PERCOLATION TRANSITION WITH CHANGED HIDDEN VARIABLES IN COMPLEX NETWORKS

Rui Yang[a], Huibin Li[a,†], Zhanli Zhang[a,b]

[a]School of Statistics and Mathematics
Central University of Finance and Economics, 100081 Beijing, China
[b]Key Laboratory of Mathematics, Information and Behavioral Semantics (LMIB)
Ministry of Education, 100191 Beijing, China

(Received March 15, 2016; revised version received July 14, 2016)

We study a class of percolation models in complex networks in which nodes are characterized by changed hidden variables reflecting the variational properties of nodes, and the occupied probability of each link is determined by the hidden variables of the end nodes at each time. By the mean field theory, we find analytical expressions for the phase of percolation transition. It is determined by the distribution of the hidden variables for the nodes and the occupied probability between pairs of them. Moreover, the analytical expressions obtained are checked by means of numerical simulations on a particular model. Besides, the general model can be applied to describe and control practical diffusion models, such as disease diffusion model, scientists cooperation networks, and so on.

DOI:10.5506/APhysPolB.47.2195

1. Introduction

Over the past years, the study of complex networks has emerged as an important tool to better understand many social, technological, and biological real-world systems ranging from communication networks like the Internet to cellular networks [1–7]. An important question regarding networks is the percolation phenomenon [8–13] which is motivated by many applications in real networks such as epidemic spreading in social networks [14, 15].

The theory of percolation applied to random networks has been proven to be one of the greatest advances in complex network science [16–20]. A network may undergo a phase transition as nodes or links are successively occupied [13, 16]. When the fraction of occupied nodes or links is greater than a threshold value, the occupied nodes or links form a giant component

† Corresponding autor: `hli_cufe@cufe.edu.cn`

of the network. On the contrary, the giant component disappears when the fraction of occupied nodes or links is less than the threshold value. The statement of the percolation phenomenon [21] is simple: in node percolation, every node is independently either occupied with probability p, or not occupied with probability $1-p$. The occupied nodes form continuous components which have some interesting properties in real networks. In particular, the system shows a continuous phase transition at a finite value of p which is characterized by the formation of a component large enough to span the whole system from one node to the others in the limit of infinite system size or the scale of the component is almost as the scale of the whole system, $\mathcal{O}(N)$. We say that such a system percolates for this value of p or the percolation transition takes place in this system. As the percolation transition is approached from small values of p, the average component size diverges in a way reminiscent of the divergence of fluctuations in the approach to a thermal continuous phase transition and, indeed, one can define correlation functions and a correlation length in the obvious fashion for percolation models, and hence measure critical exponents for the transition [22–24].

Besides, there is a deep link percolation in which the links of the lattice are occupied (or not) with probability p (or $1 - p$). This system shows qualitatively similar behavior though different in some details from node percolation. That is, the occupied links form a giant component when the occupied probability p is greater than a threshold value. In the opposite case, the giant component disappears and all occupied links disintegrate into small components.

In the past several years, explosive percolation in networks has been proposed and heavily studied [25, 26]. Potential edges, sampled uniformly at random from the complete graph, are considered one at a time and either added to the graph or rejected provided that the fraction of accepted edges is never smaller than a decreasing function asymptotically approaching the value of $\alpha = 1/2$. It is shown that multiple giant components appear simultaneously in a strongly discontinuous percolation transition and remain distinct. Furthermore, tuning the value of α determines the number of such components with smaller α leading to an increasingly delayed and more explosive transition [27–29].

In these examples, the occupied probability is the same for every node or link. However, it is not necessarily the same for the nodes or links. As an example, in the disease diffusion process, the probability of person infected is different for the immunity of persons, and so on. Then, how to control the diffusion process with different occupied probabilities for different links, which is decided by the properties of the end nodes of links, is utmost important to control the diffusion process in real networks, such as to control disease diffusing in social networks.

In this paper, we investigate a class of percolation models with changed hidden variables on nodes in complex networks. In this percolation model, each node is assigned according to its property with a hidden variable of time, which is independently drawn from some probability distribution, and each link is occupied with some probability related to the hidden variables (or the properties) of the end nodes. Armed with the mean field theory, the analytical expressions for the phase transition of this percolation model is obtained, which is determined by the distribution of the hidden variables of nodes and the occupied probability for the links. In the end, the theoretical expressions for the phase transition of this percolation model are checked by means of numerical simulations on a particular networks.

The paper is organized as follows. In Sec. 2, we introduce the percolation with changed hidden variables on nodes in complex networks and deduce the theoretical condition for the percolation transition model. In Sec. 3, we simulate the model on some special networks and obtain the numerical results which dovetail into the theoretical results perfectly. The conclusion is given in Sec. 4.

2. Percolation with hidden variable model on complex networks

We define the class of percolation transition model in complex networks with changed hidden variables on nodes as follows. Let us consider a connected undirected network with N nodes, where $N \gg 1$. The percolation model in this network is generated by the following rules.

(1) Each node is assigned with a hidden variable $h_i(t)$ at time t, which is independently drawn from a probability distribution $\rho(h)$ with $h \geq 0$.

(2) At time t, for each pair of nodes (i, j) whose hidden variables are $h_i(t)$ and $h_j(t)$, the edge (i, j) is occupied with probability $r(h_i(t), h_j(t))$ (*the occupied probability*), where $r(h_i(t), h_j(t)) \geq 0$ is a symmetric function of h_i and h_j.

That is, given a probability distribution $\rho(h)$ and the symmetric occupied probability function $r(x, y)$, the percolation transition model with hidden variables is determined.

For the generated mechanism, the average number of occupied edges on a node at time t with hidden variable h is [30]

$$k_h(t) = N \int_0^\infty \rho\left(h'(t)\right) r(h(t), h'(t)) \mathrm{d}h'(t)\,, \tag{1}$$

and the average number of occupied edges on a node globally is

$$\begin{aligned}\langle k(t)\rangle &= \int_0^\infty k_h(t)\rho(h(t))\mathrm{d}h(t)\\ &= N\int_0^\infty\int_0^\infty \rho(h(t))\rho(h'(t))r(h(t),h'(t))\mathrm{d}h'(t)\mathrm{d}h(t)\,,\end{aligned} \tag{2}$$

which illustrates that at time t, the average degree is directly determined by the probability distribution of hidden variables on nodes and the occupied probability for each link.

To reveal the size distribution of the occupied component at time t, we start with a single occupied vertex by revealing its occupied neighbors following occupied edges, then go with their neighbors, *etc.* [31]. Let n_l be the number of nodes exposed for the first time in step l of this revealing process. Given the previous numbers $n_0 = 1, n_1, \ldots, n_{l-1}$, the distribution of n_l is

$$P(n_l) = \binom{N-\sum_{l=0}^{l-1} n_s}{n_l}(1-q^{n_{l-1}})^{n_l}(q^{n_{l-1}})^{N-\sum_{s=0}^{l} n_s}\,, \tag{3}$$

where q is the probability that a node is disconnected to the n_{l-1} nodes which are exposed in step $l-1$ on average, and that

$$\begin{aligned} q &= p\,(x \text{ is disconnected to } y|\text{ the hidden variables of}\\ &\qquad \text{node in } n_l, n_{l-1} \text{ is } x, y \text{ respectively at time } t)\\ &= \textstyle\int\int \rho(x)\rho(y)(1-r(x,y))\mathrm{d}x\mathrm{d}y\\ &= \textstyle\int\int \rho(x)\rho(y)\mathrm{d}x\mathrm{d}y - \int\int \rho(x)\rho(y)r(x,y)\mathrm{d}x\mathrm{d}y\\ &= 1-\frac{\langle k(t)\rangle}{N}\,.\end{aligned} \tag{4}$$

In the large N limit with fixed $\langle k(t)\rangle$, $p(n_l)$ tends to $e^{-n_{l-1}\langle l\rangle}(n_{l-1}\langle k\rangle)^{n_l}/n_l!$. Thus, the revealing process reduces to a Poisson branching tree model, with each node independently branching to a number of new nodes, where this number is a Poisson random variable with average $\langle k\rangle$. The distribution p_n over the order n of the resulting tree is conveniently analyzed by the generating function $F(z) = \sum_n p_n z^n$, which satisfies

$$F(z) = z\exp\left[\langle k\rangle(F(z)-1)\right]\,. \tag{5}$$

In Ref. [31], we get that the transition point for this model is

$$\langle k\rangle = 1\,. \tag{6}$$

Thus, when $\langle k \rangle > 1$, the occupied links and nodes form a giant component of the network, while when $\langle k \rangle < 1$, the giant component disappears.

For

$$\langle k(t) \rangle = N \int_0^\infty \int_0^\infty \rho(h(t))\rho\left(h'(t)\right) r\left(h(t), h'(t)\right) \mathrm{d}h(t)\mathrm{d}h(t)\,,$$

we have

$$N \int_0^\infty \int_0^\infty \rho(h(t))\rho\left(h'(t)\right) r(h(t), h'(t))\mathrm{d}h'(t)\mathrm{d}h(t) = 1\,. \tag{7}$$

So if $N \int_0^\infty \int_0^\infty \rho(h(t))\rho(h'(t))r(h(t), h'(t))\mathrm{d}h'(t)\mathrm{d}h(t) > 1$ at time t, the giant component of the occupied edge takes place and the percolation happens; while if $N \int_0^\infty \int_0^\infty \rho(h(t))\rho(h'(t))r(h(t), h'(t))\mathrm{d}h'(t)\mathrm{d}h(t) < 1$ at time t, the occupied edges are all small clusters whose scales are far smaller than the size of the whole network.

3. Numerical simulations on networks

For applications, we simulate the model on a network with $N = 5000$ nodes. In this model, each node is assigned with a hidden variable h, which is independently drawn from the probability distribution $\rho(h(t)) = \lambda t e^{-\lambda h t}$ with exponent parameter λ for $h \geq 0$. For each pair of nodes (i, j) whose hidden variables are $h_i(t)$ and $h_j(t)$, respectively, at time t, the edge (i, j) is occupied with probability $r(h_i, h_j) = \Theta(h_i + h_j - c)$, where

$$\Theta(x) = \begin{cases} 1\,, & x > 0\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

In this model, at time t, the degree distribution $p(k)$ is [32]

$$p(k) = N e^{-\lambda c t} \frac{1}{k^2} \theta_k \left(N e^{-\lambda c t}, N\right) + e^{-\lambda c t} \delta(k - N)\,,$$

where $\delta(x)$ is the Dirac function, $\theta_x(a, b)$ is

$$\theta_x(a, b) = \begin{cases} 1\,, & a \leq x \leq b\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

That is, the networks simulated by this model exhibit a scale-free degree distribution, with degree exponent $\gamma = 2$, for degrees in the range of $N e^{-\lambda c t} \leq k \leq N$, with an accumulation point at $k = N$, given by the δ function, with weight $e^{-\lambda c t}$.

Substituting $\rho(h(t))$ and $r(h,h')$ into equation (6), we get

$$N\int_0^\infty\int_0^\infty \lambda t e^{-\lambda ht}\lambda t e^{-\lambda h' t}\Theta\left(h+h'-c\right)\mathrm{d}h'\mathrm{d}h = 1\,,$$

that is

$$\langle k(t)\rangle = N\int_0^\infty\int_0^\infty \lambda t e^{-\lambda ht}\lambda t e^{-\lambda h' t}\Theta\left(h+h'-c\right)\mathrm{d}h'\mathrm{d}h\,.$$

Integrating it, we have

$$\langle k(t)\rangle = Ne^{-\lambda ct}(1+\lambda ct)\,.$$

The transition point for this model is

$$\langle k(t)\rangle = Ne^{-\lambda ct}(1+\lambda ct) = 1\,,$$

hence, we can get the relationship among the exponential distribution parameter λ, the window parameter c and time t.

In the following, we give the numerical simulations of this model.

Especially, given $c=\ln N$, the percolation transition point for λ satisfies

$$\langle k(t)\rangle = Ne^{-\lambda t\ln N}(1+\lambda\ln N) = 1\,.$$

That is

$$N^{1-\lambda t}(1+\lambda\ln N) = 1\,.$$

In Fig. 1, with $c=\ln N$, $t=1$, the scale of the largest component as a function of the mean value $\frac{1}{\lambda}$ of the exponent distribution $\rho(h,t)=\lambda te^{-\lambda th}$ is shown. It indicates that percolation takes place in the network when the average value $\frac{1}{\lambda}$ of the exponent distribution $\rho(h)$ exceeds a certain value. For the fixed window parameter $c=\ln N$, as the mean value of the hidden variables for nodes increases, the occupied probability for each link is increasing, and then the number of nodes in the largest component is increasing. Besides, as the mean value exceeds a threshold value, the giant component takes place, as it is shown by the subgraph in Fig. 1. From the detail in the subgraph of Fig. 1, we know that the percolation transition point λ_z belongs to the interval $[0.6, 0.9]$ in this model.

In Fig. 2, with $c=\ln N$, $t=1$, we give the average degree $\langle k\rangle$ as a function of the mean value $\frac{1}{\lambda}$ of the exponent distribution $\rho(h,t)=\lambda te^{-\lambda th}$. It shows that the average degree $\langle k\rangle$ is increasing with the increase of the mean value $\frac{1}{\lambda}$ of the exponent distribution. It is because the increase of

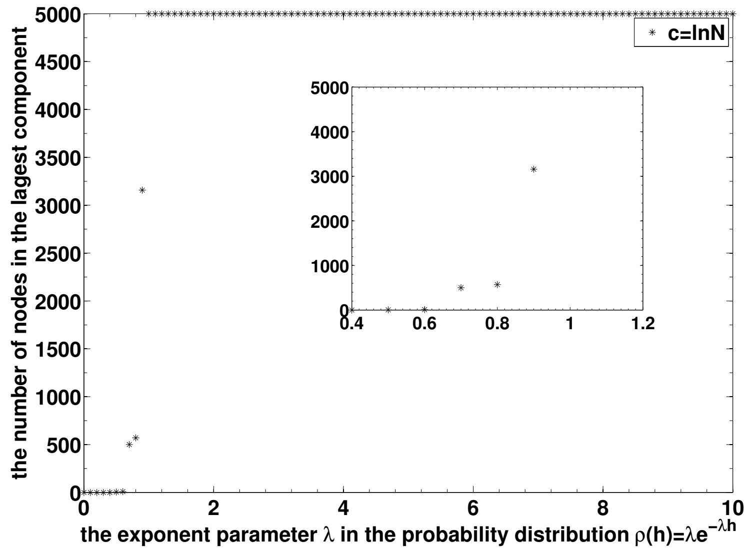

Fig. 1. The largest component as a function of the exponent parameter λ is shown in the main graph, and the details for $0.4 \leq \lambda \leq 1.2$ are shown in the subgraph.

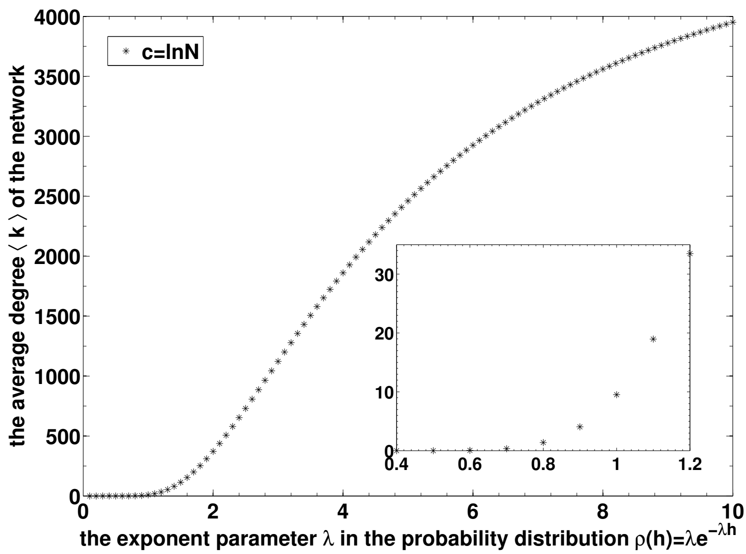

Fig. 2. The average degree $\langle k \rangle$ as a function of the exponent parameter λ is shown in the main graph, and the subgraph shows the details for $0.4 \leq \lambda \leq 1.2$.

mean value of the hidden variables for nodes results in the increase of the average degree of the network. Furthermore, from the subgraph in Fig. 2, we find that the transition point is $\langle k \rangle = 1$ with $\lambda_z \in [0.6, 0.9]$, which matches the numerical simulations in Fig. 2 perfectly.

Actually, since

$$N^{1-\lambda}(1 + \lambda \ln(N)) = 1\,,$$

we know that $\lambda < 1$ must be satisfied.

In Fig. 3, fixing $\lambda = 1$, $t = 1$, we give the scale of the largest component as a function of the window parameter c. According to our rule, when the time t is fixed, the distribution of node's hidden variable is fixed. With the increase of the parameter c, the occupied probability for each link is decreasing, and then the number of nodes in the largest component is decreasing. Besides, as the window parameter c exceeds a threshold value, the giant component disappears, which is shown in the subgraph of Fig. 3. From the details in the subgraph of Fig. 3, we know that the percolation transition point c_z belongs to the interval $[8, 14]$.

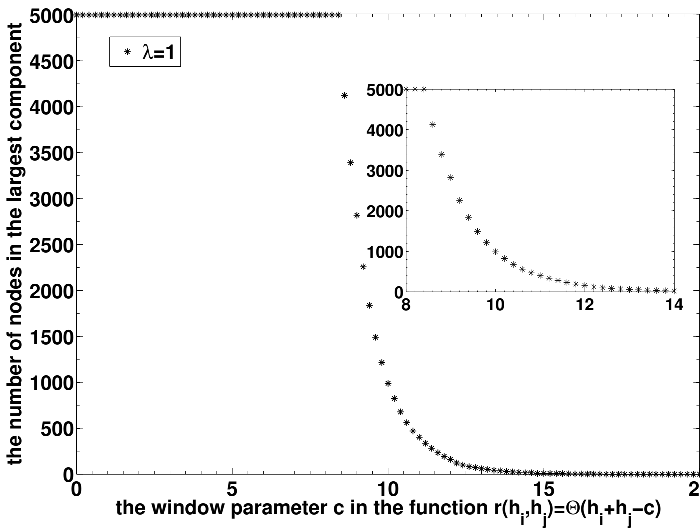

Fig. 3. The largest component as a function of the window parameter c is shown in the main graph, and the subgraph shows the details for $8 \leq c \leq 14$.

In Fig. 4, for $\lambda = 1$, $t = 1$, we give the average degree $\langle k \rangle$ as a function of parameter c. As increase of the window parameter c results in the decrease of the occupied probability according to $r(h_i, h_j) = \Theta(h_i + h_j - c)$, the average degree $\langle k \rangle$ decreases with the increase of c. Furthermore, from the subgraph in Fig. 4, we find that the transition point is $\langle k \rangle = 1$ with $c_z \in [8, 14]$, which matches the numerical simulations in Fig. 4 perfectly.

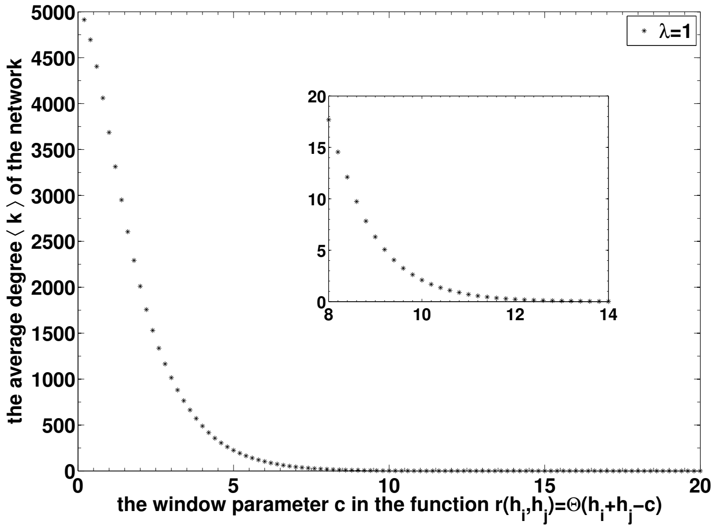

Fig. 4. The average degree $\langle k \rangle$ as a function of the window parameter c is shown in the main graph, and the subgraph shows the details for $8 \leq c \leq 14$.

From the above numerical simulations, which match the theoretical expressions perfectly, we know that the percolation transition takes place in our new class of percolation model.

4. Conclusion

In summary, we study a new class of percolation model in complex networks with hidden variables on nodes. In this model, each node is assigned with hidden variable which represents the property of the node, and each link is occupied with some probability based on the hidden variables of the end nodes. With the mean field theory, we derive the theoretical condition for the appearance percolation transition for this model above which the occupied edges form a giant component of the network, while below it, the giant component disappears and all occupied links disintegrate into small components. As applications, we take a special hidden variable distribution and a special occupied function as an example to check our model. It matches the theoretical results perfectly.

This work is supported by the National Natural Foundation of China, grant Nos. 71503292, 11401602, and 11472315.

REFERENCES

[1] R. Albert, A.L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).

[2] S. Yoon, S.H. Yook, Y. Kim, *Phys. Rev. E* **76**, 056104 (2007).

[3] Y.-Y. Liu, E. Csóka, H. Zhou, M. Pósfai, *Phys. Rev. Lett.* **109**, 205703 (2012).

[4] T. Nogawa, T. Hasegawa, *Phys. Rev. E* **89**, 042803 (2014).

[5] D.J. Watts, S.H. Strogatz, *Nature (London)* **393**, 440 (1998).

[6] A.L. Barabási, R. Albert, *Science* **286**, 509 (1999).

[7] S.N. Dorogovtsev, J.F.F. Mendes, *Adv. Phys.* **51**, 1079 (2002).

[8] P. Grassberger, *Phys. Rev. E* **91**, 062806 (2015).

[9] E. Massaro, F. Bagnoli, *Phys. Rev. E* **90**, 052817 (2014).

[10] G.J. Baxter, S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, *Phys. Rev. E* **83**, 051134 (2011).

[11] M. Cuquet, J. Calsamiglia, *Phys. Rev. A* **83**, 032319 (2011).

[12] G.J. Baxter, S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, *Phys. Rev. E* **82**, 011103 (2010).

[13] M.E.J. Newman, R.M. Ziff, *Phys. Rev. E* **64**, 016706 (2001).

[14] Y. Chen *et al.*, *Phys. Rev. E* **75**, 046107 (2007).

[15] M. Boguñá, R.P. Satorras, *Phys. Rev. E* **66**, 047104 (2002).

[16] J.D. Noh, *Phys. Rev. E* **76**, 026116 (2007).

[17] P. Colomer-de-Simón, M. Boguñá, *Phys. Rev. X* **4**, 041020 (2014).

[18] H. Hooyberghs, B. Van Schaeybroeck, *Phys. Rev. E* **83**, 032101 (2011).

[19] R.A. da Costa, S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, *Phys. Rev. E* **91**, 032140 (2015).

[20] Y. Berezin, A. Bashan, S. Havlin, *Phys. Rev. Lett.* **111**, 189601 (2013).

[21] D. Stauffer, A. Aharony, *Introduction to Percolation Theory*, 2$^{\rm nd}$ edition, Taylor and Francis, London 1992.

[22] J. Fan, M. Liu, L. Li, X. Chen, *Phys. Rev. E* **85**, 061110 (2012).

[23] Y. Hu, B. Ksherim, R. Cohen, S. Havlin, *Phys. Rev. E* **84**, 066116 (2011).

[24] R. Parshani, S.V. Buldyrev, S. Havlin, *Phys. Rev. Lett.* **105**, 048701 (2010).

[25] R.M. D'Souza, J. Nagler, *Nature Physics* **11**, 531 (2015).

[26] D. Achlioptas, R.M. D'Souza, J. Spencer, *Science* **323**, 1453 (2009).

[27] W. Chen, R.M. D'Souza, *Phys. Rev. Lett.* **106**, 115701 (2011).

[28] W. Chen, Z. Zheng, R.M. D'Souza, *Europhys. Lett.* **100**, 66006 (2011).

[29] W. Chen *et al.*, *Phys. Rev. E* **87**, 052130 (2013).

[30] G. Caldarelli, A. Capocci, P.D.L. Rios, M.A. Munñoz, *Phys. Rev. Lett.* **89**, 258702 (2002).

[31] B. Söderberg, *Phys. Rev. E* **66**, 066126 (2002).

[32] M. Boguñá, R.P. Satorras, *Phys. Rev. E* **68**, 036112 (2003).