International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:4, No:12, 2010

Percolation Transition with Hidden Variables in Complex Networks

Zhanli Zhang, Wei Chen, Xin Jiang, Lili Ma, Shaoting Tang and Zhiming Zheng

Abstract—A new class of percolation model in complex networks, in which nodes are characterized by hidden variables reflecting the properties of nodes and the occupied probability of each link is determined by the hidden variables of the end nodes, is studied in this paper. By the mean field theory, the analytical expressions for the phase of percolation transition is deduced. 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.

Keywords—complex networks, percolation transition, hidden variable, occupied probability.

I. 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], [2], [3], [4], [5], [6]. An important question regarding networks is the percolation phenomenon[7], [8], [9], [10], [11], [12], [13] which is motivated by many applications in real networks such as epidemic spreading in social networks[14], [15], [16].

The theory of percolation applied to random networks has been proven to be one of the most notorious advances in complex network science[17], [18], [19], [20], [21], [22]. A network may undergo a phase transition as nodes or links are successively occupied[13], [17]. When the fraction of occupied nodes or links is greater than a threshold value, the occupied nodes or links form a giant component of the network. By contraries, the giant component disappears when the fraction of occupied nodes or links is less than the threshold value. The statement of the percolation phenomenon[23] 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 contiguous 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 other 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 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.

Besides, link percolation is also deeply researched in which the links of the lattice are occupied (or not) with probability p (or $1-p$). This system shows behavior qualitatively similar to, 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.

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

In this paper, a new class of percolation model with hidden variables on nodes in complex networks is investigated. In this percolation model, according to the various properties of the nodes, each node is assigned with a hidden variable, 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, the percolation with hidden variables on nodes in complex networks is introduced and the theoretical condition for the percolation transition model is deduced. In Sec.3, this model is simulated on some special networks and the numerical results which dovetail into the theoretical results perfectly is achieved. The conclusion is given in Sec.4.

Zhanli Zhang and Wei Chen are with the School of Mathematical Sciences, Peking University, 100871, China, e-mail: zhangzhl@pku.edu.cn.
Xin Jiang, Lili Ma, Shaoting Tang and Zhiming Zheng are with the Key Laboratory of Mathematics, Information and Behavioral Semantics (LMIB), Ministry of Education, 100191, China.

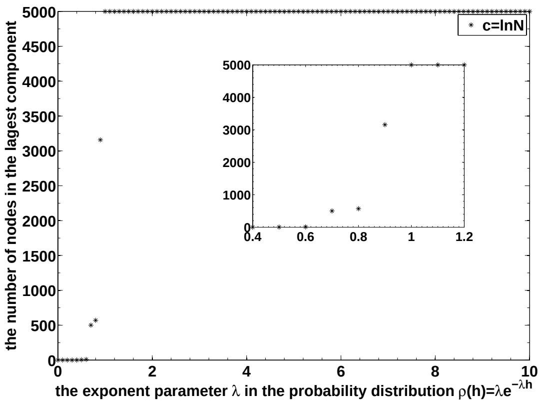

Fig. 1. The largest component as a function of the exponent parameter λ is shown in the main graph, and the subgraph is the details for $0.4 \leq \lambda \leq 1.2$.

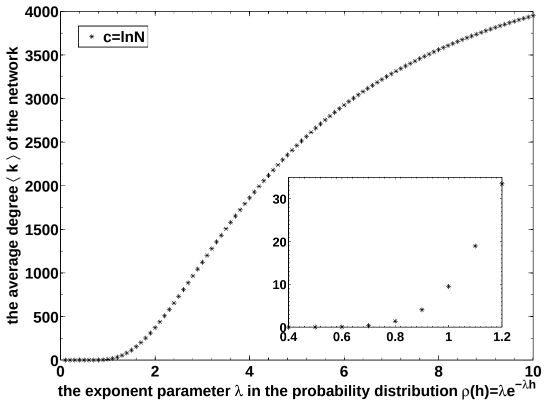

Fig. 2. The average degree as a function of the exponent parameterλ is shown in the main graph, and the subgraph is the details for $0.4 \leq \lambda \leq 1.2$.

II. Percolation with Hidden Variable Model on Complex Networks

The class of percolation transition model in complex networks with hidden variables on nodes is defined 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, which is independently drawn from a probability distribution $\rho(h)$ with $h \geq 0$.

(2) For each pair of nodes (i,j) whose hidden variables are h_i and h_j, the edge (i,j) is occupied with probability $r(h_i,h_j)$ (*the occupied probability*), where $r(h_i,h_j) \geq 0$ 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 incident on a node with hidden variable h is[24]

$$k(h) = N\int_0^{\infty} \rho(h')r(h,h')dh', \tag{1}$$

and the average number of occupied edged incident on a node globally is

$$\langle k\rangle = \int_0^{\infty} k(h)\rho(h)dh = N\int_0^{\infty}\int_0^{\infty} \rho(h)\rho(h')r(h,h')dh'dh \tag{2}$$

which illustrates that 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, start with a single occupied vertex, reveal its occupied neighbors following occupied edges, then their neighbors, etc.[25]. Let n_t be the number of nodes exposed for the first time in step t of this revealing process. Given the previous numbers $n_0 = 1, n_1, \cdots, n_{t-1}$, the distribution of n_t is

$$P(n_t) = \begin{pmatrix} N - \sum_{l=0}^{t-1} n_l \\ n_t \end{pmatrix} (1-q^{n_{t-1}})^{n_t}(q^{n_{t-1}})^{N-\sum_{l=0}^{t} n_l}, \tag{3}$$

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

$$\begin{aligned} q &= p(x \text{ is disconnected to y}|\text{ the hidden variables of} \\ &\qquad \text{node in } n_t, n_{t-1} \text{ is } x, y \text{ respectivelly}) \\ &= \int\int \rho(x)\rho(y)(1-r(x,y))dxdy \\ &= \int\int \rho(x)\rho(y)dxdy - \int\int \rho(x)\rho(y)r(x,y)dxdy \\ &= 1 - \frac{\langle k\rangle}{N} \end{aligned} \tag{4}$$

In the large N limit with fixed $\langle k\rangle$, $p(n_t)$ tends to $e^{-n_{t-1}\langle t\rangle}(n_{t-1}\langle k\rangle)^{n_t}/n_t!$. 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

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

For ref. [25], 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

$$\langle k\rangle = N\int_0^{\infty}\int_0^{\infty} \rho(h)\rho(h')r(h,h')dh'dh,$$

thus

$$N\int_0^{\infty}\int_0^{\infty} \rho(h)\rho(h')r(h,h')dh'dh = 1 \tag{7}$$

So if $N\int_0^{\infty}\int_0^{\infty} \rho(h)\rho(h')r(h,h')dh'dh > 1$, the giant component of the occupied edge takes place and the percolation happens; while if $N\int_0^{\infty}\int_0^{\infty} \rho(h)\rho(h')r(h,h')dh'dh < 1$, the occupied edge are all small clusters whose scales are far smaller than the size of the whole network.

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:4, No:12, 2010

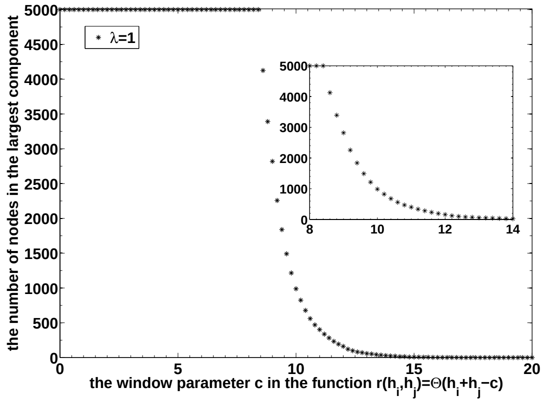

Fig. 3. The largest component as a function of the window parameter c is shown in the main graph, and the subgraph is the details for $8 \le c \le 14$.

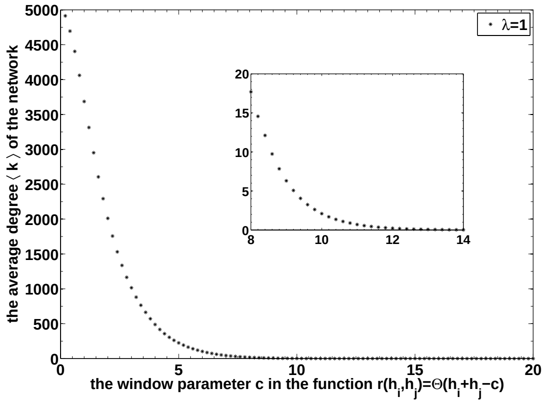

Fig. 4. The average degree $\langle k\rangle$ as a function of the window parameter c is shown in the main graph, and the subgraph is the details for $8 \le c \le 14$.

III. Numerical Simulations on Networks

For applications, the model is simulated 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 $\rho(h) = \lambda e^{-\lambda h}$ with exponent parameter λ for $h \ge 0$. For each pair of nodes (i, j) whose hidden variables are h_i and h_j respectively, 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, the degree distribution $p(k)$ is[26]

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

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

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

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

Submitting $\rho(h)$ and $r(h, h')$ into equation (6), then

$$N\int_0^\infty \int_0^\infty \lambda e^{-\lambda h}\lambda e^{-\lambda h'}\Theta(h + h' - c)dh'dh = 1,$$

that is

$$\langle k\rangle = N\int_0^\infty \int_0^\infty \lambda e^{-\lambda h}\lambda e^{-\lambda h'}\Theta(h + h' - c)dh'dh.$$

After the integral,

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

The transition point for this model is

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

and then the relationship between the exponential distribution parameter λ and the window parameter c is achieved.

The numerical simulations of this model is in the following.

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

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

That is

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

In Fig. 1, with $c = \ln N$, the scale of the largest component as a function of the mean value $\frac{1}{\lambda}$ of the exponent distribution $\rho(h) = \lambda e^{-\lambda h}$ 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 increase mean value of the hidden variables for nodes, 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 the subgraph shown in Fig. 1. From the detail in the subgraph of Fig. 1, the result shows that the percolation transition point λ_t is $\lambda_t \in [0.6, 0.9]$ in this model.

In Fig. 2, with $c = \ln N$, the average degree $\langle k\rangle$ as a function of the mean value $\frac{1}{\lambda}$ of the exponent distribution $\rho(h) = \lambda e^{-\lambda h}$ is given. 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 that the increase of 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, the result shows that the transition point is $\langle k\rangle = 1$ with $\lambda_t \in [0.6, 0.9]$, which matches the numerical simulations in Fig. 1 perfectly.

Actually, because

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

then $\lambda < 1$ must be satisfied.

In Fig. 3, fixing $\lambda = 1$, the scale of the largest component as a function of the window parameter c is given. For the distribution of node 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 detail in the subgraph of Fig. 3, the result shows that the percolation transition point c_t is $c_t \in [8, 14]$.

In Fig. 4, for $\lambda = 1$, the average degree $\langle k \rangle$ as a function of parameter c is given. 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 increases of c. Furthermore, from the subgraph in Fig. 4, the results shows that the transition point is $\langle k \rangle = 1$ with $c_t \in [8, 14]$, which matches the numerical simulations in Fig. 3 perfectly.

From the above numerical simulations which matches the theoretical expressions perfectly, the results show that the percolation transition takes place in our new class of percolation model.

IV. Conclusion

In summary, a new class of percolation model in complex networks with hidden variables on nodes is studied. 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, the theoretical condition for the appearance percolation transition for this model is derived, above which the occupied edges forms a giant component of the network, while below which, the giant component disappears and all occupied links disintegrate into small components. As applications, a special hidden variable distribution and a special occupied function are taken as an example to check our model, which matches the theoretical results perfectly. However, the hidden variables for nodes can be changed with the variety environment of the network, which has been not considered in our current work but will be in our future work.

Acknowledgments

This work is supported by the National Key Basic Research Project of China Grant No.2005CB321902.

References

[1] R. Albert and A. L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2] S. Yoon, S. H. Yook, and Y. Kim, Phys. Rev. E **76**, 056104 (2007).
[3] D. J. Watts, and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[4] A. L. Barabási, and R. Albert, Science **286**, 509(1999)
[5] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).
[6] S. N. Dorogovtsev, and J. F. F. Mendes, Adv. Phys. **51**, 1079 (2002).
[7] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
[8] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, Phys. Rev. E **78**, 051105 (2008).
[9] M. Á. Serrano and P. D. L. Rios, Phys. Rev. E **76**, 056121 (2007).
[10] A. Zen, A. Kabakçıoğlu, and A. L. Stella, Phys. Rev. E **76**, 016110 (2007).
[11] E. Lóez, R. Parshani, R. Cohen, S. Carmi, and S. Havlin, Phys. Rev. Lett. **99**, 188701 (2007).
[12] J. D. Noh, Phys. Rev. E **76**, 026116 (2007).
[13] M. E. J. Newman and R. M. Ziff, Phys. Rev. E **64**, 016706 (2001).
[14] R. P. Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[15] C. Moore and M. E. J. Newman, Phys. Rev. E **61**, 5678 (2000).
[16] M. Boguñá and R. P. Satorras, Phys. Rev. E **66**, 047104 (2002).
[17] J. D. Noh, Phys. Rev. E **76**, 026116 (2007).
[18] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, Phys. Rev. E **78**, 051105 (2008).
[19] A. Zen, A. Kabakcıoğlu, and A. L. Stella, Phys. Rev. E **76**, 026110 (2007).
[20] M. A. Serrano and P. D. L. Rios, Phys. Rev. E **76**, 056121(2007).
[21] E. López, R. Parshani, R. Cohen, S. Carmi, and S. Havlin, Phys. Rev. E **99**, 188701 (2007).
[22] L. Apolo, O. Melchert, and A. K. Hartmann, Phys. Rev. E **79**, 031103 (2009).
[23] D. Stauffer and A. Aharony, Introduction to Percolation Theory, 2nd edition, Taylor and Francis, London (1992).
[24] G. Caldarelli, A. Capocci, P. D. L. Rios, and M. A. Munñoz5, Phys. Rev. Lett. 89, 258702 (2002).
[25] B. Söderberg, Phys. Rev. E **66**, 066126 (2002).
[26] M. Boguñá and R. P. Satorras, Phys. Rev. E **68**, 036112 (2003).