a social network. As a working example, we consider the diffusion of an innovation that can be transmitted and adopted at no cost throughout a population of previously unaware agents. In a pure percolation setting, agents become informed about the existence of the innovation only locally, meaning through the observation of adoption by their neighbors. Consequently, the structure of the social network where agents are embedded in, is a determinant of the outcome of the process (Tur et al., 2018; Qiao et al., 2019). Earlier studies on word-of-mouth diffusion have also considered percolation processes, but only in regular networks as a two-dimensional lattice (Solomon et al., 2000; Hohnisch et al., 2008; Cantono and Silverberg, 2009; Zheng et al., 2013) or in completely random networks (Campbell, 2013).

The percolation model of social diffusion processes consists of agents that are the nodes of a social network, and are heterogeneous in their preferences towards the innovation, or, put differently, their reluctance to adopt the innovation. This model has a clear counterpart in models of epidemic diffusion, where this reluctance corresponds to the resistance of an agent to disease infection. In social percolation, agents are characterized by a *minimum utility requirement* (u) for adopting an innovation. The innovation itself is characterized by an intrinsic

[1] An alternative algorithm proposes to add links instead of rewiring the existing links (Newman and Watts, 1999). This alternative, however, changes the density of the networks, which influences the process of diffusion through word-of-mouth. To ease the comparison between networks, we use the original algorithm. However, we numerically checked that the algorithm did not create disconnected networks that could be a problem for our analyses.

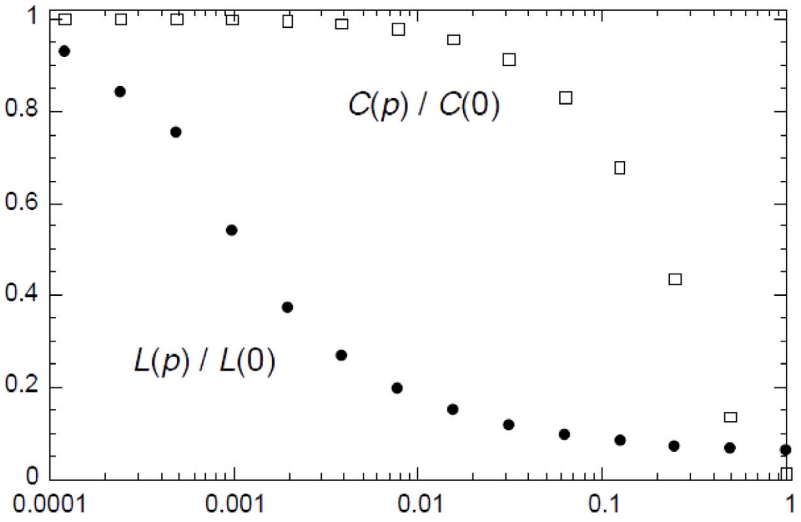


Fig. 1. Clustering coefficient C (white squares) and average path length L (black dots), as a function of the rewiring probability in small-world networks. *Source:* Watts and Strogatz (1998)

utility level, which is represented by a number $\xi \in [0,1]$. The higher the minimum utility requirement u – i.e. the more reluctant an agent is – the higher the utility she requires from an innovation for adopting it. For this study, the minimum utility requirement of agents is a random variable uniformly distributed, $u \sim U[0,1]$.

The theoretical framework just presented corresponds to the so-called social percolation model (Solomon et al., 2000; Flores et al., 2012). The main idea of this model is that adopting agents inform their neighbors, who then decide whether to adopt. In this framework time is discrete, and an agent i adopts the innovation at a given time t if three conditions are satisfied:

1. the agent has not adopted before t,
2. the agent is informed about the innovation (which occurs if at least one neighbor has adopted at time $t-1$),
3. the utility of the innovation is equal or higher than the minimum utility requirement of the agent, that is $u_i \leq \xi$.

Notice that as soon as the intrinsic utility parameter ξ is set to some value, it is already known for which agents it holds that $u_i \leq \xi$. We name them "willing-to-adopt" agents. However, these are only potential adopters. Without a network structure of social contacts, in what is often referred to as a "complete network", there is perfect information about the innovation. When the innovation is introduced, all agents are informed, and all willing-to-adopt agents adopt, while the rest of the population does not. Since u is uniformly distributed as $u_i \sim U[0,1]$, a proportion $100 \cdot \xi\,\%$ of the population on average will adopt an innovation of intrinsic utility $\xi \in [0,1]$. Consequently, in a population of size N, the expected number of agents adopting the innovation $N_{adopters}$ is the expected number of agents with $u \leq \xi$, and is given by

$$N_{adopters} = N \cdot Prob(adoption) = N \cdot Prob(q \leq \xi) = N \cdot \xi \tag{1}$$

Thus, a uniform distribution of individual preferences u_i in a complete network gives a linearly increasing diffusion size as a function of the intrinsic utility.

When agents are embedded in a social network structure instead, and information travels only through social contacts, two different regimes arise in the intrinsic utility space $\xi \in [0,1]$: a *high-diffusion* regime, where the diffusion size is about the same that one obtains in a complete network, and a *low-diffusion* regime, where diffusion is almost absent. These two regions are separated by a critical value of the utility of the innovation, the so-called percolation threshold ξ_c, as the result of a second-order critical transition (Stauffer and Aharony, 1994).

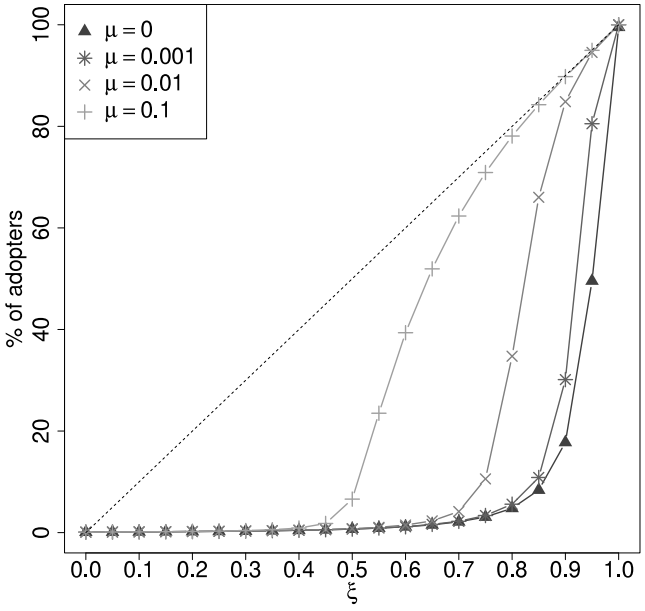


Fig. 2. The percolation benchmark model. Simulated diffusion size for different utility levels of an innovation $\xi \in [0,1]$ (horizontal axis) in different small-world networks with decreasing rewiring $\mu \in [0,1]$ (where $\mu = 0$ is the fully clustered structure of a regular lattice). Values are averages over 500 runs. The network size is $N = 10{,}000$ nodes, with 10 seeds (early adopters).

In this case, as further shown below, the introduction of innovations with utility just above or just below the percolation threshold will have very different outcomes in terms of high or low diffusion, respectively. In product diffusion, this threshold is also known as the fine line separating a "hit" from a "flop" (Solomon et al., 2000).

We analyze the percolation model by means of batch simulation experiments. These are simulations of the diffusion process repeated several times (500 replications in our case) in the same configuration of parameters. For a given set of parameters we compute the average value of the final diffusion size. Here we do batch simulations experiments for the rewiring probability of networks μ and the intrinsic utility of an innovation ξ.

Fig. 2 compares the final percentage of adopters as a function of the intrinsic utility ξ in different small-world networks (Watts and Strogatz, 1998), with rewiring probability $\mu \in \{0, 0.001, 0.01, 0.1\}$. The case $\mu = 0$

corresponds to the regular lattice. All networks have $N = 10{,}000$ nodes, each one representing a potential adopter of the innovation, with $k = 4$ neighbors on average. The u values of agents are random draws from a uniform distribution, $u_i \sim U[0, 1]$. In all simulations the diffusion process is initialized with 10 randomly chosen early adopters, the seeds of the simulation.[2] The dashed line shows the cumulative distribution of the uniformly distributed u values. This is the percentage of willing-to-adopt agents for innovation of utility ξ, as shown in Eq. (1), and would be the final diffusion size in a complete network.

As we see in Fig. 2, the social structure creates "information failures" compared to the complete network with perfect information. Some willing-to-adopt agents never become informed about the existence of the innovation because none of their neighbors have adopted. As a consequence, the final diffusion size is lower than the linear function of Eq. (1) representing full information (dashed line). This information failure undergoes a critical transition, and the diffusion size presents two different phases, or regimes: a low-diffusion phase and a high-diffusion phase. Below the percolation threshold, virtually zero agents adopt, no matter the utility of the innovation. Above the percolation threshold, the number of adopters increases sharply.[3] The percolation threshold is approximately 0.3 for the fully random network, 0.5, 0.7 and 0.8 for the small-worlds with rewiring probabilities 0.1, 0.01 and 0.001, respectively, and again 0.8 for the regular lattice.[4] As the utility of the innovation increases above the threshold, the number of adopters converges to the line representing the cumulative distribution of the minimum utility requirement u of agents, that is the diffusion size of a complete network.

The information failure causing inefficiency in diffusion stems from the network structure of the population: when we compare networks with a fixed number of links, such as the small-world networks generated with different probability of rewiring links, clustering hampers diffusion. Fig. 2 shows that percolation thresholds decrease with μ, and that diffusion sizes increase with μ. Low values of rewiring probabilities μ produce networks with large values of the shortest path length and a large clustering coefficient. In such networks with high clustering, most links are redundant and cannot be used to reach new sources of information.[5] Ultimately, this is a result of an inefficient use of the social capital embedded in social links.

3. Homophily and social reinforcement

3.1. Homophily

There are two main ways to account for homophily in word-of-mouth diffusion. First, one can assume that an agent is most sensitive to the adoption decision of a neighbor who is most similar, putting more trust in the information that comes from sources similar to herself. That is, an agent is more likely to adopt if informed by a similar neighbor, compared to a dissimilar neighbor who has adopted before. This allows to introduce confirmation bias in the diffusion process (Konc and Savin, 2019). In such a case, the network structure may not display any homophily; rather, neighbors that happen to be more similar in a network with otherwise randomly distributed preferences, will exert more influence on each other. For such an effect of homophily to take effect, one would have to assume that agents can observe not only the adoption decision of a neighbor, as in the standard percolation model, but also the neighbor's underlying preference regarding the specific innovation that is diffusing.

Second, one can take homophily into account as a property of the network and its structure, defining homophily as the extent to which agents with similar preferences are clustered together. We follow this second approach, as we then do not have to assume that agents can observe the preferences of their neighbors. Indeed, while one can assume that agents can observe who adopts an innovation in their neighborhood, it is less obvious to assume that agents can also observe the minimum utility level that adopting neighbors hold, as the underlying driver of their adoption decision.

We introduce homophily in the model as a modification of the small-world network algorithm presented in the previous section. Instead of drawing the u values of each agent in the initial regular lattice from a uniform distribution (Fig. 3(a)), we first draw the u values for the whole population and then set them in the regular lattice in an orderly fashion, with values decreasing on either side of the agent with the highest u, up until the agent with the lowest u (Fig. 3(b)). This orderly distribution of preferences represents the fully homophilous one, with all agents having neighbors with very similar preferences. We then consider different degrees of homophily in the network, with the aim of understanding in full detail its effect on diffusion. In particular, we tune the homophily level $\rho \in [0, 1]$ by reordering the preferences of a fully homophilous network. Starting from the configuration of perfect homophily (Fig. 3(b)) we take each node with probability $1 - \rho \in [0, 1]$ and swap it with another chosen at random (Fig. 3(c)). After the reordering of preferences, we apply the parameter μ rewiring a certain fraction of links in the regular network as to create a small-world network (Fig. 3(d)). The whole procedure can be summarized as follows:

- start with the regular lattice (Fig. 3(a))
- place nodes with similar preferences (minimum utility requirement) next to each other (Fig. 3(b))
- swap each node randomly, with probability $1 - \rho$ (Fig. 3(c))
- rewire the links with probability μ (Fig. 3(d))

This procedure thus leads us from the original regular lattice with full clustering and randomly distributed preferences as in Fig. 3(a), to a regular lattice with full clustering and full homophily in Fig. 3(b), to a regular lattice with only a certain degree of homophily (using the ρ-parameter) in Fig. 3(c), and eventually to the final network characterized by a certain degree of homophily and by a certain degree of clustering (using the μ-parameter) in Fig. 3(d). If $\rho = 1$, there is no reordering, which is the perfect homophily scenario just studied. If $\rho = 0$, all nodes are reordered. This case is equivalent to the absence of homophily (as in Fig. 3(a)). Between these two extreme cases, the algorithm provides different degrees of homophily in the social network. Notice that we first tune the homophily of the network using the parameter ρ, and then rewire links using the parameter μ. The latter operation will further affect the homophily of the network as it changes the neighborhood of some of the nodes. This could have consequences for our analyses, since rewiring links further decreases the overall homophily of the network.

In order to check the extent to which homophily is affected by μ, we compute the average difference of preferences in the neighborhood of each node. The average of these differences for all nodes in the network gives an indication of the overall homophily of the network: the higher the average difference, the lower the homophily. Fig. 4 presents the average values over 1000 repetitions of the same network structure, for different values of $\rho \in [0, 1]$.

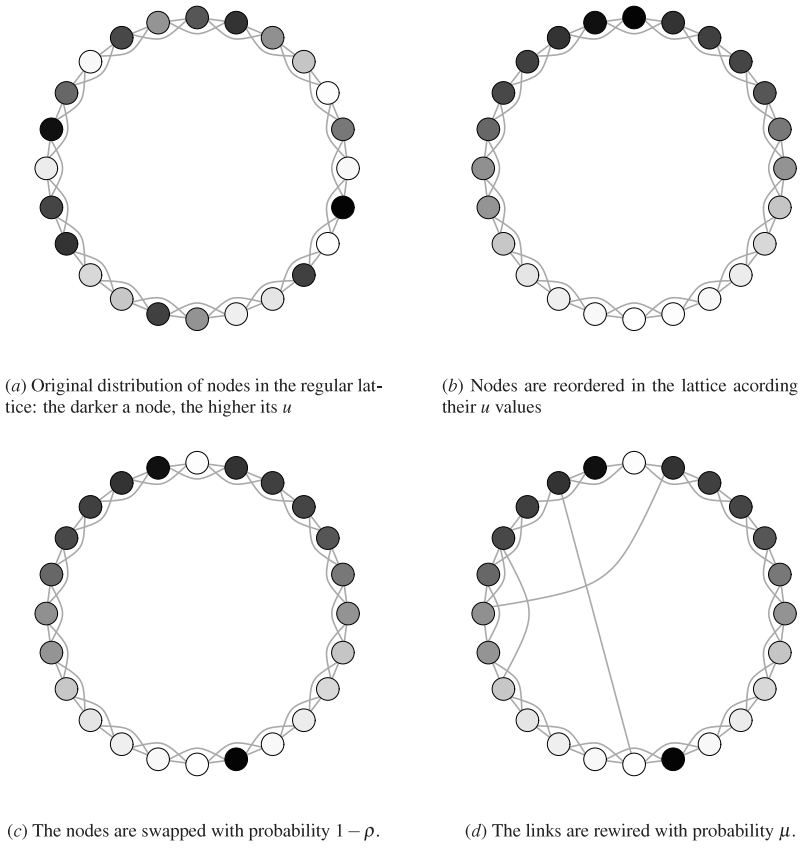
Fig. 4 shows the average difference in preferences among nodes linked to a given node (ego network). Results for $\mu = 0$, 0.001 and 0.01 are virtually indistinguishable. Homophily levels in these three networks are almost identical after rewiring the initial regular lattice, for all values of ρ. It is only when rewiring 10% of the nodes as with

[2] Simulations with a smaller number of seeds (5) and a larger number of seeds (20) yield the same results (available upon request).

[3] This change would be a sharp discontinuity in the case of an infinite population. With finite populations the gap is smoothed.

[4] Newman and Watts (1999) show that the theoretical threshold value for infinite network size is p that solves $\mu = 1 - \frac{(1-p)^2}{4p}$.

[5] Note that information failures would be less pronounced if early adopters, the seeds of simulations, were not random, but were located in nodes with special properties, e.g. nodes with high centrality in the network (targeted seeding).



(*a*) Original distribution of nodes in the regular lattice: the darker a node, the higher its u

(*b*) Nodes are reordered in the lattice acording to their u values

(*c*) The nodes are swapped with probability $1-\rho$.

(*d*) The links are rewired with probability μ.

Fig. 3. Implementing homophily of nodes in small-world networks structures.

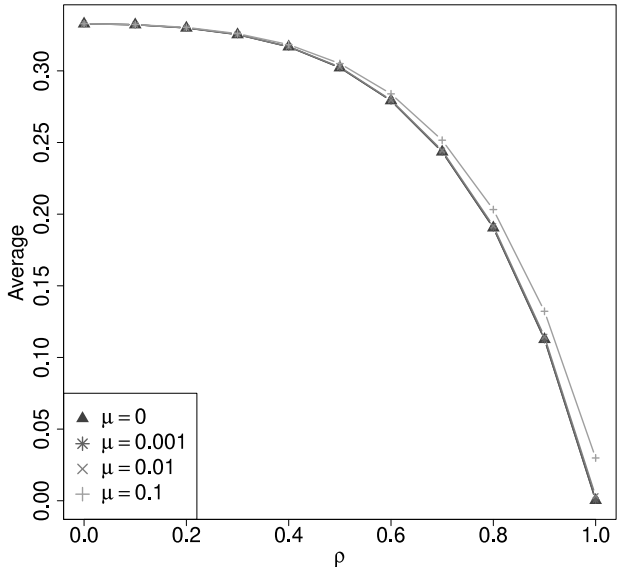


Fig. 4. Effect of swapping nodes (with probability $1-\rho$) on the difference of preferences of the ego networks.

$\mu = 0.1$, that we observe a difference, albeit marginal. For $\rho = 0$, nodes are re-reordered and there is no homophily in the initial regular lattice. In this case, all networks have a comparable average difference of preferences among nodes. As ρ increases, the level of homophily in the initial regular lattice also increases. Intuitively, the rewiring procedure "opens" the neighborhood of an agent, as it gives her access to other agents with different preferences. However, this difference remains small, and the effect of rewiring on homophily is markedly smaller that the effect of reordering.

3.2. Social reinforcement

Social reinforcement is the effect that adoption by neighbors have on an individual's adoption decision (Centola, 2015; Liang, 2021). For some innovations, social reinforcement may be more intense than for other innovations (McMillan et al., 2018), so we introduce social reinforcement as a parameter. In the context of our social network model, following Tur et al. (2018), we introduce social reinforcement in decisions at the local level, by means of a minimum utility requirement u which decreases with the number of adopting neighbors. This effect creates social reinforcement: the more neighbors of an individual i have adopted in the past, the lower the individual minimum utility requirement u_i becomes.

In the benchmark percolation model of diffusion, an agent makes her first adoption decision after the first adoption event in her neighborhood. With social reinforcement this decision is not definitive: an agent may not adopt the innovation the first time she was informed, but she may adopt later, if additional adoptions occur in her neighborhood. In other words, some agents are willing-to-adopt after being informed by a first adopting neighbor, while others need more adopting neighbors to be willing-to-adopt (Berry et al., 2019).

We follow Tur et al. (2018) and model social reinforcement at local level with the following expression of an agent's u, as described in Eq. (2).

$$u_t^i = u_0^i \cdot \left(\frac{1}{a_t^i}\right)^{\gamma} \tag{2}$$

As the number of adopting neighbors a_t^i increases, the minimum utility requirement u_t^i decreases. The parameter γ represents an "intensity" of social reinforcement: for a given number of adopting neighbors, the

updated value of u will be lower for higher γ. Arguably, the larger γ is, the wider is diffusion. For $\gamma = 0$ (no social reinforcement), the updated u coincides with the benchmark case $u_t^i = u_0^i \ \forall t$. As demonstrated in Tur et al. (2018), this function satisfies four conditions: (1) it is decreasing in a_t^i, (2) it is decreasing in γ, (3) with no social reinforcement ($\gamma = 0$) it replicates the benchmark percolation model, and (4) the first consideration of adoption is based on the initial minimum utility requirement u_0^i only. In particular, the first adoption event in a neighborhood is independent on social reinforcement, since when only one neighbor adopts $u_t^i = u_0^i$.

Note that, in a way, this model is a type of stochastic actor-oriented model (Snijders, 1996): the actors of the model are connected by a network, and they have attributes which can change: their minimum utility requirement and whether they have adopted. The only stochastic element in our model, however, lies in the initial conditions. Once the minimum utility requirements are randomly drawn, the network is set, and the seeds are decided, then all the following steps are deterministic.

4. Simulation results

We simulate percolation in different networks with different levels of homophily and social reinforcement, using the same general setting as in the benchmark model: with a population of $N = 10{,}000$ nodes, $k = 4$ neighbors on average, and 10 seeds. Fig. 5 shows the final average percentage of adopters in the different cases. In this figure, each panel presents a different combination of homophily (ρ) and social reinforcement (γ). Each column keeps the level of social reinforcement fixed, and homophily increases as we move down the panels. Each row keeps the level of homophily fixed, and social reinforcement increases as we move right. The top left figure ($\rho = 0, \gamma = 0$) shows the benchmark case as in Fig. 2, while the bottom right panel shows the case with extreme homophily and extreme social reinforcement.

4.1. Homophily

We first focus on the bottom left figure in Fig. 5, the case with extreme homophily but no social reinforcement ($\rho = 1$ and $\gamma = 0$). All homophilous networks present almost identical adoption sizes, including the regular lattice. In other words, the diffusion size does not differ substantially for different values of the rewiring probability μ. Moreover, the diffusion pattern is almost the same as for a complete network, with a linear correspondence between diffusion size and utility of the innovation. This is equivalent to saying that in a homophilous network, the network structure does not play a role in the diffusion process.

The intuition behind the effect of extreme homophily can be explained as follows. In the original regular lattice with a perfect homophily ordering, all agents that are willing-to-adopt the innovation are situated in a same connected component of agents with u higher than the utility of the innovation, the so-called operational network. In a small-world network, a tiny fraction of links are rewired, and it is unlikely that rewiring disconnects the operational network connected component. If at least one early adopter belongs to this component, all agents in this component will eventually be informed and adopt as well. In this case, the number of adopting agents approaches the number of willing-to-adopt agents. Since the utility thresholds of agents follow a uniform distribution on the support $[0, 1]$, the expected number of agents willing-to-adopt an innovation of utility $\xi \in [0, 1]$ is $\xi \cdot N$ (Eq. (1)). Thus, the expected number of adopters increases linearly with the utility of the innovation. That is to say, homophily corrects almost completely the information failure stemming from clustering in networks.

Homophily introduces order in the network. In the limit of perfect homophily, its effect is a "linearisation" of diffusion, as it eliminates the sharp transition between high-diffusion and low-diffusion regimes of the network. Increasing homophily enhances diffusion, and this effect is stronger in more clustered networks (low rewiring μ). In the left column panels of Fig. 5 from the top ($\rho = 0$) to the bottom ($\rho = 1$) we observe how increasing homophily boosts the diffusion size of clustered networks, with the largest effect obtained in regular lattices ($\mu = 0$).

The important implication of results presented above is that homophily corrects the information failure of percolation in clustered networks. Agents are located "at the right place": with high probability, agents with similar preferences are connected to each other. Homophilous networks bring information exactly where it needs to be, providing a more efficient diffusion process. With perfect homophily, once an agent adopts, all willing-to-adopt agents are informed. Only unconnected network components without early adopters are lost, which in a small-world are a negligible fraction.

This result has important practical applications, e.g. in the context of information campaigns and marketing strategies. Since the real structure of a social network is hardly known, the degree of homophily is a highly valuable information: the more homophily among agents connected in a network, the less important the network structure is for diffusion, and the more diffusion size depends proportionally on the utility of the innovation. By contrast, in networks without homophily, there is a critical threshold value for utility, which means that achieving a utility above this threshold is pivotal.

4.2. Adding social reinforcement

Looking at the different columns of Fig. 5 from left to right, we move from a context without social reinforcement ($\gamma = 0$) to a context with strong social reinforcement ($\gamma = 1$). Here, we observe that social reinforcement supports diffusion. The reason is that social reinforcement from neighbors triggers adoption by agents who otherwise, based on their initial individual preference, would not have adopted. The effect of social reinforcement, however, decreases markedly for higher levels of homophily. Social reinforcement has only a small positive effect on diffusion for highly homophilous networks. In homophilous networks agents are clustered in terms of their preferences, and social reinforcement has no effect in a cluster of agents who are all willing-to-adopt anyway, while it can exert little effect in a cluster of unwilling-to-adopt agents.

The most subtle results regarding social reinforcement are obtained for high levels of homophily ($\rho = 0.9$, $\rho = 0.8$). Comparing to the case without homophily this means a larger diffusion size for low utility of the innovation ξ, and lower diffusion size for high utility of the innovation. One implication of this result is that social networks with moderate levels of both clustering and homophily are good terrains for the diffusion of early innovations in their infancy, for which we may expect that utility is still relatively low. When dealing with mature innovations instead, likely characterized by high utility, the positive feedback of social reinforcement boosts diffusion by adding further demand in a less homophilious network. This effect emerges from the interaction of social reinforcement and homophily: without social reinforcement (first column of Fig. 5), homophily always leads to more diffusion. With social reinforcement, two regimes of no-diffusion and full-diffusion emerge, each time around the same critical threshold ξ. The second order phase transition of diffusion on networks disappears, and a first order transition occurs, which separates the two regimes. A first order critical transition between alternative coordination equilibria is typical of processes of social reinforcement (Brock and Durlauf, 2001, 2002).

These results are relevant to all cases of social diffusion, and particularly to initiatives – public or private – that have the objective of promoting the extent of social diffusion: if we assume, as it is natural, that utility of an innovation is proportional to a form of effort (e.g. investment, time, etc.), in a scenario with moderate levels of homophily it does not pay off to put any effort on utility, until a level high enough is reached. Above this sort of threshold, or critical mass,

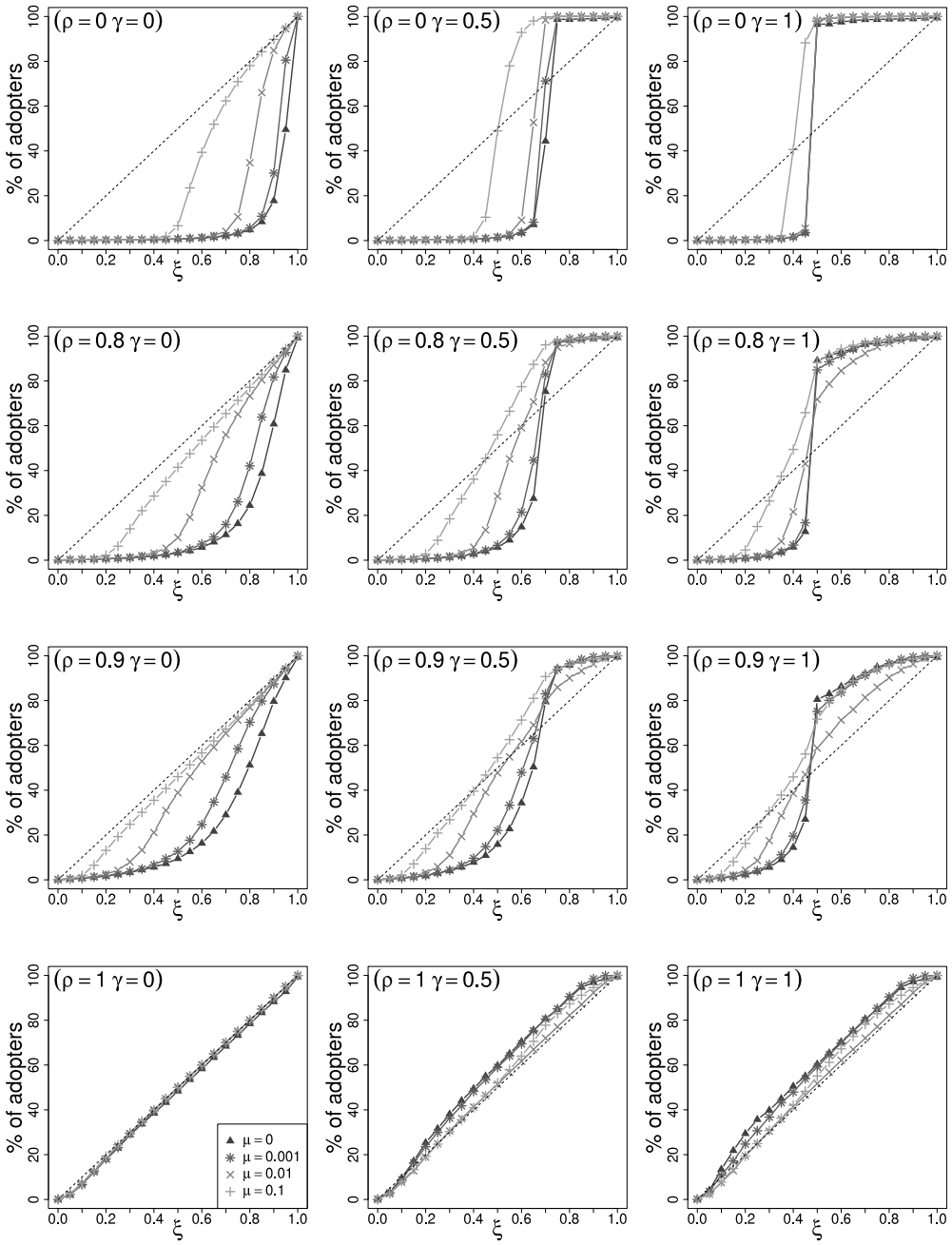


Fig. 5. Diffusion size (vertical axis) for different intrinsic utility levels ξ (horizontal axis) and for different small-world network structures (μ). Results are presented for increasing homophily levels (panels in different rows, where $\rho = 0$ is no homophily, while $\rho = 1$ is perfect homophily) and for increasing social reinforcement intensities (where $\gamma = 0$ is no social reinforcement, and where $\gamma = 1$ is maximum social reinforcement). Reported values are averages over 500 simulation runs. The network size is $N = 10{,}000$ nodes, with 10 early adopters.

the diffusion outcome takes off. Social reinforcement moves this threshold to lower utility levels, making it easier to obtain large diffusion outcomes. Around this threshold the marginal return to investment is largest.

Another interesting result concerns how social reinforcement affects the relative efficiency of different network structures. Less clustered networks remain the most efficient network for diffusion for low levels of homophily, regardless of social reinforcement. However, by looking at diffusion size for high levels of social reinforcement (right panels of Fig. 5) and relatively high levels of homophily, we observe that more clustered networks are relatively more efficient. The results indicate that social reinforcement renders clustered networks more efficient in terms of diffusion size for network with strong homophily. These theoretical results may shed a new light on mixed evidence found on the effect of homophily and of social reinforcement in empirical diffusion studies (for a recent review, see Ertug et al., 2022).

For the sake of comparison, we have also simulated diffusion on undirected scale-free networks with and without homophily and social reinforcement, based on the algorithm by de Almeida et al. (2013). The results are in Fig. A.7 of the Appendix. As we can see, the role of homophily on diffusion size is very marginal. This is most likely due to the fact that scale-free networks have very little clustering, thus preventing cascading effects (Centola, 2010; Jeong and Yu, 2022).

4.3. The non-linear effects of homophily on diffusion

The impact of homophily on diffusion is non-linear. In order to better evaluate such impact, we have calculated the absolute difference

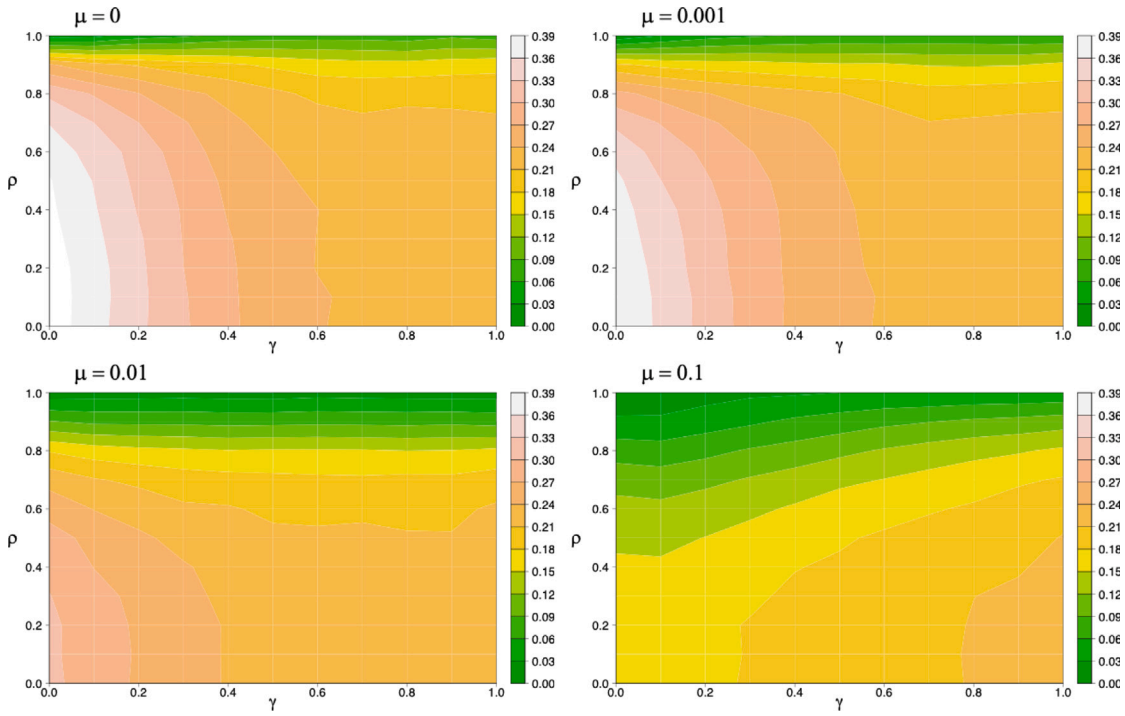


Fig. 6. The marginal effect of homophily: the heat maps represent the absolute difference in simulated diffusion between a small-world network and the diffusion reference level of a complete network. A given color refers to a given difference range with respect to the complete network. For instance, areas in yellow indicate a difference between 3 and 4. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

between the actual diffusion size in small-world networks and the potential level of diffusion that realizes in a complete network (the 45° line), and integrated over utility values. This measure represents inefficiency, or what in economics goes under the name of loss of consumer welfare as it counts the number of agents who would have liked to adopt the innovation, but do not. It amounts to an inefficient use of the social capital contained in the structure of a social network. The heat maps of Fig. 6 report the results. The values of the heat maps represent the area between the diffusion size curve of a given network and the complete network diffusion. For a network with no innovation diffusion this area is maximum, and measures 0.5 (the integral of the 45° line over the range $[0, 1]$).

The impact of changes in the homophily level is very different depending on whether we consider high homophily networks or low homophily. For instance, in a regular lattice ($\mu = 0$) without social reinforcement, an increase in homophily from 0.9 to 1 generates an increment of diffusion size more than twice as large as the increment that we obtain for a change in homophily from 0 to 0.5. The same criticality emerges in cases where lower levels of homophily bring more diffusion. In a small-world with $\mu = 0.01$ and strong social reinforcement, lowering homophily from $\rho = 1$ (perfect ordering of agents) to 0.6 induces a change of diffusion size at least seven times as large as lowering homophily further from 0.6 to 0 (no homophily).

The heat maps analysis also allows to observe an interesting phenomenon about the interplay of social reinforcement and network structure: for a given level of homophily, in the more regular networks ($\mu = 0, 0.001, 0.01$) an increase of social reinforcement induces an ever-smaller departure from the diffusion of a complete network (from white to pink to orange regions). Instead, in the relatively more disordered network ($\mu = 0.1$) the opposite realizes, as increasing social reinforcement makes this departure larger (from yellow to orange). This reversal of effects can be seen by observing the different direction in the pattern of contour lines in the bottom-right panel compared to the other three panels of Fig. 6.

Concluding, for any level of clustering in small-world networks (μ) and any intensity level of social reinforcement (γ), the effect of a spatial ordering of intrinsic preferences as described by homophily in a network has the following non-linear connotation: introducing homophily with ordering more than half the agents of the network has very moderate effects while further spatial ordering of the preferences of agents in the population generates relatively large effects on diffusion.

5. Conclusion

We have proposed a simulation model of diffusion through word-of-mouth in a social network with homophily and social reinforcement. A theoretical understanding of the combined roles of homophily and social reinforcement is important, as it has proven difficult to disentangle their effect on diffusion empirically. Innovations tend to diffuse in clustered parts of social networks, which may equally point to homophily as a driving factor (as agents with similar preferences tend to be connected) or to social reinforcement (as adopting agents may convince connected agents to adopt as well), or a combination of the two. Our study confirms that both homophily and social reinforcement support diffusion, but in very different ways: while homophily in a network makes it possible to fulfill preferences, so that in the limit of perfect homophily all potential adopters do adopt, social reinforcement 'overcomes' preferences, adding adoption for those consumers who initially would not have adopted.

We start from a benchmark model of percolation in small-world networks without homophily nor social reinforcement, where i. all network structures exhibit a percolation threshold separating a low-diffusion and a high-diffusion regime, and ii. the less clustered a network's structure, the more efficient the percolation of information. The latter finding can be understood from the detrimental role of clustering generating redundancy in information flows among agents resulting in an inefficient percolation process.

Compared to the benchmark model, our findings on homophily and social reinforcement can be summarized as follows. Firstly, homophily greatly improves the efficiency of diffusion compared to the benchmark case, especially for clustered networks. Homophily means that agents with similar preferences are connected, leading information to percolate easily to agents who are willing-to-adopt. As a consequence, homophily eliminates the phase transition of diffusion in social networks. Instead, diffusion scales linearly with the utility of the innovation diffusing.

Secondly, social reinforcement also greatly improves the efficiency of diffusion compared to the benchmark case, especially for clustered networks. This can be readily understood as social reinforcement renders agents more willing-to-adopt when more neighbors have adopted before. In contrast to the effect of homophily, diffusion with social reinforcement retains and amplifies the separation between a low-diffusion and high-diffusion regime. In presence of social reinforcement the transition between no-diffusion and diffusion phases is more abrupt, and it occurs at the same critical utility level for all networks.

Thirdly, the "linearization" of diffusion observed for homophilous networks is robust to the introduction of social reinforcement: innovations hardly diffuse more in homophilous networks with social reinforcement than without social reinforcement. This means that social reinforcement has little effect on diffusion among agents that are alike, while it has a strong effect on diffusion among agents that are different. This finding can be understood as a "preaching-to-the-converted" effect: in homophilous networks, agents that are willing-to-adopt are already clustered allowing information to percolate easily, and social reinforcement has little or no additional effect on diffusion.

The theoretical findings of our model can inform empirical studies on diffusion in several ways. First, we have established that both homophily and social reinforcement support diffusion of innovations, but in ways that lead to very different diffusion patterns. Homophily renders diffusion more in line with the utility of an innovation, as agents who are willing-to-adopt will generally be informed in the percolation process. Hence, given a good ex ante estimation of the utility of an innovation – whatever its precise definition or meaning in particular context – one can readily predict how much an innovation will diffuse. By contrast, social reinforcement (without homophily) makes diffusion much less predictable. While innovations with very low utility will certainly hardly diffuse and innovations with very high utility will certainly almost fully diffuse, there is high uncertainty about the diffusion size of innovations with moderate utility. There is a very fine line with a total hit or total flop, and the utility level associated with this threshold is not easy to predict as it depends on the exact network structure as well as the exact level of social reinforcement. Given this theoretical results, divergent findings between empirical studies may be related to different levels of homophily and social reinforcement.

The model also provides insights in strategies that actors can employ that wish to achieve high diffusion of some innovation. For example, authorities may want to reach out to as many people as possible in public health campaigns, and firms may want to market their new product as many potential consumers as possible (Solomon et al., 2000). The theoretical implications of our model for investment strategies, then, are twofold. First, the marginal returns to investments in the utility of an innovation in terms of diffusion size are close to constant in social networks with high degrees of homophily (regardless of social reinforcement). In such settings, a cost–benefit analysis can easily be made. By contrast, investments in an innovation's utility in the context of high social reinforcement (and low homophily) are very risky. One may underinvest in utility leading to little diffusion, thus generating very low return on investment. Reversely, one may over-invest in utility as full diffusion could have been reached already with lower utility. Second, rather than reasoning from a given and fixed social network, investment strategies should take into account the relevant network at play for the particular innovation they want to diffuse, in particular, to what extent the network is homophilous and exert social reinforcement. It may be that regarding certain innovations, agents interact very much with neighbors with similar preferences, while for other innovations they may interact more with neighbors that exert social reinforcement.

The theoretical model can be extended in a number of ways. Starting with the benchmark model of percolation through word-of-mouth, a number of variations can be introduced. First and foremost, the assumption that information about a new innovation is always passed on by adopters and never by non-adopters is arguably a strong one, which can be relaxed such that adoption raises the probability of passing on information, instead of fully determining it. Rather than assuming that adopters always pass on information and non-adopters never pass on information, one can introduce a parameter expressing the likelihood of agents passing on information depending on whether or not they themselves have adopted it. A second promising extension is to have the exertion of social reinforcement depend on the degree of homophily between two agents, as for example in the model by Konc and Savin (2019). This extension would also speak closer to experimental evidence that suggests that similar agents exert greater social reinforcement (Centola, 2011). A third topic for future research is to compare random seeds, as we applied, to targeted seeds common in marketing campaigns modeled as percolation processes (Campbell, 2013). For example, one may assume that organizations promoting an innovation will target agents with high centrality (Valente and Davis, 1999; Valente and Vega Yon, 2020). As a final avenue for research, we foresee applications of the percolation model with homophily and social reinforcement to network structures other than small-world networks, including two-dimensional regular lattices (Solomon et al., 2000) and scale-free networks (Konc and Savin, 2019).

Appendix. Simulations for the scale-free network

See Fig. A.7.

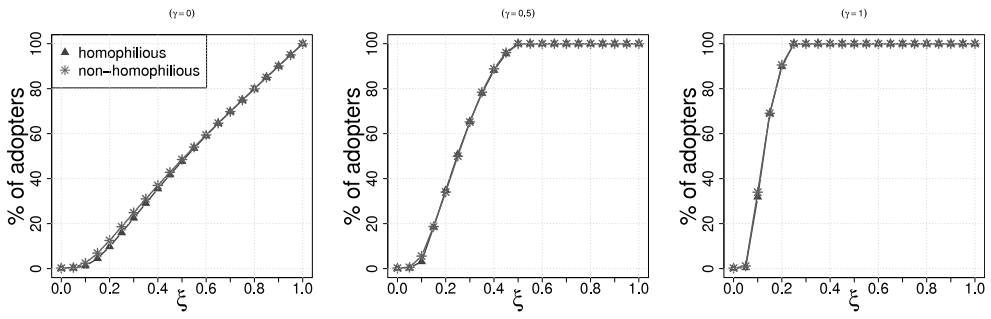


Fig. A.7. Diffusion size (vertical axis) for different intrinsic utility levels ξ (horizontal axis) and for homophilious and non-homophilious scale-free networks. Reported values are averages over 500 simulation runs. The network size is $N = 10{,}000$ nodes, with 10 early adopters.

References

Albert, R., Barabasi, A.L., 2002. Statistical mechanics of complex networks. Rev. Modern Phys. 74 (1), 47–97.

Aral, S., Muchnik, L., Sundararajan, A., 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proc. Natl. Acad. Sci. 106 (51), 21544–21549.

Berry, G., Cameron, C.J., Park, P., Macy, M., 2019. The opacity problem in social contagion. Social Networks 56, 93–101.

Brock, W.A., Durlauf, S.N., 2001. Discrete choice with social interactions. Rev. Econom. Stud. 68 (2), 235–260.

Brock, W.A., Durlauf, S.N., 2002. A multinomial-choice model of neighborhood effects. Amer. Econ. Rev. 92 (2), 298–303.

Campbell, A., 2013. Word-of-mouth communication and percolation in social networks. Amer. Econ. Rev. 103 (6), 2466–2498.

Cantono, S., Silverberg, G., 2009. A percolation model of eco-innovation diffusion: the relationship between diffusion, learning economies and subsidies. Technol. Forecast. Soc. Change 76 (4), 487–496.

Centola, D., 2010. The spread of behavior in an online social network experiment. Science 329 (5996), 1194–1197.

Centola, D., 2011. An experimental study of homophily in the adoption of health behavior. Science 334 (6060), 1269–1272.

Centola, D., 2015. The social origins of networks and diffusion. Am. J. Sociol. 120 (5), 1295–1338.

Centola, D., Macy, M., 2007. Complex contagions and the weakness of long ties. Am. J. Sociol. 113 (3), 702–734.

Cowan, R., Jonard, N., 2004. Network structure and the diffusion of knowledge. J. Econom. Dynam. Control 28 (8), 1557–1575.

de Almeida, M.L., Mendes, G.A., Madras Viswanathan, G., da Silva, L.R., 2013. Scale-free homophilic network. Eur. Phys. J. B 86 (2), 38.

Delre, S.A., Jager, W., Janssen, M.A., 2007. Diffusion dynamics in small-world networks with heterogeneous consumers. Comput. Math. Organ. Theory 13 (2), 185–202.

Dorogovtsev, S., Mendes, J., 2013. Evolution of Networks: From Biological Nets To the Internet and WWW. Oxford University Press, Oxford.

Ertug, G., Brennecke, J., Kovács, B., Zou, T., 2022. What does homophily do? A review of the consequences of homophily. Acad. Manag. Ann. 16 (1), 38–69.

Flores, R., Koster, M., Lindner, I., Molina, E., 2012. Networks and collective action. Social Networks 34 (4), 570–584.

Hohnisch, M., Pittnauer, S., Stauffer, D., 2008. A percolation-based model explaining delayed takeoff in new-product diffusion. Ind. Corp. Chang. 17 (5), 1001–1017.

Jeong, W., Yu, U., 2022. Effects of quadrilateral clustering on complex contagion. Chaos Solitons Fractals 165, 112784.

Konc, T., Savin, I., 2019. Social reinforcement with weighted interactions. Phys. Rev. E 100 (2), 022305.

Liang, H., 2021. Decreasing social contagion effects in diffusion cascades: Modeling message spreading on social media. Telemat. Inform. 62, 101623.

McMillan, C., Felmlee, D., Osgood, D.W., 2018. Peer influence, friend selection, and gender: How network processes shape adolescent smoking, drinking, and delinquency. Social Networks 55, 86–96.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annu. Rev. Sociol. 27 (1), 415–444.

Newman, M., Watts, D., 1999. Scaling and percolation in the small-world network model. Phys. Rev. E 60 (6), 7332–7342.

Qiao, T., Shan, W., Zhang, M., Liu, C., 2019. How to facilitate knowledge diffusion in complex networks: The roles of network structure, knowledge role distribution and selection rule. Int. J. Inf. Manage. 47, 152–167.

Shalizi, C.R., Thomas, A.C., 2011. Homophily and contagion are generically confounded in observational social network studies. Sociol. Methods Res. 40 (2), 211–239.

Snijders, T.A., 1996. Stochastic actor-oriented models for network change. J. Math. Sociol. 21 (1–2), 149–172.

Solomon, S., Weisbuch, G., de Arcangelis, L., Jan, N., Stauffer, D., 2000. Social percolation models. Physica A 277 (1–2), 239–247.

Stauffer, D., Aharony, A., 1994. Introduction to Percolation Theory, second ed. Routledge, London, UK.

Tur, E.M., Zeppini, P., Frenken, K., 2018. Diffusion with social reinforcement: The role of individual preferences. Phys. Rev. E 97 (2), 022302.

Valente, T.W., 1996. Social network thresholds in the diffusion of innovations. Social Networks 18 (1), 69–89.

Valente, T.W., Davis, R.L., 1999. Accelerating the diffusion of innovations using opinion leaders. ANN. Am. Acad. Politi. Soc. Sci. 566 (1), 55–67.

Valente, T.W., Vega Yon, G.G., 2020. Diffusion/contagion processes on social networks. Health Educ. Behav. 47 (2), 235–248.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393 (6684), 440–442.

Zheng, M., Lu, L., Zhao, M., 2013. Spreading in online social networks: the role of social reinforcement. Phys. Rev. E 88 (1), 012818.