The role of edge weights in social networks: modelling structure and dynamics

Riitta Toivonen[a,*], Jussi M. Kumpula[a], Jari Saramäki[a], Jukka-Pekka Onnela[a,b], János Kertész[a,c] and Kimmo Kaski[a] .

[a]Laboratory of Computational Engineering, Helsinki University of Technology, P.O. Box 9203, FIN-02015 HUT, Finland;
[b]Department of Physics, Clarendon Laboratory, University of Oxford, OX 1 3PU, UK;
[c]Department of Theoretical Physics, Budapest University of Technology and Economics, Budapest, Hungary

ABSTRACT

The structure of social networks can influence various dynamic processes of human interaction and communication, such as opinion formation and spreading of information or infectious diseases. To facilitate simulation studies of such processes, we have developed a weighted network model to mimic the structure of real social networks, in particular taking into account the recent observations on weight-topology correlations.[1] The model iterates on a fixed size network, reaching a steady state through processes of weighted local searches, global random attachment, and random deletion of nodes, and it has essentially two parameters which can be used to tune network properties. The generated networks display community structure, with strong internal links and weak links connecting the communities.[2] Similarly to empirical observations, strong ties correlate with overlapping neighborhoods, and under edge removal, the network becomes fragmented faster when weak ties are removed first.[1] As an example of the effects that such structural properties have on dynamic processes, we present preliminary results from studies of social dynamics describing the competition of two non-excluding opinions in a society. Our results show that the weighted community structure slows down the dynamics as compared to randomized reference networks.

Keywords: Social networks, Weighted networks, Community structure, Opinion dynamics

1. INTRODUCTION

Over the past decade or so the structural and functional properties of complex networks have been under intensive study.[3–5] Social networks of human interaction, in particular, have been found to possess several interesting characteristics.[1,6–8] Most prominently, the nodes in such networks are structured into cohesive groups, or "communities". The term community is typically used in the context of groups of vertices with dense internal and sparse external connections, but the exact definitions differ.[9–11] The community structure naturally leads to high values of the average clustering coefficient $\langle c \rangle$, which measures the density of edges within the nearest neighborhoods of vertices. The distribution $p(k)$ of node degrees k (number of neighbors) has been found in many cases to be broad; however, due to obvious limitations, there has to be a cutoff and the maximum degrees cannot be very large. Furthermore, social networks have been found to be assortative, i.e. have positive degree-degree correlations, which in everyday terms means that popular people tend to know other popular people.

The structural properties of social networks are likely to affect their functional properties, such as spreading of information in the network and opinion formation among people. Network models which capture the essential characteristics of interaction networks are necessary for simulation studies of the effects social structure can have on dynamics. Several models have been designed to imitate the characteristics of social networks, based on various mechanisms such as emphasizing connections within neighborhoods,[12,13,16,20] homophily effects,[8] or geographical proximity.[15] It turns out that even models with simple rules can reproduce many topological

∗ e-mail: Riitta.Toivonen@tkk.fi

Noise and Stochastics in Complex Systems and Finance,
edited by János Kertész, Stefan Bornholdt, Rosario N. Mantegna,
Proc. of SPIE Vol. 6601, 66010B, (2007) · 0277-786X/07/$18 · doi: 10.1117/12.725557

characteristics of social networks, as demonstrated by the network growth model with random and nearest-neighborhood attachment, generating tightly connected groups of vertices.[16]

However, most of these models for social networks do not consider link weights, i.e. the strengths of social ties between people. Evidently, tie strengths are important not only to dynamic processes taking place on social networks, but also to the formation of the network structure itself, and in a realistic model tie strengths should play a major role. Earlier, the role of edge weights has been taken into account in e.g. models inspired by empirical data on transport networks.[17,18] There the weight-topology correlations are very different from social networks in that the largest weight links form a skeleton of main highways which carry the flow to be distributed in the denser but weaker local connections. In a social network, the arrangement of link weights is the opposite, i.e., strong links are found within cohesive groups while weaker links connect various communities.[1,2] Here, we discuss a weighted social network model,[19] designed to yield proper weight-topology correlations. The model is based on generating and strengthening connections with weighted local searches. Allowing for reinforcement of connections leads to more pronounced community structure than what is typically obtained by unweighted models. Somewhat similar unweighted processes have been proposed in two earlier models.[20,21] In Section 2 we present the essentials of the model, and in Section 3 we explore the effect of weighted community structure on an opinion formation model.

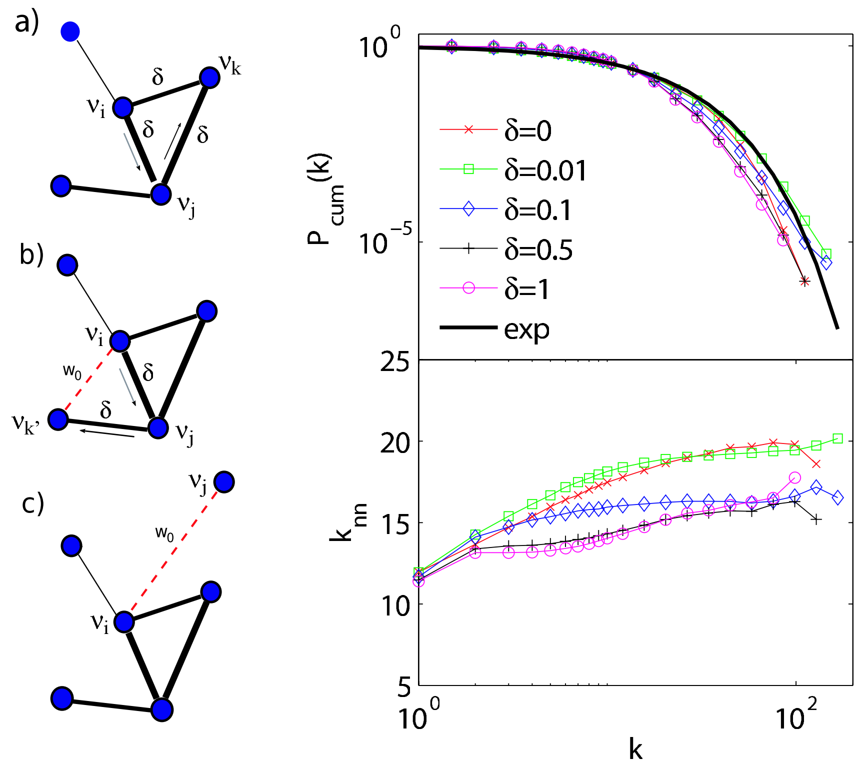

Figure 1. Left: a schematic of edge and weight dynamics in the model. (a): a weighted two-step search starts from node ν_i and proceeds through ν_j to ν_k, which happens to be a neighbor of ν_i. Links along the search path, as well as the link between ν_i and ν_k, are strengthened by δ. (b): the local search from ν_i ends at $\nu_{k'}$, which is not a neighbor of ν_i. In this case, link $e_{ik'}$ is established with probability p_Δ, with initial weight w_0. (c): with probability p_r, node ν_i creates a link to random node ν_j. Right: cumulative degree distribution and average nearest-neighbor degree k_{nn} as a function k for networks generated with various values of δ. Network size $N = 50\,000$.

2. WEIGHTED SOCIAL NETWORK MODEL

The model considered here has a fixed number N of nodes. This can be justified by the observation that internal structural changes in social networks are usually much faster than changes in network size. In the model, the initial starting point is an "empty" network of N nodes with no edges. The following rules are then repeated until distributions stabilize: at each time step each node ν_i has an opportunity to create and strengthen its links by local search (Fig. 1, a-b), and random attachment (Fig. 1, c). If node ν_i has at least one neighbor, it starts a weighted two-step search for new acquaintances among its neighboring nodes. The weighted search represents the fact that you are more likely to meet the friends of your closest friends. More precisely, ν_i chooses one of its neighbors, ν_j, with probability $w_{ij}/\sum_{m\neq i} w_{im}$, where w_{ij} is the link weight between nodes ν_i and ν_j. Further, if the chosen node ν_j has other neighbors than ν_i, the search continues to node ν_k, which is similarly chosen with probability $w_{jk}/\sum_{m\neq i,j} w_{jm}$. Both links along the search path gain a weight increment δ, motivated by the idea that each interaction strengthens the tie between two people. Now, if ν_i and ν_k are already connected, this link is also reinforced by δ as depicted in Fig. 1(a), with the same motivation. Otherwise, this link is established with probability p_Δ with an initial weight w_0, Fig. 1(b). The second way a node can obtain new neighbors is by random attachment, see Fig. 1(c). Here, with probability p_r, the node establishes a link to a random node chosen uniformly among the other N nodes. Also, if the node does not have any neighbors, it creates a random link with initial weight w_0. The initial weight w_0 sets the baseline for link weights, while the ratio δ/w_0 sets the increment with respect to this baseline.

It should be noted that if there were no mechanisms for link removal, the above-mentioned two processes would eventually fill the entire network with links. The two simplest link removal mechanisms are to remove either a single link or a single node with all its links, with some probability p_d at each time step. The essential difference between these methods is that deleting nodes removes links in a more correlated way. The resulting degree distributions are not identical, and it was observed that node removal generates a broader tail, which is the alternative we choose. Therefore, in our model, each node can be deleted with probability p_d at each time step. This can be interpreted as the node moving away from our observable scope of the social network. Evidently, in order to keep the network from shrinking, whenever a node is deleted a new one will enter the network at the next time step. The constant node removal probability leads to exponentially distributed lifetimes for nodes and links with respective means of p_d^{-1} and $(2p_d)^{-1}$.

In summary, our model parameters are p_r, p_Δ, p_d and δ/w_0. For this study, we fix $w_0 = 1$ as the arbitrary baseline for link weights. Node deletion probability is set to $p_d = 0.001$, which fixes the time scale of the simulations, and $p_r = 5 \cdot 10^{-4}$ corresponding to the addition of one random link on average for each node in 1000 time steps. This leaves us with two adjustable parameters, namely δ and p_Δ. It should be noted that the network becomes weighted when $\delta > 0$. The local search tends to favor "familiar paths" when δ is large, which increases the probability that a local search ends at a node which is already a neighbor of the node performing the search. This reduces the number of links created, unless p_Δ is increased accordingly. Therefore, p_Δ is used to adjust the average degree $\langle k\rangle$ for each δ. In the following, the target value for $\langle k\rangle$ was set to 10. After roughly 10 average node lifetimes the average degree as well as the distributions for degree, nearest neighbor degree and strength appeared to have stabilized. To be sure, we ran the network generation algorithm for 25 lifetimes, which amounts to 25 000 time steps when $p_d = 0.001$.

Figure 1 shows the cumulative degree distribution $P_{cum}(k)$ and the average nearest neighbor degree k_{nn} as a function of k for several values of δ. For comparison, an exponentially decaying curve is plotted along with the cumulative distributions. Although the difference is not large, the $P_{cum}(k)$ appear to decrease slower than exponential. Evidently, the highest degrees are $O(10^2)$, which is realistic for social networks. Degree-degree correlations also appear realistic and the resulting networks are assortative, as shown by the increasing trend for k_{nn} vs. k.

In order to illustrate the structure of our model networks, we show in Fig. 2 a 'snowball' sample, or a view into the vicinity of a randomly selected node, of a network of $N = 50\,000$ nodes. It is evident that the network has community structure, and the correlation of link weights and topology is compatible with the Granovetter hypothesis, i.e. weak links are mostly between communities whereas strong links are mostly confined within communities. This can be quantified by comparing the weight of each link against a characteristic called *overlap* O_{ij}.[1,22] This measure tells us the fraction of neighbors the end nodes of the edge e_{ij} share and hence, the value of this measure is high for edges within a community, and low for edges connecting communities. Formally, the overlap of an edge connecting vertices ν_i and ν_j is defined as $O_{ij} = \frac{n_{ij}}{(k_i-1)+(k_j-1)-n_{ij}}$, where n_{ij} is the number of common neighbors of nodes $node_i$ and $node_j$, and k_i and k_j are their degrees. The inset in Fig. 2 shows the average overlap $\langle O_{ij}|w_{ij}\rangle$ as a function of link weight. Here it is seen that links with high overlap, i.e. edges

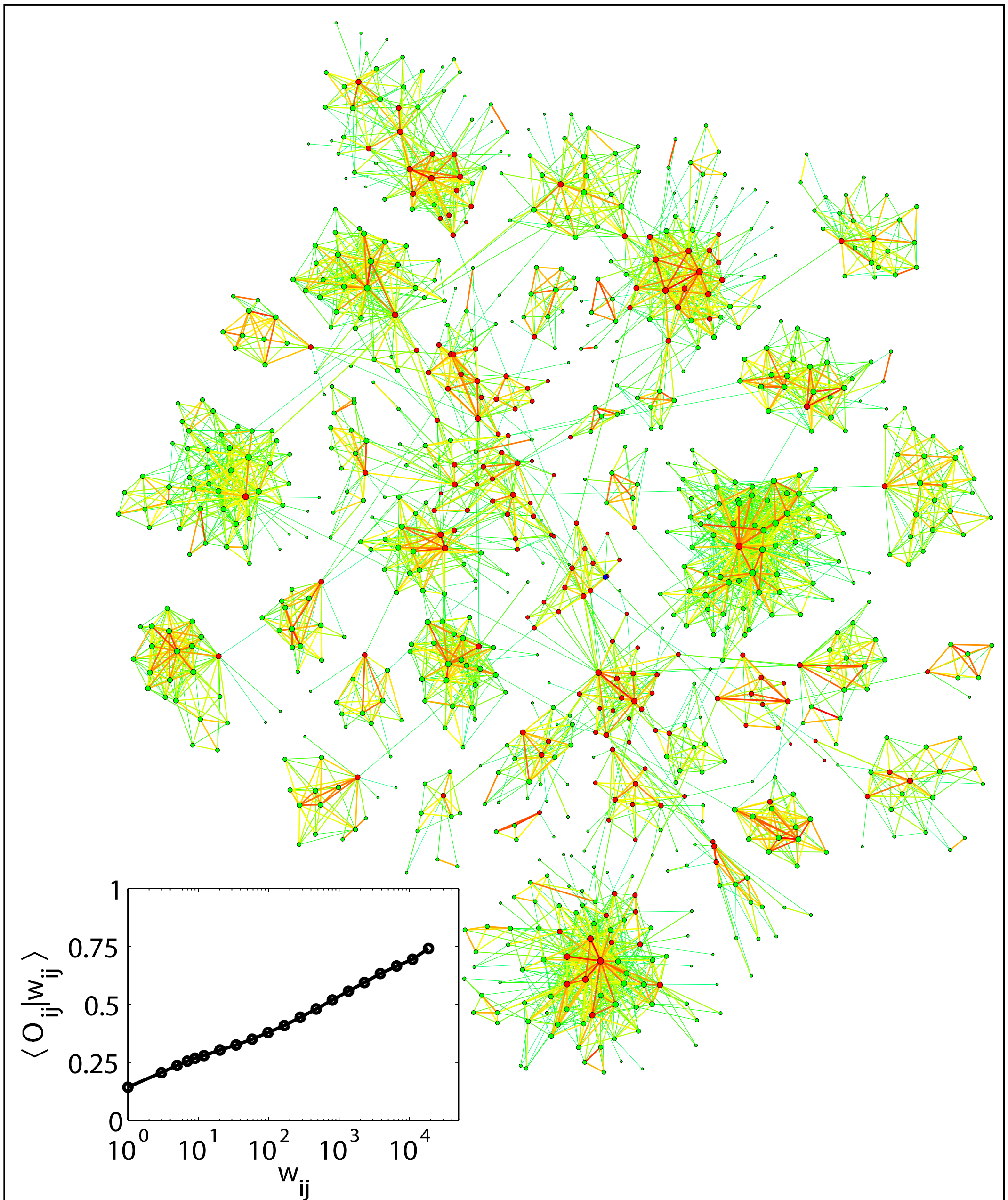

Figure 2. (Color online). 'Snowball' sample of a $N = 5 \cdot 10^4$ network with $\delta = 0.5$. The range of colors indicates link weights, from light green (small w_{ij}) through yellow to red (large w_{ij}). Link thickness is also indicative of weight. Inset: Average overlap as a function of link weight, averaged over 20 realizations of networks of size $N = 5 \cdot 10^4$

connecting nodes sharing many neighbors, have strong weight and *vice versa*, thus in line with the Granovetter hypothesis and recent findings.[1]

Another way to probe weighted network structure is link percolation, which is motivated by the observation that if weak links are mostly located between communities, the network should fragment faster when weak links are removed first. Here, following the analysis in Ref. [1], we remove links in ascending and descending order of link weights, and monitor the relative giant component size R_{LCC} as a function of fraction of removed links f. In addition to R_{LCC}, we monitor the susceptibility-like quantity $\tilde{S} = \sum_{s<s_{max}} n_s s^2/N$, where n_s is the number of disconnected components of s nodes. According to percolation theory, a collapse of the network (vanishing of the

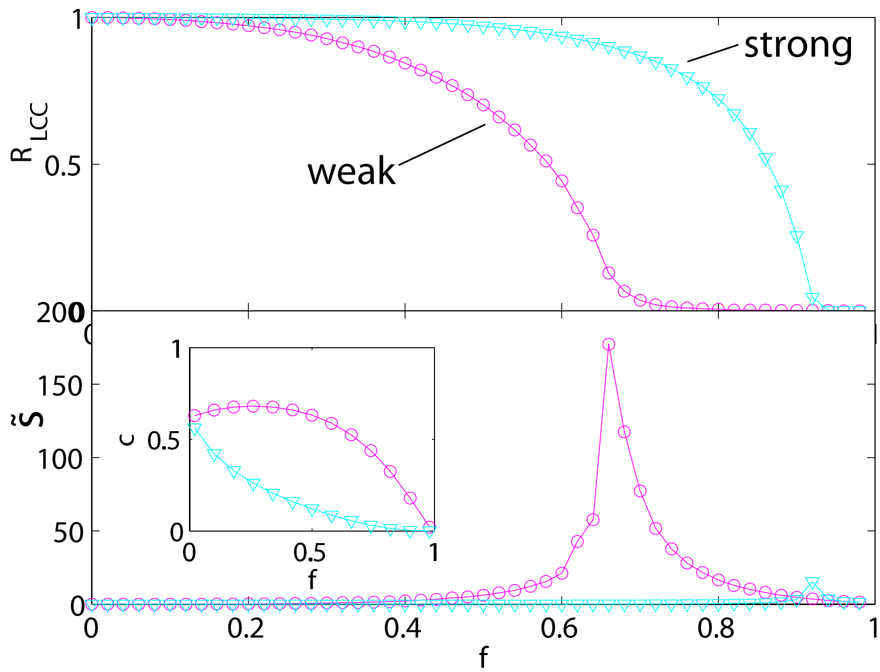

Figure 3. Relative giant component size and susceptibility as a function of f when links are removed in order of ascending (circles) and descending (triangles) weight. Inset: Average clustering. Results are averaged over 20 realizations of $N = 5 \cdot 10^4$ networks with $\delta = 0.5$.

giant component) via a phase transition results in the divergence of $\tilde{S}$ at the critical point. Figure 3 shows R_{LCC} and $\tilde{S}$ for networks with $\delta = 0.5$ for both orders of link removal, where it is evident that the giant component vanishes much earlier when weak links are removed first. This is corroborated by the high peak of $\tilde{S}$ at the fraction of links where the relative giant component size becomes vanishingly small. These results demonstrate again the correlation between link weights and (global) network topology, i.e. weak links are important for the global connectivity, as observed for the empirical weighted social network.[1] A more extensive analysis of the various unweighted and weighted characteristics of the networks will be published elsewhere.[19]

3. WEIGHTED OPINION DYNAMICS

Let us move on to look at the dynamics of opinion formation in the framework of weighted social networks. The question we address is how the inherent community structure of the network and the related weight-topology correlations affect the dynamics of a weighted three-state opinion formation model. Earlier, mesoscopic network structure, such as communities and their interconnections which are not described by global level statistics, has been found to influence several dynamical processes. For example, agents in social networks communicate intensively with only a small group of close contacts, which causes "trapping" of information to the local communities, unless the weak connections are actively utilized.[1] Another recently discovered effect of community structure is that it promotes cooperation in the Prisoner's Dilemma game.[23] Furthermore, the effect of community structure on two models of competing options has been explored recently.[24] We briefly review the latter models and results, and reformulate one of these models in the weighted context.

In both of the models studied in Ref. [24], nodes change their state depending on the states of their nearest neighbors. The *voter model* concerns the competition of two options A and B, and the state of a node is updated by imitation of a randomly chosen neighbor. The exact transition probabilities are $p_{A\to B} = \sigma_B$, $p_{B\to A} = \sigma_A$, where σ_A and σ_B are the *local densities* or the fraction of neighbors in states A or B. In words, each node is the more likely to adopt opinion A the more of its neighbors hold opinion A, and similarly for B. The *AB model*[25] includes a third, non-excluding state AB in between the two extremes A and B. A node cannot change its state directly from A to B, but has to pass through the AB state. In Ref. [24], the states A and B are interpreted as competing languages, and agents in state AB as bilingual individuals. Equivalently, we can consider the states A and B as competing extreme opinions, in which case the AB state can be interpreted either as neutral or ambivalent. The update rules are formulated as follows: $p_{A\to AB} = \frac{1}{2}\sigma_B$, $p_{B\to AB} = \frac{1}{2}\sigma_A$, $p_{AB\to A} = \frac{1}{2}(1-\sigma_B)$, $p_{AB\to B} = \frac{1}{2}(1-\sigma_A)$. A node in state A or B is the more likely to change to the intermediate state AB the more of its acquaintances are in the "opposite" state B or A. A node is not affected by its neighbors in the AB state - only A and B have the 'power of persuasion'. Both models are symmetric with respect to the states A and B.

In Ref. [24], these dynamical models were considered in a class of social type networks produced by the model presented in Ref. [16], and in randomized reference networks. It was found that the qualitative behavior of the volatile voter model was not affected by the topology. On the other hand, the network topology dramatically changed the dynamics of the AB model, where the times to reach consensus (i.e. all nodes in the network being in the same state) were found to be distributed as power law for the networks with community structure, while the distribution was exponential for the randomized counterparts. This was explained by the variety of "topological traps" related to the community structure, which produced long-lived metastable states.

Here we define a simple weighted extension of the AB-model, taking into account interaction weights. We will call it the *social influence model*, or *si*-model. It is plausible to assume that our contacts can influence us more the more time we spend with them, or the closer friends they are to us, characteristics which are depicted in the weight of the link. Instead of the local density σ, we therefore consider the *influence* ι of nodes in each state, defined as follows:

$$\iota_A(\nu_i) = \sum_{\substack{\nu_j \in \mathcal{N}(\nu_i),\\ \text{state}(\nu_j) = \text{A}}} \frac{w_{ij}}{s_i}, \tag{1}$$

where $s_i = \sum_{\nu_j \in \mathcal{N}(\nu_i)} w_{ij}$ is the strength of node ν_i, and similarly for B. In words, each neighbor ν_j influences agent ν_i with linear dependence on the strength of the tie w_{ij}. The influence values are thus in the range $(0, 1)$. The probabilities of a node changing its state are defined as follows:

$$p_{A\to AB} = \frac{1}{2}\iota_B, \qquad p_{B\to AB} = \frac{1}{2}\iota_A \tag{2a}$$

$$p_{AB\to A} = \frac{1}{2}(1-\iota_B), \qquad p_{AB\to B} = \frac{1}{2}(1-\iota_A) \tag{2b}$$

As in the AB model, a node cannot change its state from A to B or vice versa without going through the intermediate state AB. If all interaction weights are equal (unweighted networks), $\iota_A = \sigma_A$, and we recover the unweighted AB model.

A question of interest regarding these models is that under which conditions consensus is reached, and what is the process of emergence and growth of spatial domains where the nodes are in the same state. In order to describe the dynamics of the system we use the *interface density* ρ, which is defined as the fraction of links that connect nodes in different states. Note that interface density decreases as domains grow in size. Finally, if one of the states becomes dominant, the interface disappears along with the other disappearing states indicating that an absorbing state of consensus has been reached.

We use the weighted community network model described in the previous section to test the effects of both the community structure and the related weight-topology correlations (WNC, weighted network with communities) on the *si*-model. For reference, we use fully randomized versions of the networks (RN), where the edges are randomly rewired carrying their weights with them and keeping the degree sequence intact, under the restriction that the network must stay connected. This eliminates structural correlations, and therefore the resulting randomized networks are locally treelike. In order to show the significance of the coupling between weights and topology, we also use another type of reference, where weights have been randomly relocated while keeping the topology (WRNC, weight-randomized network with communities). All simulations begin with a random initialization of the nodes in states A, B or AB. At each time step each node is updated in a random order (random sequential update) according to the transition probabilities (2a-2b) until we reach either an absorbing state where all nodes are in the same state, or a predetermined maximum number of time steps.

The left panel of Fig. 4 shows the average interface density $\langle \rho \rangle$ in these cases, averaged over 100 realizations of the network with 10 runs of the *si*-model in each. In the WNC networks (black), interface density decreases very slowly, with shape resembling a power law. At $t = 10^5$, none of the runs had yet reached the absorbing state. Comparing this with the WRNC networks (light gray), we observe that the layout of weights in the network significantly increases the time it takes for the system to reach the absorbing state. In the fully randomized case (RN, dark gray), the system reaches consensus much faster, indicating that also the structural correlations play their part in slowing down the dynamics. The interface density is lower in WNC and WRNC than in RN, indicating that in the networks with topological community structure the characteristic size of the homogenized regions is larger.

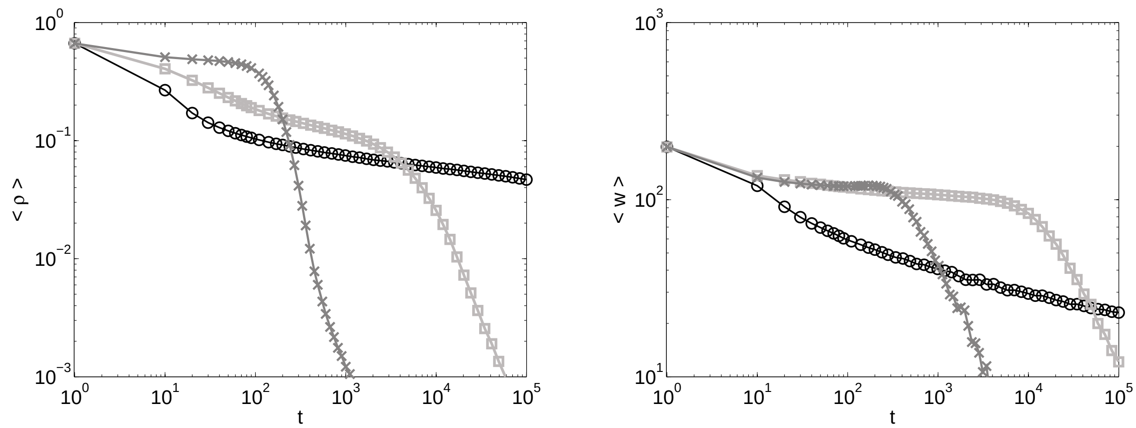

Figure 4. Left: Average interface density $\langle \rho \rangle$ as a function of time. Right: Average weight $\langle w \rangle$ in the interface as a function of time. Original networks with weighted community structure (WNC, black), networks with weights randomly relocated in the topology (WRNC, light gray) and networks with randomized topology while keeping the degree and weight sequences intact (RN, dark gray). Averages are taken over 100 network realizations, with 10 runs of the *si*-model in each. Parameters of the original networks: $N \approx 3000$, $\delta = 0.5$, $p_\Delta = 0.02$, $p_d = 0.001$, and $p_r = 5 \cdot 10^{-4}$, which result in average degree $\langle k \rangle \approx 7$.

The right panel of Fig. 4 shows the average weight $\langle w \rangle$ of the interface edges, similarly averaged. In all of the networks, average weight decreases slightly during the first few time steps. Only in the networks with strong weight-topology correlations (WNC, black), however, the average interface weight keeps decreasing as the interface density decreases. The ever fewer interfaces are thus settling on the weak links, which we know to be located between the communities. Thus, we can infer that the nodes in each community in the network are becoming homogenized to the same opinion.

4. SUMMARY

In this study we have presented a model of social networks which takes into account interaction strengths or link weights between the nodes. While link weights make the model more realistic in describing human interaction in a social network, they also generate a clear community structure. By using this social network model we have explored the effect of weights on a three-state model of opinion formation and found that both the distribution of weights within network topology and the community structure have a pronounced effect on the time it takes for the system to reach an absorbing state. We conclude that the interaction weights play an important role in social dynamics.

Acknowledgments

We would like to thank Victor Eguíluz, Maxi San Miguel and Xavi Castelló for the idea of using a three-state model of social dynamics and related discussions. This work was supported in part by Cost Action P10 (COST-STSM-P10-02707), and by the Academy of Finland (Center of Excellence program 2006-2011).

REFERENCES

1. J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási, "Structure and tie strengths in mobile communication networks", *PNAS*, **104**, pp. 7332-7336, 2007.
2. M. Granovetter, "The strength of weak ties", *Am. J. Soc.* **78**, pp. 1360–1380, 1973.
3. R. Albert and A.-L. Barabási, "The statistical mechanics of complex networks", *Rev. Mod. Phys.* **74**, p. 47, 2002.
4. S. N. Dorogovtsev and J. F. F. Mendes, "Evolution of networks", *Adv. Phys.* **51**, p. 1079, 2002.
5. M. Newman, "Structure and function of complex networks," *SIAM Review* **45**, pp. 167–256, 2003.
6. J. Scott, *Social network analysis: A Handbook, 2nd ed.*, Sage, London, 2000.
7. L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, "Classes of small-world networks", *PNAS* **97**, pp. 11149–11152, 2000.
8. M. Boguña, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, "Models of social networks based on social distance attachment", *Phys. Rev. E* **70**, 056122, 2004.
9. M. Newman and M. Girvan, "Finding and evaluating community structure in networks", *Phys. Rev. E* **69**, 036104, 2006.
10. G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society", *Nature* **435**, pp. 814–818, 2005.
11. J. Reichardt and S. Bornholdt, "Statistical mechanics of community detection," *Phys. Rev. E* **74**, 016110, 2006.
12. E. M. Jin, M. Girvan, and M. E. J. Newman, "Structure of growing social networks," *Phys. Rev. E* **64**, 046132, 2001.
13. J. Davidsen, H. Ebel, and S. Bornholdt, "Emergence of a small world from local interactions: Modeling acquaintance networks," *Phys. Rev. Lett.* **88**, 128701, 2002.
14. A. Grönlund and P. Holme, "Networking the seceder model: Group formation in social and economic systems", *Phys. Rev. E* **70**, 036108, 2004.
15. L. H. Wong, P. Pattison, and G. Robins, "A spatial model for social networks", *Physica A* **360**, pp. 99–120, 2006.
16. R. Toivonen, J.-P. Onnela, J. Saramäki, J. Hyvönen, and K. Kaski, "A model for social networks", *Physica A* **371(2)**, 2006.
17. A. Barrat, M. Barthélémy, and A. Vespignani, "Weighted evolving networks: Coupling topology and weight dynamics", *Phys. Rev. Lett* **92**, 228701, 2004.
18. A. Barrat, M. Barthélémy, and A. Vespignani, "The effects of spatial constraints on the evolution of weighted complex networks", *J. Stat. Mech.* , P05003, 2005.
19. J. Kumpula, J.-P. Onnela, J. Saramäki, K. Kaski, and J. Kertész, "Emergence of communities in weighted social networks", *Manuscript in preparation*
20. M. Marsili, F. Vega-Redondo, and F. Slanina, "The rise and fall of a networked society: A formal model", *PNAS* **101**, pp. 1439–1442, 2004.
21. J. Davidsen, H. Ebel, and S. Bornholdt, "Emergence of a small world from local interactions: Modeling acquaintance networks", *Phys. Rev. Lett.* **88**, 128701, 2002.
22. J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, M. A. de Menezes, K. Kaski, A.-L. Barabási, and J. Kertész, "Analysis of a large-scale weighted network of one-to-one human communication", *New J. Phys.*, *in press*, arXiv:physics/070631, 2007.
23. S. Lozano, A. Arenas, and A. Sánchez, "Mesoscopic structure conditions the emergence of cooperation on social networks", arXiv:physics/0612124, 2006.
24. X. Castelló, R. Toivonen, V. Eguíluz, J. Saramäki, K. Kaski, and M. San Miguel, "Anomalous lifetime distributions and topological traps in ordering dynamics", arXiv:0705.2560, 2007.
25. X. Castelló, V. Eguíluz, and M. San Miguel, "Ordering dynamics with two non-excluding options: bilingualism in language competition", *New J. Phys.* **8**, 2006.