Multilayer weighted social network model

Yohsuke Murase,[1,2,*] János Török,[3] Hang-Hyun Jo,[4,5] Kimmo Kaski,[4] and János Kertész[6,3,4]

[1]*RIKEN Advanced Institute for Computational Science, 7-1-26, Minatojima-minami-machi, Chuo-ku, Kobe, Hyogo, 650-0047, Japan*

[2]*CREST, Japan Science and Technology Agency 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan*

[3]*Department of Theoretical Physics, Budapest University of Technology and Economics, HU-1111 Budapest, Hungary*

[4]*Department of Biomedical Engineering and Computational Science, Aalto University School of Science, P.O. Box 12200, FI-Espoo, Finland*

[5]*BK21plus Physics Division and Department of Physics, Pohang University of Science and Technology, Pohang 790-784, Republic of Korea*

[6]*Center for Network Science, Central European University, Nádor u. 9, HU-1051 Budapest, Hungary*

(Received 6 August 2014; published 17 November 2014)

Recent empirical studies using large-scale data sets have validated the Granovetter hypothesis on the structure of the society in that there are strongly wired communities connected by weak ties. However, as interaction between individuals takes place in diverse contexts, these communities turn out to be overlapping. This implies that the society has a multilayered structure, where the layers represent the different contexts. To model this structure we begin with a single-layer weighted social network (WSN) model showing the Granovetterian structure. We find that when merging such WSN models, a sufficient amount of interlayer correlation is needed to maintain the relationship between topology and link weights, while these correlations destroy the enhancement in the community overlap due to multiple layers. To resolve this, we devise a geographic multilayer WSN model, where the indirect interlayer correlations due to the geographic constraints of individuals enhance the overlaps between the communities and, at the same time, the Granovetterian structure is preserved.

DOI: 10.1103/PhysRevE.90.052810 PACS number(s): 89.75.Hc, 89.65.−s

I. INTRODUCTION

The abundance of data due to the rapid development of the information and communication technology (ICT) has generated entirely new, multidisciplinary approaches in social sciences [1,2], in which physics plays a considerable role both in terms of data analysis and modeling. One of the major challenges in this context is the understanding of the structure of the society, which is crucial for many applications ranging from epidemiology to urban planning. While traditional techniques based mainly on questionnaires focused on small scale organization of the society [3] the new tools enable one to uncover the structure on many scales up to the societal level. A broad range of ICT data has been used to study empirically these questions. Examples include email [4], mobile phone call (MPC) [5–8], short-message communication, social network services (SNS) [9], and scientific collaborations [10,11].

Mobile phone data have a special role in this endeavor as the coverage in the adult population approaches 100% and much of the interpersonal communication runs today over mobile phones. Therefore the records of the calls can be used to map out the network of social interactions [5,6,12]. In this mobile call network the famous Granovetter hypothesis about the "strength of weak ties" [13] turned out to be correct. According to this hypothesis links between individuals have different strengths corresponding to the intensity of the relationship, the time spent together, mutual confiding, etc., and the stronger is a tie, the larger is the overlap between the further contacts of those who form the tie. This local property has severe consequences on the entire structure: The society consists of communities, which are strongly wired and these communities are then connected by weak ties, thus playing an important role to hold society together. The duration or the frequency of calls serves as a natural measure of the strength of ties for mobile phone calls and in this way it was possible to prove the Granovetter hypothesis on this data set [5,6].

In order to demonstrate the global consequences a link percolation analysis was carried out. Provided that links are sorted according to their weights, removing the weakest links first one by one results in a sharp transition at a relatively early stage, indicating the fragmentation of the society. In the opposite case, when links were eliminated in the descending order of their weights, the percolation threshold set in at a much higher portion of removed links because strong links are within the communities, where a large number of paths between nodes exist. In this sense, the difference Δf_c between the two percolation thresholds can be considered as the measure of the Granovetterian character of the network.

After this empirical verification of the Granovetter hypothesis, the next step was to understand the mechanisms leading to the formation of these structures in a social network by constructing a model, which incorporates basic link-formation processes between individuals. Two main mechanisms were taken into account, namely local and global attachment rules together with tie strength reinforcement [14,15]. Here the local and global attachment rule correspond to cyclic and focal closure mechanisms [4], with the former referring to the link formation with one's network neighbors, or with friends of friends and the latter to the attribute-related link formation which is independent of the local network topology. The reinforcement step corresponds to the general observation that social ties get strengthened by using them. With these simple processes, the complex Granovetterian weight-topology relation of social networks could be successfully reproduced as demonstrated for the large value of Δf_c [14].

The community structure of complex networks is an extensively studied topic [16]. The identification of communities or structural modules, i.e., groups of nodes having more connections among themselves than outside the group is a

*yohsuke.murase@gmail.com

highly nontrivial task and much effort has been devoted to its solution (see, e.g., [17–19].) Most of the methods produce a partition of the network, meaning that a node can belong to only one community. However, as pointed out in [20] this cannot lead to an appropriate description of many complex networks, especially of social ones, where there is usually considerable overlap between the communities due to the fact that nodes can belong simultaneously to several of them. A number of algorithms have been suggested to uncover overlapping communities [20–22].

The community detection method of Ahn *et al.* [22] was based on the identification of link communities; at the same time they suggested a remarkable mechanism as the origin of overlapping communities. Using the language of social networks (what we are interested in here), they propose that a person can be in different types of relationships, like kinship, collaboration, friendship, etc. Moreover, people are switching their social contexts and communication channels depending on the occasions, and the social network should strongly depend on the context [23,24]. To handle these aspects, it is necessary to represent the social networks as a multilayer or multiplex network [25–27], where each layer corresponds to a different type of relationship. Since these contexts are hardly distinguishable from the available data, the networks observed in this way are usually considered as projections or an aggregate of multiple layers. Such a projection of multilayer networks should be in line with the observed stylized facts faithfully with empirical data, such as the Granovetter-type structure. An important aspect of the multilayer structure is missing both from the original Granovetter paper [13] and the above described model [14]. The aim of the present paper is to investigate the possibilities to model the combination of the multilayer structure of the society with the Granovetterian relationship between tie strengths and topology. In order to do so, we start from the simple, single-layer model by Kumpula *et al.* [14] and introduce the multilayer structure in different ways.

This paper is organized as follows. In the next section, a naive multilayer network is investigated and it is shown that it leads to a breakdown of the Granovetter-type structure as correlations are suppressed. We therefore introduce the copy-and-shuffle model, where a parameter tunes the correlations. We find a regime with Δf_c significantly different from zero, however, there the average number of overlapping communities, a node participates in, is low. To overcome this difficulty, we formulate a model in Sec. III, where correlations are caused by the dependence on the geographic distance. This model has a parameter region where both the Granovetterian structure and a considerable enhancement of the average overlap are observed. The last section is devoted to a summary and discussion.

II. MULTILAYER WEIGHTED SOCIAL NETWORK (WSN) MODEL

A. Single-layer WSN Model

Let us first summarize the original WSN model by Kumpula *et al.* [14]. It considers an undirected weighted network of N nodes. The links in the networks are updated by the following three rules. The first rule is called *local attachment* (LA). Node i chooses one of its neighbors j with probability proportional to w_{ij}, which stands for the weight of the link between nodes i and j.

Then, node j chooses one of its neighbors except i, say k, randomly with probability proportional to w_{jk}. If node i and k are not connected, they are connected with probability p_Δ with a link of weight w_0, but if they are already connected this link weight and the other two link weights w_{ij} and w_{jk} in a triangle are increased by δ. The second rule is *global attachment* (GA), where if a node has no links or otherwise with probability p_r, it is connected to a randomly chosen node with weight w_0. Finally, the third rule *node deletion* (ND) is introduced to the model, where with probability p_d, a node loses all its links. At each time step, LA, GA, and ND are applied to all nodes. Starting from a network without any links, the network reaches a statistically stationary state after a sufficient number of updates. As a function of the reinforcement parameter δ this model shows a gradual transition from a module free topology to a Granovetterian structure with strongly wired communities connected by weak ties.

B. Generalization to the multilayer case

In order to study multilayer effects we generalize the single-layer WSN model in the following naive way. We consider L layers of the same set of nodes and we assume that each layer corresponds to a different type of relationship or communication context. For each layer, we independently construct a network in the same way as in the original single-layer WSN model. For simplicity, the same parameters are used for all the layers. After the stationary networks are constructed in each layer, the aggregate network is constructed by summing up the edge weights: $w_{ij} = \sum_{k=1}^{L} w_{ij}^k$, where w_{ij}^k is the weight of the link between nodes i and j in the kth layer [28]. It is this aggregate network for which we expect the Granovetterian structure.

In the following, $N = 50\,000$, $p_r = 0.0005$, $p_\Delta = 0.05$, $p_d = 0.001$, $\delta = 1$, and $w_0 = 1$ are used. The results are obtained after 25×10^3 time steps and averaged over 50 realizations. To see whether the multilayer model reproduces a realistic social network of the kind the mobile phone call (MPC) graph is a proxy [5,6], a link percolation analysis is carried out for the model. We removed fraction f of the links from the generated networks in both ascending and descending orders, and measured the relative size R_{LCC} of the largest connected component and the normalized susceptibility $\chi = \sum n_s s^2/N$, where n_s is the number of components of size s and the sum is taken over all but the largest component. At the percolation threshold the order parameter R_{LCC} vanishes and χ diverges in the thermodynamic limit. For finite systems the former quantity shows a fast decay and the latter one a sharp peak at the threshold value f_c. The significant difference Δf_c in the thresholds for the two sequences of link removal is characteristic by the Granovetter structure; $\Delta f_c = f_c^d - f_c^a$, where the upper index d (a) stands for descending (ascending) sequences of removed links.

Figure 1 shows R_{LCC} and χ as a function of f for a single-layer network ($L = 1$) and a double-layer network ($L = 2$). The two plots in each figure show the results for ascending

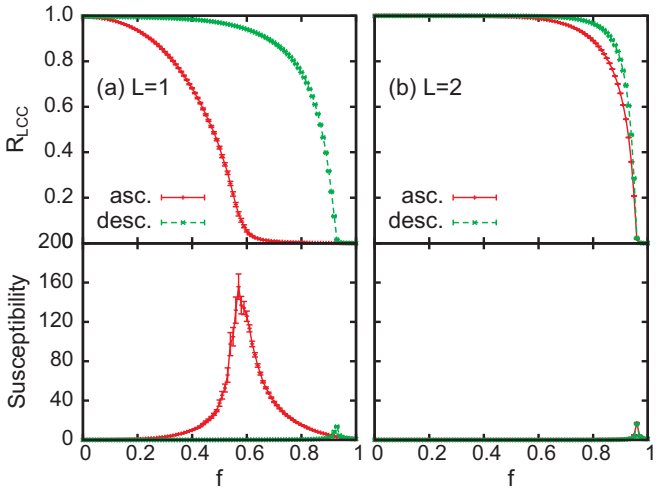

FIG. 1. (Color online) Link percolation analysis for $L = 1$ (left) and $L = 2$ (right). The upper figures show the relative size of the largest connected component, R_{LCC}, as a function of the fraction of the removed links f. The lower figures show the susceptibility χ. Red solid (green dashed) lines correspond to the case when links are removed in ascending (descending) order of the link weights. The error bars show standard errors.

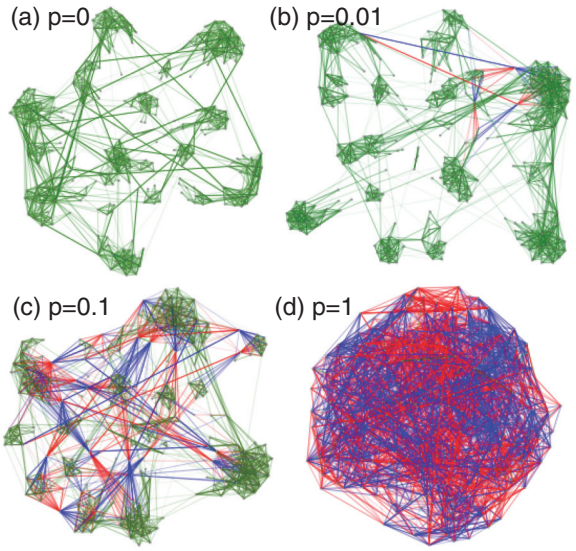

FIG. 2. (Color) Snapshots of the copy-and-shuffle model with different p shuffling parameter values and $N = 300$. Red (blue) links are in the first (second) layer, and green links are in both layers.

and descending orders. For $L = 1$ we get $\Delta f_c \approx 0.35$, while for $L = 2$ the figure shows that the percolation threshold for ascending order f_c^a is not significantly different from that for descending order f_c^d (i.e., $\Delta f_c \approx 0$).

The percolation thresholds for $L = 2$ are approximately the same, $f_c \approx 0.95$, indicating that the introduction of a second layer destroys the Granovetterian structure. The percolation threshold agrees well with that of an Erdős-Rényi (ER) random network having the same average degree $\langle k \rangle$ as the simulated model: $f_c = 1 - 1/\langle k \rangle$ with the measured $\langle k \rangle = 21.9$. (Note that this is twice the average degree of a single layer.) This observation shows that combining already two independent layers from the original single-layer WSN model leads to a high level of randomization in the aggregate [29]. One may think that the observed effect is due to the increasing total degree when two layers are merged. However, we carried out simulations, where the total degree was controlled by p_Δ and found that for $L = 2$ the thresholds are always very close to each other; $\Delta f_c \approx 0$.

C. Copy-and-shuffle WSN model

Due to the fact that merging two layers of WSN models destroys the Granovetterian structure, we investigated how the correlation between layers affects the properties of the network. We created the second layer by copying the first layer and then shuffled the fraction p of the nodes in the second layer. Shuffling nodes i and j means that all original links (i,k) become (j,k) and vice versa. This is just a relabeling of the nodes in that layer, meaning that the topology remains the same, i.e., both layers correspond to single-layer WSN models but with increasing p the correlations between them decrease. This is called the "copy-and-shuffle" model (see Fig. 2).

When $p = 0$, the aggregate network is equivalent to the single-layer network whose link weights are doubled. For $p = 1$, it is the same as the double-layer model; the Granovetterian structure gets entirely destroyed by randomization. By controlling p between 0 and 1, a transient behavior is observed. Figure 3 shows how the threshold values f_c^a get closer to f_c^d as p is increased and for $p \to 1$ we get $\Delta f_c \to 0$. The reason is that strong links in the second layer connect the communities more randomly since the correlation between the first and the second layer diminishes.

The randomization has the consequence that the percolation threshold f_c^a gets closer to that of the corresponding Erdős-Rényi random network. However, for a reasonably large range

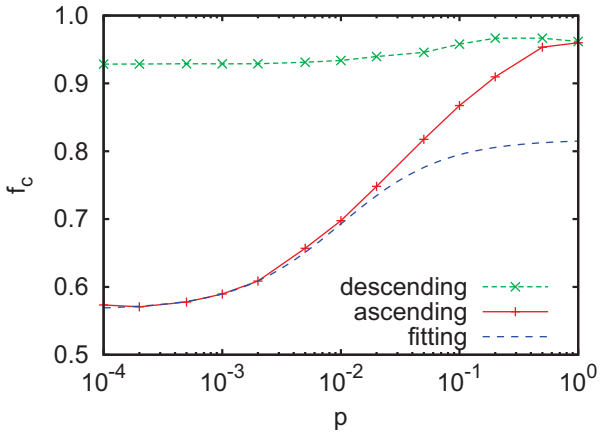

FIG. 3. (Color online) Percolation thresholds for various shuffle fraction values p for the copy-and-shuffle model. The green upper and red lower lines denote the critical points f_c^d and f_c^a, respectively. The critical points are determined by the peak of the susceptibility. The points are calculated for 50 independent runs. The blue dashed line is calculated using Eq. (3).

of p, we can clearly differentiate f_c^a and f_c^d thus the similarity between the layer assures the Granovetterian structure.

The gradual transition can be understood in the following way. Let us make the assumptions that the original network (first layer) is composed of strongly connected groups interconnected by weak links [see Fig. 2(a)] and the average size of these groups is $\bar{s}$ which is small and independent of the total number of agents N. Eventually the number of groups is $N_g = N/\bar{s}$. In the link percolation analysis starting from the weak links (ascending order) we can consider the groups as "supernodes" and we have to solve the percolation problem for the links connecting them.

Not only the intergroup links turn out to be weak but also some intragroup ones. Let us denote the number of links by M, the total number of weak links by M_w, and the intergroup weak links by M_g which are a subset of M_w. Let us remove the f^a fraction of the total links in ascending order. In this case the number of weak links gets $M_w(f^a) = M_w - Mf^a$ since we only removed weak links. Thus it is clear that this approximation will not work for $f^a > M_w/M$. The intergroup links are removed by the same rate as weak links, so the number of intergroup links after removing the f^a part of the total links in ascending order is

$$M_g(f^a) = M_g\left(1 - f^a\frac{M}{M_w}\right). \tag{1}$$

In order to have quantitative results we need the number of strong links. This can be estimated if we assume that $M_w = M(\langle k\rangle - 2)/\langle k\rangle$. This means that each node has two strong links while the rest are weak. This can be justified by considering the effect of cyclic closure: The cyclic closure is the most frequent interaction which strengthens two links of a focal node with positive feedback of their weights. Assuming a random network for the groups at the percolation threshold, one should get $M_g(f_c^a) = N_g/2$. This gives $\langle k_g\rangle \equiv 2M_g/N_g = 3.26$ for the single-layer model, where $\langle k_g\rangle$ is the average degree of the connections between the groups, i.e., that of the supernodes. Snapshots of the single-layer model as in Fig. 2(a) justify the low contact number for the groups.

The shuffling of agents creates new connections from one group to another. These connections increase the intergroup connection degree by

$$\langle k_g\rangle(p) \approx \langle k_g\rangle(p=0) + p\bar{s}\langle k\rangle. \tag{2}$$

Thus there is a linear increase of the average degree of the groups with p.

Now we can use Eq. (1) to go in the reverse direction, namely that knowing M_g/N_g we can get the percolation threshold (note that M_w and M also depend on p in a trivial way):

$$\begin{aligned} f_c^a &= \frac{M_w(p)}{M(p)}\left(1 - \frac{N_g}{2(M_g + pN\langle k\rangle)}\right) \\ &= \frac{\langle k\rangle - 2}{\langle k\rangle}\left(1 - \frac{1}{\langle k_g\rangle(p=0) + 2\bar{s}\langle k\rangle p}\right). \end{aligned} \tag{3}$$

The average size of the groups for the single layer can be obtained by an infomap analysis [19] and was found to be $\bar{s} = 15.1$. The resulting curve is shown in Fig. 3 as a dashed line and is compared to the empirical threshold values. The calculated line fits the initial part very well, where the above picture is expected to work.

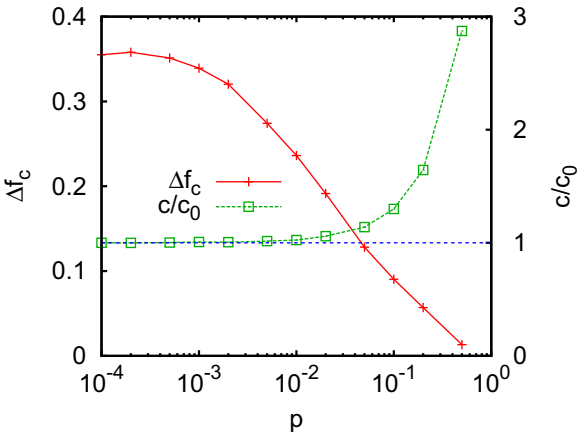

FIG. 4. (Color online) Characteristic quantities for the copy-and-shuffle multilayer WSN model. The difference Δf_c between the percolation thresholds decreases with the shuffling probability p, while $\bar{c}/\bar{c}_0$, the ratio of the average number of communities a node belongs to at parameter value p and $p = 0$, decreases. There is no regime, where Δf_c is significantly larger than zero and $\bar{c}/\bar{c}_0$ is considerably larger than one. The results are averaged by 50 independent samples and the errors are smaller than the symbol size.

The copy-and-shuffle model produces a region of p, where a multilayer Granovetterian structure exists. Now we have to check whether our construction has led to enhancement of the overlapping of the communities, too. We have analyzed the aggregate networks by the method of Ahn *et al.* [22] and calculated $\bar{c}/\bar{c}_0$, the ratio of the average numbers of communities a node belongs to at parameter value p and $p = 0$ [30]. We expect that $\bar{c}/\bar{c}_0$ should increase as shuffling goes on. Figure 4 shows the dependence of this quantity on p. The overlap starts to increase only when the Granovetterian correlation between link weight and topology is already wiped away. In the next section we make another attempt to produce a model, where the Granovetterian structure and overlapping communities coexist.

III. GEOGRAPHIC MULTILAYER WSN MODEL

The above results show that some correlations are needed between layers in order to have Δf_c significantly different from zero for a multilayer model. Previous studies have reported that there are strong geographic constraints on social network groups even in the era of the Internet [31] and this is reflected in the MPC data [12,32,33]. For example, intercity communication intensity is inversely proportional to the square of their Euclidean distance, which is reminiscent of the gravity law [12,32].

Motivated by these observations, we consider now a model embedded into a two-dimensional geographic space. At the beginning of the simulation nodes are distributed randomly in the unit square with periodic boundary condition. These geographic positions are fixed and shared by all the layers. We assume that the probability for making a new connection in the global attachment (GA) step in the WSN model is higher if the two nodes are geographically close. The probability

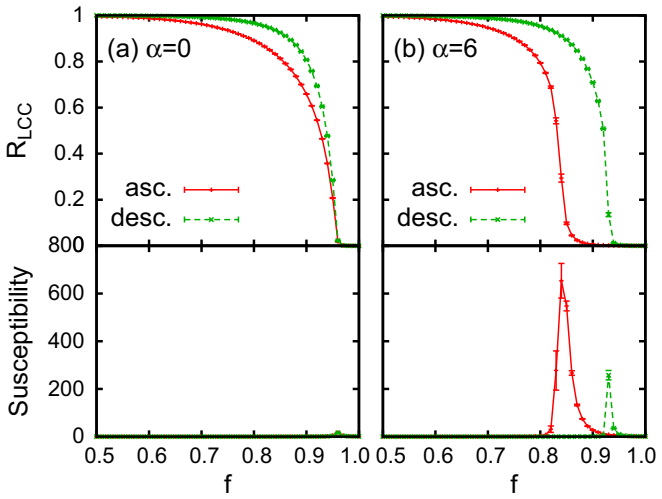

FIG. 5. (Color online) Link percolation analysis for (a) $\alpha = 0$ and (b) $\alpha = 6$. The upper figures show the relative size of the largest connected component, R_{LCC}, as a function of the fraction of the removed links f. The lower figures show the susceptibility χ. Note that the scale of the horizontal axis is different from Fig. 1. Red solid (green dashed) lines correspond to the case when links are removed in ascending (descending) order of the link weight. The results are obtained by 50 independent samples. The error bars show standard errors.

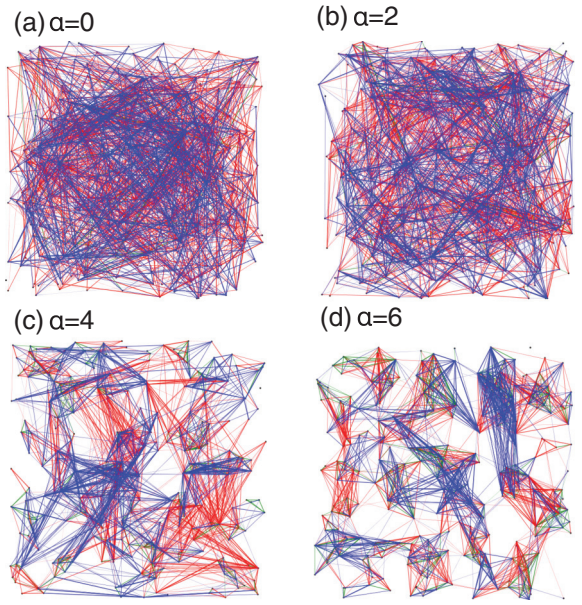

FIG. 6. (Color) Sample of double-layer networks of the geographic model for $N = 300$. Links only in the first and the second layers are shown in blue and red, respectively, and links shared by both layers are depicted in green.

that node i makes a new connection to node j by GA is proportional to $r_{ij}^{-\alpha}$, where r_{ij} is a distance between nodes i and j, and α is a new parameter controlling the dependence on geographic distance as in [34,35]. When $\alpha = 0$, this probability is independent of the geographic distance, thus the model is equivalent to the uncorrelated multilayer model we presented in the previous section. When α is larger, the nodes tend to be connected with geographically closer nodes yielding the correlation between the networks in different layers. Since only nonconnected pairs are considered, the probability for node i to make connection with a node j which is not yet connected to i is given by

$$p_{ij} = \frac{r_{ij}^{-\alpha}}{\sum_{k \in S_i} r_{ik}^{-\alpha}}, \tag{4}$$

where S_i is the set of the nodes not connected to the node i. The other rules such as LA or ND are kept the same as in the original WSN model.

Figure 5 shows the results for link percolation analysis for the geographic model with $\alpha = 0$ and 6. Because the network for larger α has a smaller average degree, we used a larger value of p_r (0.002) in order to keep the average degree comparable to the results for the nongeographic model ($\langle k \rangle = 18.0$ for $\alpha = 6$ and $\langle k \rangle = 27.6$ for $\alpha = 0$). As shown in the figure, the network for $\alpha = 6$ exhibits a Granovetterian structure as f_c^a and f_c^d are significantly different with $\Delta f_c \approx 0.1$.

Small samples of networks ($N = 300$) for different α are shown in Fig. 6. While the network for $\alpha = 0$ and 2 look similar to the uncorrelated double-layer nongeographic network in Fig. 2(d), the networks for larger α clearly show a nice community structure. About 19% of the links are shared by two layers for $\alpha = 6$, while less than 0.1% percent of the total links are shared for $\alpha = 0$; see also the inset of Fig. 7. This already indicates the possibility of overlapping communities.

The dependence of f_c on α is shown in Fig. 8, which summarizes the main results for the geographic model. Δf_c becomes larger with increasing α and seems to get saturated around 0.15. The ratio $\bar{c}/\bar{c}_0$ decreases rather rapidly and reaches the limit value of 2. This means that for sufficiently large α we have *both* Granovetterian properties and the enhancement of the number of overlapping communities due

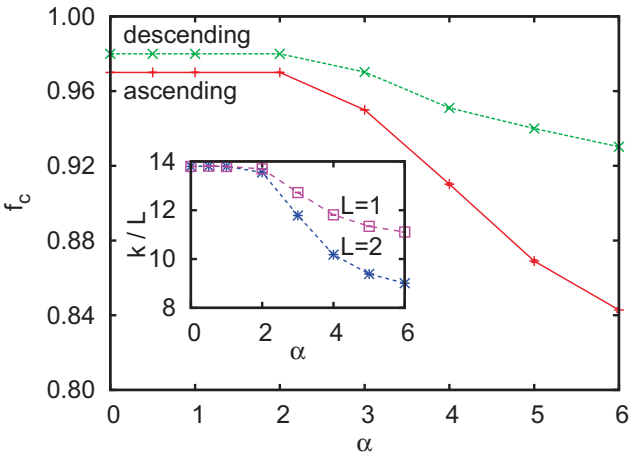

FIG. 7. (Color online) Percolation thresholds f_c^a and f_c^d for the geographic model as a function of α. The percolation thresholds are determined as the point where maximum susceptibility is observed and then averaged over 50 independent samples. (Inset) Average degrees divided by L as a function of α for $L = 1$ and $L = 2$. The values for $L = 2$ are smaller than those for $L = 1$ when α is sufficiently large because there are links appearing in both layers.

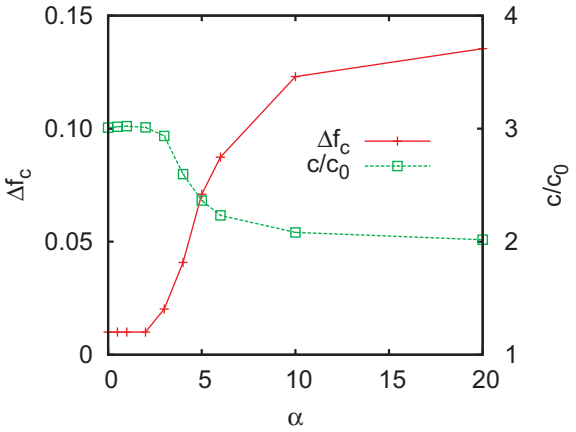

FIG. 8. (Color online) This figure is a similar plot in Fig. 4 for the geographic multilayer WSN model. Here Δf_c and $\overline{c}/\overline{c}_0$ are shown as a function of α. Note that for this model $\overline{c}_0$ also depends on α. For $\alpha \geqslant 6$ we have Δf_c significantly larger than 0 and $\overline{c}/\overline{c}_0$ close to 2. The results are obtained by 50 independent samples.

to the multilayer structure. We note that for $\alpha > 4$ both the average degree in one layer $\langle k(L=1)\rangle \simeq 11$ and $\overline{c}_0(\alpha > 4) \simeq 3$ is the same as in the nongeographic case indicating similar structure.

We note that neither the percolation thresholds nor the average degree show significant dependence on α for $\alpha < 2$. We speculate that this is because the network dimension becomes infinite for $\alpha < 2$ even when it is embedded in a two-dimensional space [35]. Since the dimensionality of the network is finite, the clustering coefficient for the network is higher compared to the network with smaller α. (For $\alpha = 0$, 2, and 6, clustering coefficients are 0.23, 0.24, and 0.55, respectively.) This also explains the change in the average degree. If a link and its neighboring link are selected by LA, the probability that the third link closing the triangle is already there will be higher for higher α thus the number of links newly created by LA is smaller leading to the decrease in the average degree.

IV. SUMMARY AND DISCUSSION

Our aim in this paper has been to model two important properties of the social network: Its Granovetterian structure and the large amount of overlapping communities due to its multilayer character. We introduced the difference Δf_c of the percolation thresholds f_c^d and f_c^a as a single variable characterizing the weight-topology relation and $\overline{c}/\overline{c}_0$, the ratio of the average number of communities a node participates in for the multilayer and the single-layer networks. We expect from a model satisfying our goal simultaneously a Δf_c considerably different from zero and $\overline{c}/\overline{c}_0$ significantly larger than one.

The naive introduction of multiple layers of single-layer WSN models breaks the Granovetter-type weight-topology relation since the communities in one layer get connected by strong ties in another layer. If we control the amount of randomness by the shuffling probability p and start from replicas of single-layer models then we arrive at a multilayer

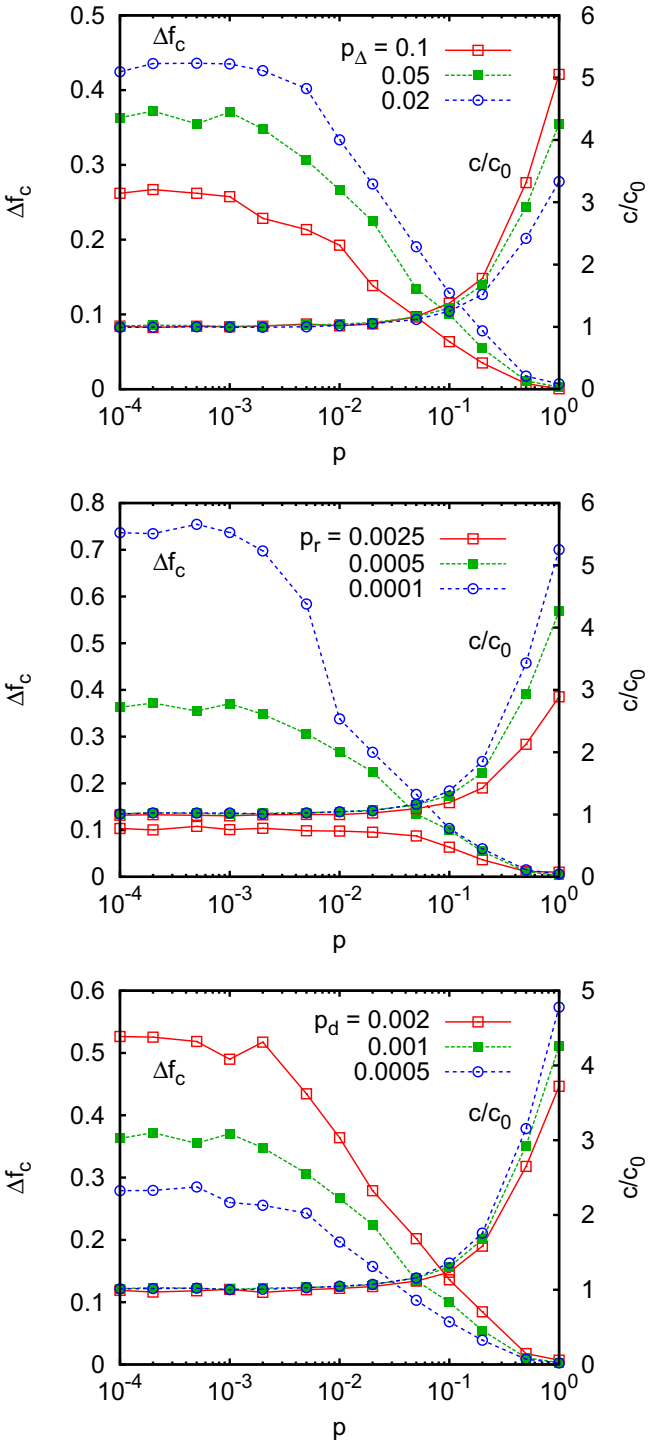

FIG. 9. (Color online) These figures show the same quantities as Fig. 4 for several values of p_Δ (a), p_r (b), and p_d (c). The results are obtained by simulation of $N = 10\,000$ and averaged over 20 independent samples.

model, which, however, has no region of the control parameter p, where both required properties can simultaneously be observed.

In order to maintain both requirements, we introduced an extension of the single-layer model such that each node has a geographic position and that geographically close nodes

have more chance to form a link via GA. The multilayer model consists then of a combination of such single layers. Controlling the distance dependence by the exponent α [as defined in Eq. (4)] we conclude that for $\alpha \geqslant 6$ we have a multilayer weighted social network, which has both the Granovetterian structure and the enhanced community overlap.

We note here that previous studies on MPC data have revealed that intercity communication density is inversely proportional to the square of the distance [12,32]. Regarding the geographic model, the relation between α and the exponent characterizing intercity communication density is not trivial because the links created by LA are not affected by the geographic position. Furthermore, we assumed that the position of the nodes are uniformly distributed, which is clearly an idealized aspect of the model since we know that population usually aggregates around city areas.

Our results have several implications. First, they show that geographic correlations play a key role to change the picture drastically in a multilayer weighted network similarly to what was observed for interdependent networks [36]. Moreover, although the models we studied are strong simplifications of the society, we believe that they have their role in the investigation of social structures. In particular, such models enable one to study the special effects of the Granovetterian and the overlapping community structure on dynamic phenomena like spreading.

Communities organize themselves along common attributes like sharing working places, classes at universities, joint interest, e.g., in sport, residential districts, etc. [34]. Geographic proximity is just one of the possibilities and other attributes can play an important role in the formation of network as well. Future work is needed to find out how to treat explicitly these attributes and their impact on the formation of the network.

Our models have also implications for further empirical studies. Unfortunately, most data sets contain only information about one channel of communication, which substantially restricts sampling of relationships even in the case of mobile call networks. An alternative approach is "reality mining," where a limited number of volunteers (of the order of one hundred) carry devices, which record several channels of communication, including face-to-face encounters [37]. This methodology could pave the way for studies of the effects of the multilayer character of human society, especially from the points of view presented in this paper.

ACKNOWLEDGMENTS

Y.M. appreciates hospitality at Aalto University and H.-H.J. acknowledges financial support from the Aalto University postdoctoral program. The systematic simulations in this study were assisted by OACIS [38]. J.T. acknowledges financial support from the European Union and the European Social Fund through project FuturICT.hu (Grant No. TAMOP-4.2.2.C-11/1/KONV-2012-0013). J.K. acknowledges support from EU Grant No. FP7 317532 (MULTIPLEX).

APPENDIX: TRIAL WITH OTHER PARAMETERS

We tested other parameters for the copy-and-shuffle model in order to verify the results are robust against the change of parameters. Figure 9 shows the results when the parameters $p_\triangle$, p_r, and p_d are modified from the ones used in Sec. II B. All the tested results are qualitatively similar to Fig. 4: Δf_c decreases to zero more quickly than the increase in $\overline{c}/\overline{c}_0$ when p is increased.

[1] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne, Science **323**, 721 (2009).
[2] P. Sen and B. K. Chakrabarti, *Sociophysics: An Introduction* (Oxford University Press, Oxford, 2013).
[3] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications* (Cambridge University Press, Cambridge, 1994).
[4] G. Kossinets and D. J. Watts, Science **311**, 88 (2006).
[5] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, M. A. De Menezes, K. Kaski, A.-L. Barabási, and J. Kertész, New J. Phys. **9**, 179 (2007).
[6] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási, Proc. Nat. Acad. Sci. **104**, 7332 (2007).
[7] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, Nature (London) **453**, 779 (2008).
[8] R. K. Pan, M. Kivelä, J. Saramäki, K. Kaski, and J. Kertész, Phys. Rev. E **83**, 046112 (2011).
[9] M. Szell, R. Lambiotte, and S. Thurner, Proc. Nat. Acad. Sci. **107**, 13636 (2010).
[10] M. E. Newman, Proc. Nat. Acad. Sci. **98**, 404 (2001).
[11] G. Menichetti, D. Remondini, P. Panzarasa, R. J. Mondragón, and G. Bianconi, PLOS ONE **9**, e97857 (2014).
[12] R. Lambiotte, V. D. Blondel, C. de Kerchove, E. Huens, C. Prieur, Z. Smoreda, and P. Van Dooren, Physica A **387**, 5317 (2008).
[13] M. S. Granovetter, Am. J. Sociol. **78**, 1360 (1973).
[14] J. M. Kumpula, J.-P. Onnela, J. Saramäki, K. Kaski, and J. Kertész, Phys. Rev. Lett. **99**, 228701 (2007).
[15] H.-H. Jo, R. K. Pan, and K. Kaski, PLOS ONE **6**, e22687 (2011).
[16] S. Fortunato, Phys. Rep. **486**, 75 (2010).
[17] M. E. Newman, Proc. Nat. Acad. Sci. **103**, 8577 (2006).
[18] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, J. Stat. Mech. (2008) P10008.
[19] M. Rosvall and C. T. Bergstrom, Proc. Nat. Acad. Sci. **105**, 1118 (2008).
[20] G. Palla, A.-L. Barabási, and T. Vicsek, Nature (London) **446**, 664 (2007).
[21] A. Lancichinetti, S. Fortunato, and J. Kertész, New J. Phys. **11**, 033015 (2009).
[22] Y.-Y. Ahn, J. P. Bagrow, and S. Lehmann, Nature (London) **466**, 761 (2010).

[23] H.-H. Jo, M. Karsai, J. Karikoski, and K. Kaski, EPJ Data Science **1**, 10 (2012).
[24] H.-H. Jo, R. K. Pan, J. I. Perotti, and K. Kaski, Phys. Rev. E **87**, 062131 (2013).
[25] M. Kivelä, A. Arenas, M. Barthelemy, J. Gleeson, Y. Moreno, and M. Porter, J. Complex Netw. **2**, 203 (2014).
[26] S. Boccaletti, G. Bianconi, R. Criado, C. I. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, Phys. Rep. **544**, 1 (2014).
[27] H.-H. Jo, S. K. Baek, and H.-T. Moon, Physica A **361**, 534 (2006).
[28] Significant difference was not found even if we took the maximum of the link weights instead of the sum. This is because the majority of the links belong to one of the layers.
[29] Of course, the resulting network has a topology different from an ER random graph as it has high clustering by construction.
[30] We get even for $p = 0$ a value larger than 1 ($\bar{c}_0 = 2.96$). This is due to the fact that the method of [22] is a partition of the link graph producing a large number of single links (diads) as communities, which enhance the overlap.
[31] J.-P. Onnela, S. Arbesman, M. C. González, A.-L. Barabási, and N. A. Christakis, PLOS ONE **6**, e16939 (2011).
[32] G. Krings, F. Calabrese, C. Ratti, and V. D. Blondel, J. Stat. Mech. (2009) L07003.
[33] P. Expert, T. S. Evans, V. D. Blondel, and R. Lambiotte, Proc. Nat. Acad. Sci. **108**, 7663 (2011).
[34] K. Kosmidis, S. Havlin, and A. Bunde, Europhys. Lett. **82**, 48005 (2008).
[35] L. Daqing, K. Kosmidis, A. Bunde, and S. Havlin, Nature Phys. **7**, 481 (2011).
[36] R. Parshani, C. Rozenblat, D. Ietri, C. Ducruet, and S. Havlin, Europhys. Lett. **92**, 68002 (2010).
[37] N. Eagle and A. Pentland, Personal and Ubiquitous Computing **10**, 255 (2006).
[38] Y. Murase, T. Uchitane, and N. Ito, Physics Procedia **57**, 73 (2014).