

Model of Community Emergence in Weighted Social Networks

J. M. Kumpula^a, J.-P. Onnela^{a,b}, J. Saramäki^a, J. Kertész^{a,c}, K. Kaski^a

^a*Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, P.O. Box 9203, FIN-02015 HUT, Finland*

^b*Physics Department, Clarendon Laboratory, Oxford University, Oxford OX1 3PU, U.K.*

^c*Institute of Physics, Budapest University of Technology and Economics, Budapest, Hungary*

Abstract

Over the years network theory has proven to be rapidly expanding methodology to investigate various complex systems and it has turned out to give quite unparalleled insight to their structure, function, and response through data analysis, modeling, and simulation. For social systems in particular the network approach has empirically revealed a modular structure due to interplay between the network topology and link weights between network nodes or individuals. This inspired us to develop a simple network model that could catch some salient features of mesoscopic community and macroscopic topology formation during network evolution. Our model is based on two fundamental mechanisms of network sociology for individuals to find new friends, namely *cyclic closure* and *focal closure*, which are mimicked by local search-link-reinforcement and random global attachment mechanisms, respectively. In addition we included to the model a node deletion mechanism by removing all its links simultaneously, which corresponds for an individual to depart from the network. Here we describe in detail the implementation of our model algorithm, which was found to be computationally efficient and produce many empirically observed features of large-scale social networks. Thus this model opens a new perspective for studying such collective social phenomena as spreading, structure formation, and evolutionary processes.

Key words: communities, weighted network, social network

PACS: 89.75.Hc, 89.65.-s, 89.75.Fb

1. Introduction

Various complex systems, from genetic transcriptions to human societies, have been approached from the perspective of network theory, turning out to contribute significantly to the understanding of their structure, function, and response [1, 2]. In case of human social systems, this approach was first adopted by social scientists [3, 4] studying fairly small data sets collected from e.g. questionnaires and then by physicists [5, 6, 7, 8, 9] focusing on large-scale data sets collected from e.g. Internet, emails etc., both giving insight into social network structure with one-to-one interactions. In the social sciences the network paradigm is that *social life consists of the flow and exchange of norms, values, ideas, and other social and cultural resources channeled through a network* [10]. By studying large-scale social networks we can get insight into collective social phenomena such as diffusion and spreading processes (epidemics), opinion formation, and evolution of language etc. In order to do so we ask (i) how are large scale social networks organized and (ii) can they be modeled with a simple model.

Answering the first question has traditionally been based on the analysis of data from questionnaires typically among $N = 10^2 - 10^3$ individuals with a wide scope social interactions but with limited resolution in quantifying unambiguously the strength of interaction between a pair of individuals. In the new approach data is obtained from electronic records of interactions between typically of $N = 10^6$ individuals with a narrower scope of social interactions but as such accurately quantifiable through measurements. We have demonstrated the success of the new approach by studying social interaction network constructed from a very large data set of mobile phone communication logs [11]. Apart from its global structure this network shows local structure due to link weights or tie strengths, such that the network was found to be modular in terms of communities with dense and stronger internal links connected sparsely with weaker external links. This network shows high degree clustering and assortative mixing demonstrating that high degree nodes in the network are connected to other high degree nodes. Perhaps the most noteworthy finding of this empirical study, however, is that the network verifies the "strength of the weak ties" hypothesis by Granovetter stating: Tie strength between two individuals increases with the overlap of their friendship circles. Moreover, the weak and strong ties have different roles such that the weak ties maintain the global integrity of the network while strong ties maintain the communities. In addition the structural prop-

erties of the network at meso-scale were found to play an important role in its functional properties, e.g. diffusion of information.

Inspired by these empirical findings from a social network the latter question can be rephrased as an aim to design a simple microscopic model to reproduce similar structural properties and then using it to simulate dynamical processes. The fact that the empirical network is structured as modules or communities raises the question "How do the communities emerge during the growth of the network". On the other hand the fact that the links of the network are weighted raises the question "how do they influence the formation of mesoscopic communities and macroscopic network topology". Then the empirical verification of Granovetter's hypothesis links these two questions as "how should we take into account the weight - topology correlation". Thus we aim to build a model, in which individual-level microscopic mechanisms translate to forming on one hand mesoscopic communities and on the other hand the whole system at the macroscopic level. In this network sociology comes to help in identifying two fundamental mechanisms for network tie formation leading to network evolution: (i) *cyclic closure* forming ties at short range with one's network neighbours, and (ii) *focal closure* forming ties independently of the range through shared activities [12]. Here we adopt a simple scenario that new ties are created preferably through strong ties with every interaction making them even stronger, mimicking people getting acquaintances with local and global search mechanisms dependent on link weights unlike the mechanisms proposed by Marsili *et al.* [13] and Davidsen *et al.* [14].

In this paper we give a detailed account of our model on social community formation [15], in terms of describing the modelling methodology and algorithms used. Although our perspective aims to be sociologically relevant and to corroborate some empirical findings we believe that our model serves as a quite general paradigm for community formation in other contexts as well. This paper is organised such that we first describe the model with account to the microscopic mechanisms, their algorithmic implementation, and parameter and initial condition settings. Then we discuss the results by describing the scaling of our algorithm, basic distributions of the formed networks, community detection in them, and their weight-topology correlations. Finally we draw conclusions.

2. The model

As described above we have built the model algorithm to mimic the network tie formation processes identified by sociology. While the actual processes taking place in the formation of social relations are undoubtedly very complex, sociology research has shown that getting new acquaintances consist of mainly two processes, which in our model are simplified to the mechanisms of local search and random (global) attachment. In sociology the effect of link weight is very complicated to investigate empirically but it is reasonable to assume that people interact mostly through their "strong friendships" such that each interaction makes this connection even stronger. In reality it is also possible that some friendships disappear corresponding to deletion of a link between two nodes. However, this process is often so slow that it is difficult to observe from available data due to it being restricted to rather narrow time windows. In our model the number of links is reduced by deletion of a node and all the links connected to it. This could correspond to a situation, in which the individual having belonged to the network for a while goes out of its scope (e.g. in case of mobile communication changing the carrier).

2.1. Microscopic rules

The microscopic rules of our model consist of two processes that create links into a network and one process that removes them. The first process of link formation is local search for new acquaintances (LA process), which is modeled by two steps of a weighted self-avoiding random-walk, see Fig. 1(a,b). In the first step, node ν_i chooses one of its neighbors, ν_j , with probability w_{ij}/s_i where, w_{ij} is the link weight connecting nodes ν_i and ν_j , and $s_i = \sum_j w_{ij}$. In the second step the search continues similarly to a neighbor of ν_j (but avoiding ν_i). The visited links are reinforced by a small amount δ . If the two-step search ends at a node that is not a neighbor of the start node, this link is established with probability p_Δ to a unit weight $w = 1$. Otherwise, the existing link is reinforced by δ . In addition to the local search each node can establish a random link with probability p_r at each time step (GA process), see Fig. 1(c). Also, if a node does not have links, it creates one random link with a unit weight $w = 1$. Finally, at each time step a node is deleted (ND process) with probability p_d , which is the only process that decreases the number of links in the network. Note that the removed node gets replaced by a new node at next time step, which keeps the network size N constant. The random deletion of nodes leads to exponential distribution for

node and link lifetimes, averages of which are $\langle \tau \rangle = p_d^{-1}$ and $\langle \tau_w \rangle = (2p_d)^{-1}$, respectively.

2.2. Algorithmic implementation

In most cases our model with microscopic rules described above produces very sparse networks, i.e., the average degree is much less than the network size, thus making their storage as full matrixes very inefficient from computational point of view. Instead we have implemented the model by storing the network Γ as a linear-probing hash-table [16], which guarantees small memory consumption, constant time retrieval of network nodes and links, and efficient looping over the neighbours of each node. At the beginning of the simulation the model network Γ is empty, i.e., it contains N nodes but no links. The simulation proceeds in time such that at each time step the nodes perform the LA and GA processes on network Γ . However, Γ is kept unchanged during the time step and a temporary network Γ^* is used to store the changes resulting from LA and GA processes such that weight increase of existing links and new links are added to Γ^* . At the end of each time step the network Γ is updated by looping over the links of Γ^* and copying the weight increases and new links to Γ , after which all links are removed from Γ^* .

As mentioned earlier, the number of links is reduced by removing nodes randomly from the network and correspondingly adding nodes without links to keep the network size constant. In practise, the node removal process is handled as follows: when a node is chosen to be removed, it is not actually deleted from the network but all links connected to it are removed. At next time step this empty node will join the network by introduction of a random link as described in Sec. 2.1.

2.3. Simulation parameters and initial conditions

The simulations start from an empty network of N nodes and proceed in time such that the changes due to local searches, random attachments, and node deletions are updated at the end of each time step. Average node lifetime $\langle \tau \rangle$ was set to 1000 time steps and random link probability to $p_r = 1 \times 10^{-3}$ corresponding to adding on average 2 random links for each node during average node lifetime. The following results were obtained after running the simulations for 25×10^3 time steps, which is long enough for all the statistics of the model to reach a steady state.

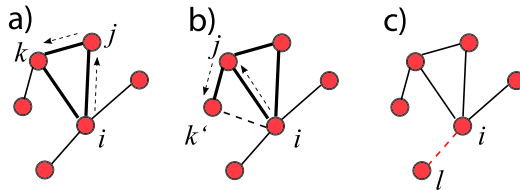


Figure 1: The microscopic rules of the model consist of two processes. The first process of tie formation is local search (cf. *cyclic closure*). (a): a weighted local search starts from i and proceeds first to j and then to k , which is a neighbor of i . (b): the local search from i ends to k' , which is not a neighbor of i . In this case a link $w_{ik'}$ of unit weight ($w = 1$) is established with probability p_{Δ} . The second process for tie formation is random link creation (focal closure). (c): node i creates a random link to a random node l with probability p_r . In cases a) and b) the weights of involved links are increased by the amount δ .

The weights enter the model dynamics through the reinforcement of visited links, which is controlled by the "friendship reinforcement" parameter δ . When δ is large some links start to dominate the local search process attracting almost all searches and becoming all the time stronger. In this case the search ends up almost always to a "familiar" node, that is, the start and end nodes are already connected. On the other hand, when $\delta = 0$ the local search follows all links with equal probability. Now the search exits the neighborhood of the start node quite easily, and a new link is then established with probability p_{Δ} . Other parameters being fixed the average degree of the network is thus controlled by parameters δ and p_{Δ} . In practice p_{Δ} is adjusted for each δ such that the average degree $\langle k \rangle$ remains constant, here we use $\langle k \rangle = 15$ unless otherwise mentioned. Keeping $\langle k \rangle$ constant allows easy comparison of networks with different δ and ensures that differences in the network structure result solely from the reorganization of the links. Below we present a simple algorithm for automatically adjusting p_{Δ} for each value of δ and average target degree $\langle k \rangle_{target}$.

First, we note that $\langle k \rangle$ is a monotonically increasing function of p_{Δ} when other parameters are kept constant. Second, the value of p_{Δ} can be adjusted during a simulation provided that one waits sufficiently long time after each adjustment such that the network structure has time to adapt to the new situation. Third, even for constant p_{Δ} there are random fluctuations in the value of $\langle k \rangle$ and therefore one should not be too avid to change p_{Δ} if $\langle k \rangle$ deviates from $\langle k \rangle_{target}$. Now, suppose we have decided $\langle k \rangle_{target}$, fixed δ and

p_r , and initiated the simulation with an initial guess $p_\Delta^0 > 0$. We measure the average degree of the network at 1000 time step intervals, and denote the i th measurement by $\langle k \rangle_i$. The value of p_Δ is adjusted after N_p such measurements; here we have used $N_p = 7$. The average of N_p measurements of $\langle k \rangle_i$ is denoted by

$$\overline{\langle k \rangle}_I = \frac{1}{N_p} \sum_{i=IN_p+1}^{IN_p+p} \langle k \rangle_i, \quad (1)$$

where $I = 0, 1, 2, \dots$ is the index of the adjustment round. The new value of p_Δ is then calculated from

$$p_\Delta^{I+1} = p_\Delta^I \left(1 - \alpha_I \frac{\overline{\langle k \rangle}_I - \langle k \rangle_{target}}{\langle k \rangle_{target}} \right), \quad (2)$$

where $\alpha_I = 0.9\alpha_{I-1}$ and $\alpha_0 = 2$. The optimal value for p_Δ is usually obtained after 100 adjustment rounds, e.g., by setting $p_\Delta = 1/50 \sum_{I=50}^{99} p_\Delta^I$. The obtained value of p_Δ can then be used for generating model networks that have the desired average degree.

3. Results

3.1. Scaling of the algorithm

Before presenting results for the formation of social networks we first investigate the cpu-time scaling of our model algorithm. It is natural to expect that the required cpu-time depends on the network size N , average degree $\langle k \rangle$, number of iterations, but also weakly on other network parameters. Figure 2 shows the cpu-time as a function of network size for networks having $\langle k \rangle = 10$ and $\langle k \rangle = 15$. The simulation runs were performed for 25000 time steps. Here it can be seen that scaling of the cpu-time is linear for both networks, which was expected because the algorithm uses only local processes (and the GA process is always performed in constant time). Note also that our algorithm could be parallelized rather easily because at each time step the LA and GA processes for each node are independent of other searches. However, we have not attempted this.

3.2. Basic distributions

In Fig. 3 we present the basic distributions for the simulated networks with varying friendship reinforcement parameter $\delta \in [0, 1.6]$ and $\langle k \rangle = 15$.

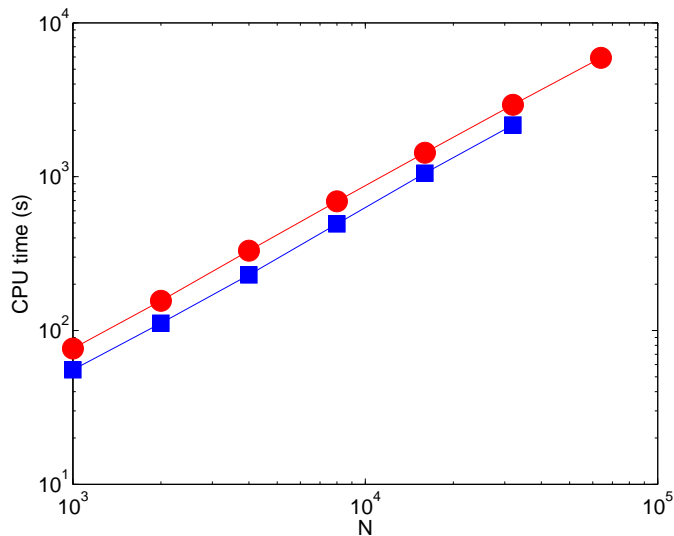


Figure 2: (Color online) The scaling of cpu-time for the model algorithm as a function of network size. Networks with $\langle k \rangle = 10$ ($p_r = 5 \cdot 10^{-4}$, $\delta = 1$, $p_\Delta = 0.044$) are denoted by (o) and networks with $\langle k \rangle = 15$ ($p_r = 1 \cdot 10^{-3}$, $\delta = 1$, $p_\Delta = 0.0719$) by (\square). Simulations were done on desktop computer with 3.0 GHz Pentium 4 processor.

There are a number of analyses of large data bases [11, 17, 18, 19] that suggest at least the following common features for social networks: *i*) degree distributions are skewed, *ii*) high-degree nodes are often connected to other high-degree nodes (assortative mixing), *iii*) the average clustering coefficient $\langle c \rangle$ is high compared to random networks, and *iv*) the average shortest path lengths are small (the small-world property). Figure 3 shows also the corresponding distributions for the mobile phone call network (note that this network has $\langle k \rangle = 3.1$). As can be seen, for all the studied values of δ the properties of our model network are in agreement with features *i*)-*iii*) and also with feature *iv*) as the network diameter grows as $\log N$ (not shown). In this sense even $\delta = 0$ can be considered as a valid choice for modelling social networks, since our results resemble those obtained from certain unweighted models [13, 14]. The effect of δ becomes important only when the community structure due to link weights is taken into account, as will be discussed next.

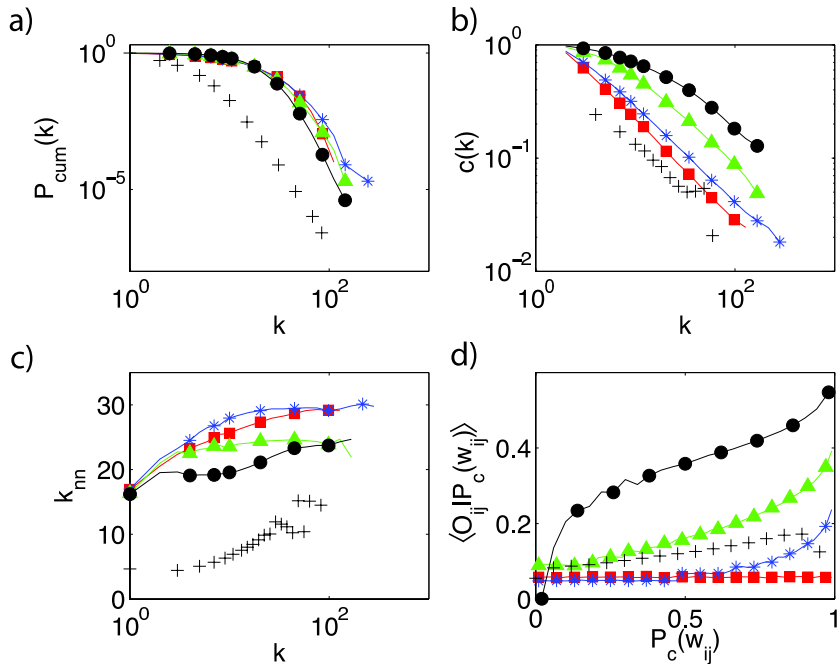


Figure 3: (Color online) Basic distributions of model networks for several values of δ . (a) Degree distribution, (b) clustering, (c) average nearest neighbor degree k_{nn} , and (d) average link overlap as a function of cumulative link weight. Results are averaged over 50 realizations for $N = 1 \times 10^4$ networks. Values for δ are 0 (\square), 2×10^{-2} ($*$), 0.2 (\triangle), and 1.6 (\circ). Results for mobile phone call data [11] are for comparison and labeled by (+).

3.3. Network community structure by clique percolation

The effect of δ on network structure is visualized in Fig. 4, where it is clearly seen that increasing δ promotes the formation of communities. Moreover, strong links seem to be confined in communities while the links between communities are mainly weak. The topological structure of the model networks can be studied by k -clique percolation method. This method defines the communities by adjacent cliques of k nodes, see Fig. 5 (a). Here we have used a recently proposed fast algorithmic implementation for k -clique percolation [20], which scales even for very large networks. We used 4-cliques approach but similar results are obtained also for $k = 3$ and $k = 5$. Figure 5 (b) shows the relative largest community size $R_{k=4}$ and average community size excluding the largest community $\langle n_s \rangle$ as a function of δ for $\delta \in [0, 1.6]$. When $\delta = 0$ the communities turn out to be quite small having at most ap-

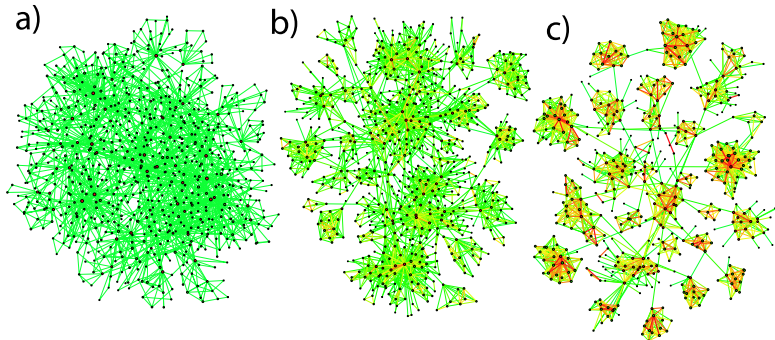


Figure 4: (Color online) Visualization of the effect of δ on network topology. (a) $\delta = 0$, (b) $\delta = 0.1$, and (c) $\delta = 1$ ($\langle k \rangle = 10$ for these networks). Weak links are green and the color changes gradually to red for strong links.

proximately 50 nodes and the average community size is $\langle n_s \rangle \approx 6$. Increasing δ changes the community structure significantly. First, the 4-cliques start to percolate through most of the network, which can be seen as a fast increase in $R_{k=4}$, but when $\delta \gtrsim 0.5$ $R_{k=4}$ starts to decrease and simultaneously $\langle n_s \rangle$ increases to about 30 nodes. This means that when δ increases above $\gtrsim 0.5$ the communities start to "condense" to the initially rather homogeneous network. The distribution of community sizes was found to be stable after the initial lead-in time of approximately 10 – 20 average node life times.

3.4. Weight-topology correlations

Finally, we discuss the correlations between link weights and network topology, and the effect of δ on them. If the strong links are confined mainly in tight communities and the links between communities are predominantly weak, as the weak ties hypothesis by Granovetter suggests, the network should break into pieces easier by removing the links in ascending than in descending order of weights. This was observed for the mobile phone call network [11] and on the basis of visual inspection of Fig. 4 one would expect also our model behave similarly when δ is sufficiently large. To confirm this we remove links from the network in both orders while monitoring the relative giant component size R_{LCC} , the susceptibility $\tilde{s} = \sum n_s s^2 / N$, and the average clustering coefficient $\langle c \rangle$ [5] as a function of the ration of removed links f . Here n_s is the number of components of size s . The results for networks for $\delta \in [0, 1.6]$ are shown in Fig. 6(a) together with the corresponding

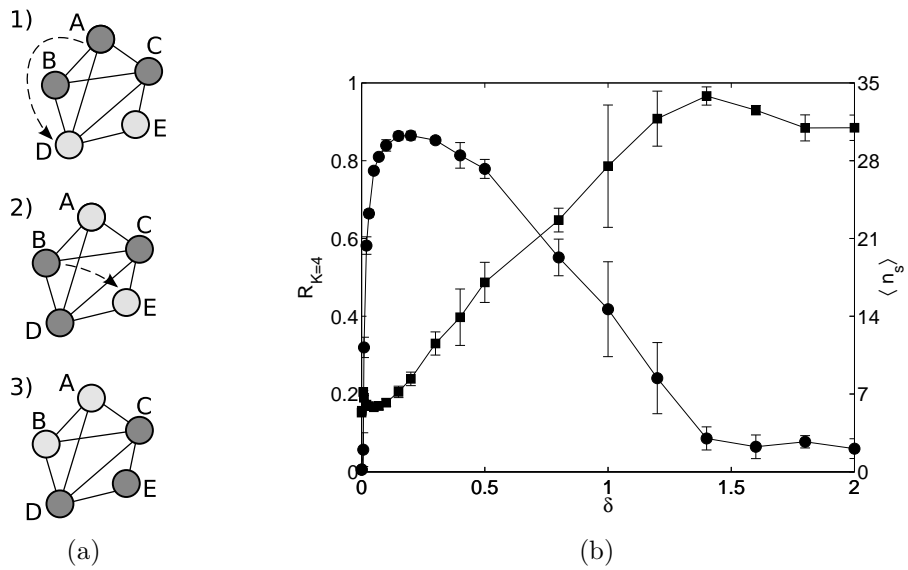


Figure 5: (a) An illustration of a 3-clique "rolling". The 3-clique ABC first rolls on to BCD by releasing node A. When node D is released the 3-clique rolls on CDE. Thus, nodes A, B, C, D, and E form a 3-clique community (but not a 4-clique community since node E does not belong to any 4-clique). (b) Largest 4-clique community size $R_{k=4}$ (\circ) and average community size (excluding the largest) $\langle n_s \rangle$ (\square) as a function of δ . Results are averaged over 50 realizations for size $N = 1 \times 10^4$ networks.

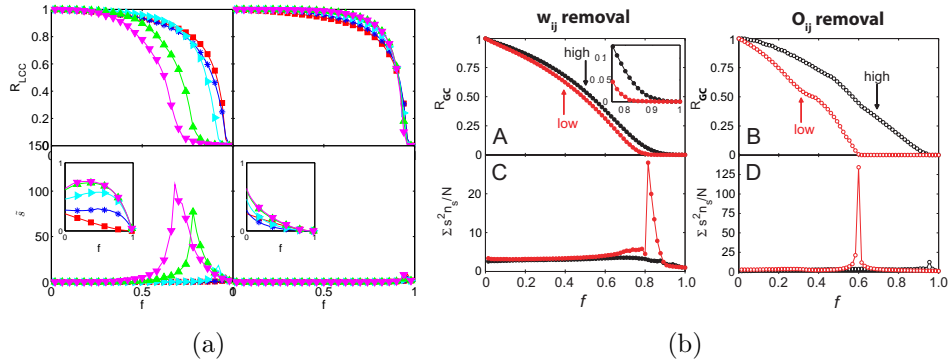


Figure 6: (Color online) (a) Link percolation results for the model networks. Upper panel shows the relative giant component size R_{LCC} and the lower panel the susceptibility \tilde{s} . On the left: weak links are removed first; on the right: strong links are removed first. In the inserts average clusterings are shown. Results are averaged over 50 realizations for size $N = 5 \times 10^4$ networks. Values of δ are 0 (\square), 2×10^{-2} ($*$), 0.2 (\triangleright), 1.0 (\triangle), and 1.6 (∇). For comparison (b) link percolation results for the mobile phone call data published in [11]. Upper and lower panels and are as in the model networks (a), and descending weight removal is in black and ascending in red. Copyright (2007) by The National Academy of Sciences of the United States of America, all rights reserved.

results for the mobile phone call network, in Fig. 6(b). When δ is small the giant component becomes disintegrated at the same point regardless of the link removal order and \tilde{s} remains very small. Therefore, coupling between weights and topology appears to be weak in this case, which is not surprising considering the results of clique percolation. The difference between link removal orders appear when $\delta \gtrsim 0.5$. Now, when weak links are removed first, the giant component starts to disintegrate earlier and the susceptibility \tilde{s} develops a significant peak, which can be interpreted such that the network breaks into many quite large components, i.e., communities. On the other hand, when strong links are removed first, \tilde{s} remains small for all values of δ and the giant component disintegrates always at the same point. The insets in Fig. 6(a) show the behavior of $\langle c \rangle$ for both link removal orders. The rapid decrease in $\langle c \rangle$ when removing the strong links first indicates that the strong links locate mainly in tightly connected components, while the opposite is true for weak links. This is corroborated also by Fig. 3(d), which shows the link overlap $o_{ij} = n_{ij}/(k_i + k_j - 2 - n_{ij})$, where n_{ij} is the number of common neighbors of nodes ν_i and ν_j , as a result of cumulative link weight [11]. There it is seen that also in our model the overlap increases with link weight for

sufficiently large δ , as the weak ties hypothesis suggests, thus supporting the observation for strong links being confined in communities. Our link percolation results are in good agreement with the corresponding results for the mobile phone call data as can be seen in Fig. 6(b).

4. Conclusions

In this paper we have introduced a simple weighted network model based on sociologically justified microscopic mechanisms to mimic mesoscopic community and macroscopic topology formation in social system. In this model the link weights are an essential part of the microscopic mechanisms establishing a coupling between network structure and interaction strengths between nodes, such that the introduction of a new link depends on the existing link weights, which are reinforced once a new link has been introduced. In addition to the local search and random attachment mechanisms that in the model correspond the cyclic closure and focal closure of network sociology, we also introduced a mechanism of random node removal with all its links simultaneously removed. The implementation of the model algorithm turned out to be computationally very efficient since the cpu-time of a simulation scales linearly with the size of the network. The simulation results show that communities emerge after nucleating from structural fluctuations but only if link weight or friendship reinforcement is sufficiently strong. In addition our model fulfills various topological and statistical properties of social networks and it exhibits weight-topology correlation complying with the Strength of Weak Tie Hypothesis by Granovetter. For this our model provides plausible mechanisms and it could in general serve as a starting point for understanding community formation in weighted networks.

Acknowledgements: J.M.K., J.-P.O., J.S. and K.K. acknowledge the Academy of Finland, the Finnish Center of Excellence program 2006-2011, proj. 213470. J.-P.O. is supported by a Wolfson College, University of Oxford, Junior Research Fellowship. J.M.K. is partly supported by the GETA graduate school. J.K. was partially supported by OTKA T049238 grant.

References

- [1] M. E. J. Newman, A. L. Barabási, and D. J. Watts, *The structure and dynamics of networks*, Princeton University Press, 2006.

- [2] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, Oxford University Press, New York, 2007.
- [3] M. Granovetter, “Economic action and social structure: The problem of embeddedness,” *The Sociology of Economic Life* **91**, pp. 481–510, 2001. The Sociology of Economic Life.
- [4] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge University Press, 1994.
- [5] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature* **393**, pp. 440–442, 1998.
- [6] R. Albert, H. Jeong, and A. L. Barabási, “Internet: Diameter of the world-wide web,” *Nature* **401**(6749), p. 130, 1999.
- [7] H. Ebel, L. I. Mielsch, and S. Bornholdt, “Scale-free topology of e-mail networks,” *Physical Review E* **66**(3), p. 35103, 2002.
- [8] J.-P. Eckmann, E. Moses, and D. Sergi, “Entropy of dialogues creates coherent structures in e-mail traffic,” *Proc. Natl. Acad. Sci. USA* **101**(40), pp. 14333–14337, 2004.
- [9] P. S. Dodds, R. Muhamad, and D. J. Watts, “An experimental study of search in global social networks,” *Science* **301**, pp. 827–829, 2003. Date: -Modified = 2007-01-20 19:03:21 +0000.
- [10] H. C. White, S. A. Boorman, and R. L. Breiger, “Social structure from multiple networks. i. blockmodels of roles and positions,” *American Journal of Sociology* **81**(4), p. 730, 1976.
- [11] 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,” *Proc. Natl. Acad. Sci.* **104**(18), pp. 7332–7336, 2007.
- [12] G. Kossinets and D. J. Watts, “Empirical analysis of an evolving social network,” *Science* **311**, pp. 88–90, January 6 2006.
- [13] M. Marsili, F. Vega-Redondo, and F. Slanina, “The rise and fall of a networked society: A formal model,” *Proc. Natl. Acad. Sci.* **101**(6), pp. 1439–1442, 2004.

- [14] J. Davidsen, H. Ebel, and S. Bornholdt, “Emergence of a small world from local interactions: Modeling acquaintance networks,” *Phys. Rev. Lett.* **88**(12), p. 128701, 2002.
- [15] J. M. Kumpula, J. P. Onnela, J. Saramäki, K. Kaski, and J. Kertész, “Emergence of communities in weighted networks,” *Phys. Rev. Lett.* **99**, p. 228701, 2007.
- [16] K. K. J. Hyvönen, J. Saramäki, “Efficient data structures for sparse network representation,” *Intl. J. Comp. Math.* **85**(8), pp. 1219–1233, 2008.
- [17] J. P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, M. 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.* **9**(6), p. 179, 2007.
- [18] M. C. González, H. J. Herrmann, J. Kertész, and T. Vicsek, “Community structure and ethnic preferences in school friendship networks,” *Physica A* **379**(1), pp. 307–316, 2007.
- [19] M. E. J. Newman, “The structure of scientific collaboration networks,” *Proc. Natl. Acad. Sci.* **98**(2), pp. 404–409, 2001.
- [20] J. M. Kumpula, M. Kivelä, K. Kaski, and J. Saramäki, “Sequential algorithm for fast clique percolation,” *Phys. Rev. E* **78**(2), p. 026109, 2008.