PHASE TRANSITIONS IN PERCOLATION PHENOMENON FOR A TREE MODEL

Yuming Yang[a,b], Baohua Teng[a]

[a]School of Physical Electronics
University of Electronic Science and Technology of China
Chengdu 610054, P.R. China
[b]School of Mathematical Sciences
University of Electronic Science and Technology of China
Chengdu 610054, P.R. China

(Received May 16, 2018; accepted December 11, 2018)

The stability of a tree model is studied by randomly and preferentially selecting edges, including adding and removing edges, respectively. Firstly, the critical point of the phase transition in the percolation phenomenon is determined according to the number of connected components with different sizes. Then the feature of the phase transition is distinguished according to the distribution of the connected components with respect to a cluster size at the critical point. The Monte-Carlo numerical results show that the phase transitions in the evolving and fragmentation processes for the Erdös–Rényi tree are continuous. For the product rule tree, however, the evolving process undergoes a discontinuous phase transition, while the phase transition in the fragmentation process is continuous.

DOI:10.5506/APhysPolB.50.77

1. Introduction

The concept of percolation was first introduced by Broadbent and Hammersley in 1957 to describe the flow of fluid in disordered porous media [1]. It told us that when the density or the concentration, *i.e.*, the number of occupied bonds or sites (this paper focuses on bond percolation) within a disorder system, increases to a certain degree, the system will suddenly appear some kind of phase change behavior with long-range correlation. Percolation theory is widely used in the study of many percolation phenomena in physical, chemical, biological and social systems, such as an artificial network of Bethe lattice [2–4], the distribution of oil or gas inside porous structures in oil field, the diffusion in a disordered medium, the fire spreading over a large forest area [2, 5], the sol–gel transition [6], the epidemic spreading in a social network [7] and so on. Interest in percolation theory has focused

mainly on the topological properties of the different percolation lattices, the percolation threshold t_c, the stability against random failures of network and the applications to thermal phase transitions [8–10], *etc.*

Originally, percolation is a random process defined on the regular geometry structure. Percolation theory [1] is used to deal mathematically with the forming of a larger connected component. It means the appearance of connectivity between two opposite sides of a regular lattice system. When this concept is generalized to the case of networks, it means the emergence of a giant connected component. In this situation, one can observe the percolation phenomenon by adding edges one-by-one to connect nodes in the network.

As a more general graph ensemble, the statistical ensemble of scale-free connected random tree networks was constructed in Ref. [11]. Through calculating the fractal and the spectral dimensions, the geometry of the graphs and the stability of the scale-free regime were characterized. In fact, actual RNA secondary structures have a heterogeneous, branched tree-like form [12]. The literature [13] found that the anomalous diffusion has connection with bond percolation on random recursive trees. Many research works [12–17] on random trees show that it is important to survey the nature of phase transition of a tree model.

Explosive percolation (EP) has recently been considered on various topological structure, such as Erdös–Rényi random graphs [18], regular lattices [19, 20] and scale-free networks [19]. Meanwhile, researchers have gotten many interesting effects of the "product rule" on these networks (see details in review article [21]). The previous studies [22, 23] dealt with random trees by branching process or fragmentation process. The difference, however, between random choosing and intentional choosing of links in tree evolution and fragmentation processes, respectively, has not been addressed.

2. Descriptions of models and method

We are interested in what are the effects of the "product rule" on random tree graph, especially the effects in evolution and fragmentation process. Thus, the purpose of this paper is to distinguish the various effects of product rules on PR random tree graph, PRT process and random graph model. Product rules work differently in these models. In random tree model and PRT process, the product rule works from the beginning to the end. However, for ER graphs, because two candidate edges are intra cluster edges, selecting an edge by product rule is equivalent to randomly selecting an edge. Therefore, the product rule does not always work in random graphs model.

In previous works [12–17], there is not any limitation on the formed trees at any time step, such that any tree with N nodes may be the ultimate tree of evolution. Here, we take a different strategy in which a specific tree will

be chosen as a target tree of an evolution process. The reason for this is that fragmentation begins with a specific tree, so evolution should also be towards that specific tree. In the model, people do not need to care about the concrete structure of the target tree, except to ensure that the evolving target tree is the tree at the beginning of fragmentation.

In order to study the effects of "product rule" on random tree graphs, two percolation models are studied in detail. The first percolation model is the random evolution and fragmentation process on tree graph. The other percolation model is the preferential evolution and fragmentation process. The latter one is our main interest. There are many preferential attachment algorithms to produce networks [18, 24, 25]. Here, product rule [18] will be used as the preferential attachment algorithm in our second model.

First of all, two target trees in both models are generated with the help of Erdös–Rényi (ER) process and product rule (PR) process, respectively. The ER process starts with N isolated nodes and adds connections randomly one-by-one under the limitation of no loop. At the end, all the nodes join together to form a big tree T_{ER}. The PR process starts with N isolated nodes and sequential preferential attachment of edges by using product rule under the limitation of no loop. At the end, all the nodes join together to form another big tree T_{PR}. In both processes, any pair nodes have the possibility of joining together under the limitation of no loop. T_{ER} and T_{PR} will be taken as the target trees in our models.

Let E_1 denote the edge set of tree T_{ER}. The ERT process is, by using set E_1, to regenerate the tree. To this end, we start with N isolated nodes again. At each time step, one edge was uniformly chosen from set E_1 to connect corresponding nodes and then be removed from E_1. Here, time t is the number of edges in the graph and $\tilde{t}$ is the density of edges of the graph. We repeat this procedure until no edge exists in E_1. In its inverse process, the edges are removed randomly one-by-one from the tree in a uniformly way until all the nodes in the original tree are isolated. As edges are removed from the tree, it fragments into multiple connected components, and each of them has a tree structure. We call this process the inverse Erdös–Rényi tree (IERT) process.

Let E_2 denote the edge set of tree T_{PR}. The product rule tree (PRT) preferential evolution process is, by using set E_2, to regenerate the tree. In each evolving step, two independent candidate edges are picked firstly from edge set E_2 and each of them is uniformly picked at random. Then one edge is selected to connect two nodes in the graph to form a new subtree, and the edge picked will avoid forming a bigger subtree [18]. This edge will be removed from E_2 after connection and the other edge will remain in E_2. Again, this procedure will be repeated until no edge exists in E_2. In its inverse process, two candidate edges are uniformly chosen at random. Then

one edge is selected to remove from the tree, which will retain the bigger subtree structure of the system. We call this process the inverse product rule tree (IPRT) process. The above definition of the IPRT process is consistent with the practice of Bastas *et al.* [20].

It is notable that the tree T_{ER} is not only the target tree of ERT process, but also the initial tree of IERT processes, and the tree T_{PR} is not only the target tree of PRT process, but also the initial tree of IPRT processes. By the way, the studies of simulations tell us the degree distribution on both trees observes poisson distribution.

For further discussion, we define two numbers in the system: $C_{\max}$, the size of the largest connected component and C_{num}, the number of connected components with different sizes. $C_{\max}$ represents the number of nodes contained in a largest connected component, so it can characterize the structural integrity of the network. Generally speaking, the bigger it is, the better structural integrity the network has. When $C_{\max}$ becomes microscale with respect to N, the network lost its integrality. However, C_{num} can characterize the crossover characteristic of the system, while $C_{\max}$ from microscale to macroscale in evolution process or from macroscale to microscale in fragmentation process. In numerical calculation, the crossover begins with C_{num} reaching its maximum value for the first time, and ends with C_{num} starting to decline for the last time. The width of this domain will become very narrow when system size N goes to infinity and it will shrink to a single point. This point can locate well the presence of a phase transition, *i.e.*, the critical point. As shown in Fig. 1, it can be seen that the number C_{num} can describe well the crossover feature in the classical Erdös–Rényi and Achlioptas processes.

In the following, the numbers of connected components with different sizes for ERT and PRT percolation processes are given by Monte-Carlo simulation. It can be seen that C_{num} in ERT process has the same trajectory as that in IERT process (see Fig. 2 (a)). It indicates that for ERT and its inverse processes, the distribution of connected components sizes is dependent only on the number of edges but independent of the linking order of edges in the graph. So, both processes have the same distribution of the sizes of connected components at the same time step for ERT and IERT processes.

The situation, however, in PRT process or IPRT process is very different. C_{num} in PRT process has no obvious similarity with that in IPRT process. C_{num} has a dramatic turning point in PRT evolution process. But in IPRT process, the turning point is not obvious, and only in log–log plot of C_{num} an obvious change can be found (see Fig. 2 (b)). It indicates that the PRT process is quite different from the ERT process.

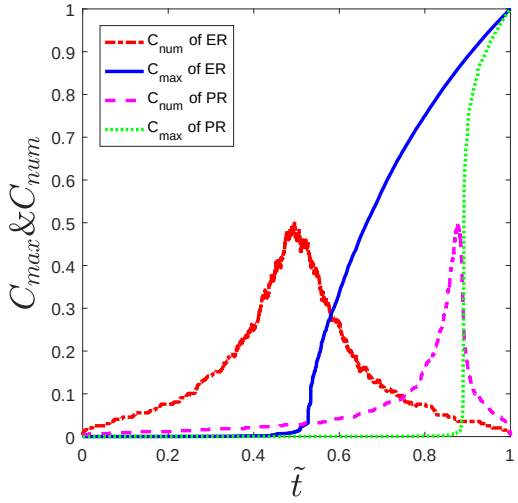
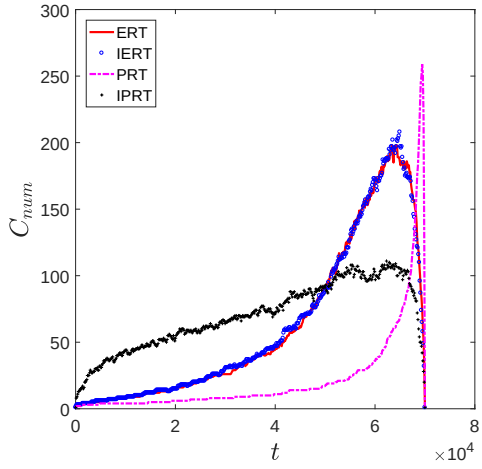
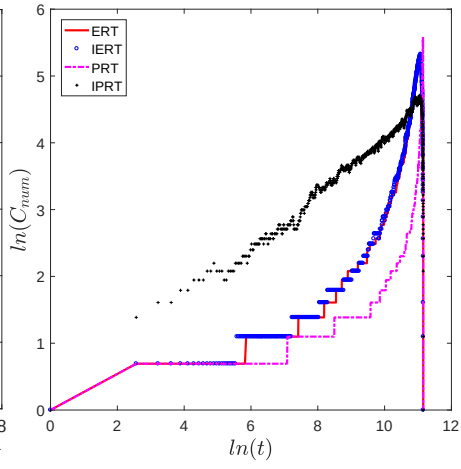


Fig. 1. C_{num} and C_{max} for classic Erdös–Rényi and Achlioptas processes. C_{num} reached its maximum value when $\tilde{t}$ approximates to 0.5 in ER process and approximates to 0.888 in Achlioptas process (or PR process). The crossover features decided by C_{num} are consistent with the results in Ref. [18]. The figure is based on one run for $N = 70\,000$.



The number of the types of clusters *versus* the time t.



The log–log plot of Fig. 2 (a) (natural logarithm). There is obvious turning point in every curve.

Fig. 2. The numbers of clusters with different sizes in various processes for $N = 70\,000$, where the number of edges existed in a graph is taken as the time step. Both sub-figures are based on one run.

For further clarification of the difference, we investigate theoretically the ERT and PRT processes as well as their inverse processes. At the time step t, e is the chosen edge joining nodes v_1 and v_2, and v_1 and v_2 are in different clusters. Let ξ and η denote the sizes of these clusters, respectively, and ζ is the size of the new formed component after edge e was added into the graph. So, we have the following probability relation:

$$P_t(\zeta = k) = \sum_{i+j=k} P_{t-1}(\xi = i, \eta = j)\,, \tag{1}$$

here $P_t(A)$ denotes the probability of the event A at the time step t. Further, let Z_v be the size of the component containing node v, then we have a recurrence relation as follows:

$$\begin{aligned} P_t(Z_v = k) \;=\; & P_{t-1}(Z_v = k) + kP_t(\zeta = k) - k\sum_{j\neq k} P_{t-1}(\xi = k, \eta = j) \\ & -k\sum_{j\neq k} P_{t-1}(\xi = j, \eta = k) - 2kP_{t-1}(\xi = k, \eta = k)\,. \end{aligned} \tag{2}$$

According to relations (1) and (2), the size of the connected components for the ERT and PRT can be calculated by Monte-Carlo simulation. Figure 3 shows the numerical results for the emergence or disappearance of a giant connected component for these processes.

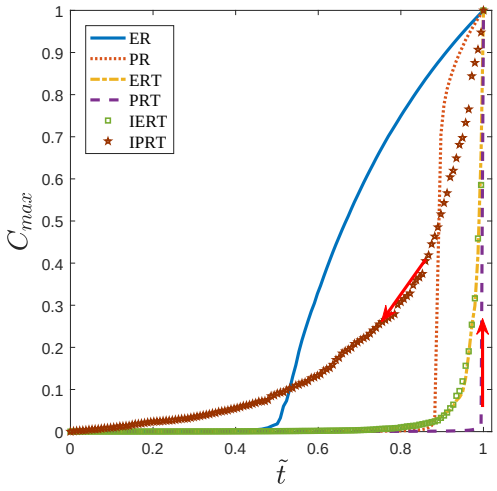


Fig. 3. The sizes of the largest connected components for various processes, as simulated on a tree for $N = 40\,000$. The data are smoothed by using over 100 runs. A forward line and a backward line indicated by two arrows form a loop.

For ERT and IERT processes, the largest component undergoes a phase transition when $t_1 N$ edges have been added or $(1 - t_2)N$ edges have been deleted. t_1 and t_2 are almost equal, *i.e.*, $t_1 \approx 0.9166$ and $t_2 \approx 0.9171$. For PRT process, the largest component does also a phase transition when $t_1 N$ edges have been added or $(1-t_2)N$ edges have been deleted. However, t_1 and t_2 have different values, *i.e.*, $t_1 \approx 0.9925$ and $t_2 \approx 0.8898$. All these values are obtained via extensive numerical simulations, typically up to $N = 70\,000$, and the values are done over many runs.

3. Discussion

In order to judge the characteristic of the phase transitions in ERT and IERT processes, the numbers of the connected components with certain size at the critical point are calculated. The distributions of connected components with respect to the size at the critical points for various processes are shown in Fig. 4.

For the sake of reasonably dealing with some outlier data, the different neighboring cluster sizes were combined in one bin according to the method in Ref. [2]. Concretely, the k^{th} bin contains all clusters with 2^{k-1} to $2^k - 1$ nodes, $k = 1, 2, \ldots, 9$. These results are plotted at the geometric mean of the two sizes $2^{k-1}$and $2^k - 1$, *i.e.*, at $\sqrt{2^{k-1}(2^k - 1)}$. The data in last several bins have been truncated. This avoids the certain systematic errors for possible outlier over there just because of the sparse distribution of cluster sizes or the lack of cluster at the interval of 2^{k-1} to $2^k - 1$ (see the rightest of the plot in the inset of Fig. 4 (b)). When the different neighboring cluster sizes were combined in one bin, it will lead to the inconsistency of the data.

It is obvious that the nice straight lines exhibit in these log–log plots except in Fig. 4 (c). The slopes of the straight lines in these plots give the exponent of the power law, *i.e.* $-\tau$, here τ is the Fisher exponent [2]. Let n_s denote the number of the clusters with size s, Eq. (3) is valid for larger s at the critical point t_c

$$n_s(t_c) \propto s^{-\tau} . \tag{3}$$

Considering that the power law for the cluster sizes distribution is a typical characteristic for continuous transitions [19, 26, 27], the phase transitions are continuous for ERT evolution processes and IERT fragmentation process. In both processes, there are linear number of N of edges existed in the graph near the critical point, thus, only when edges with a linear number of N of edges have been added to the network, the size of the largest cluster go from sublinear in N to linear in N. Therefore, the transitions are continuous.

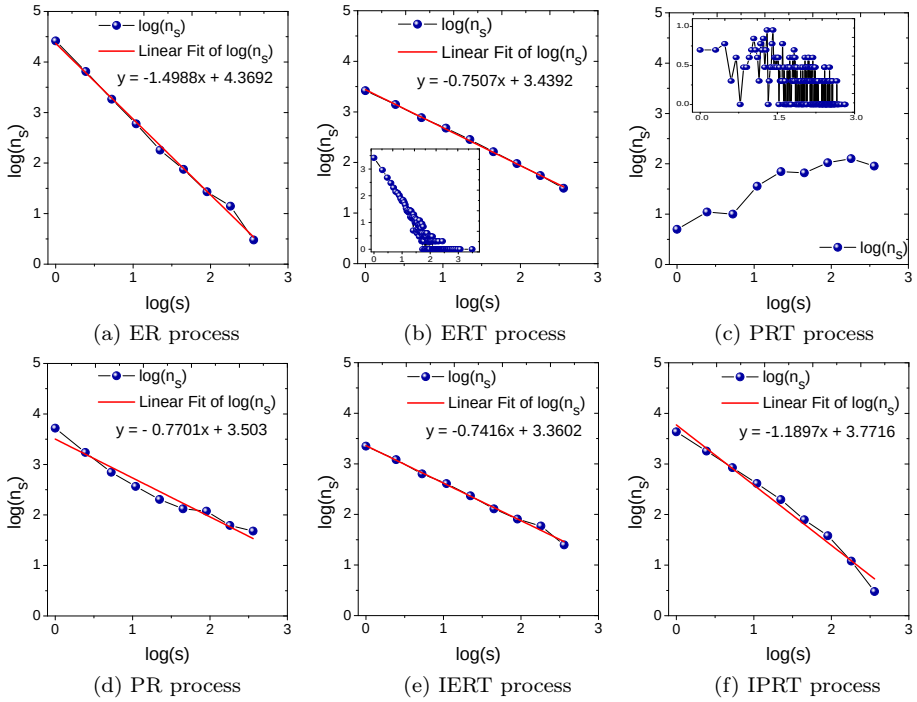


Fig. 4. (Color online) The distributions of the clusters with respect to the size at the critical points. The straight/red lines are the best fits. Figures 4 (a) and 4 (d) were chosen from the ER and PR evolution processes. There is not any limitation on the target tree's structure. Figures 4 (b) and 4 (e) were chosen from an ERT process and its inverse process, and Figs. 4 (c) and 4 (f) were done from a PRT process and its inverse process, respectively. The logarithm in these plots is the base 10 logarithm. The log–log plots in inset of Fig. 4 (b) and Fig. 4 (c) have no combination of neighboring clusters and no cut-off too. Inset of Fig. 4 (b) shows a power law behavior for the distribution of cluster sizes. Inset of Fig. 4 (c) shows that the data has almost uniform distribution if compared with the scope of the cluster sizes. If the different neighboring cluster sizes are combined in one bin, the height of the combined bin is similar to the area of a rectangle with the width of $2^k - 1$–2^{k-1} and the constant height. When the bin is plotted at the geometric mean of the two sizes 2^{k-1} to $2^k - 1$, the curve approximates to be a straight line with positive slope in theoretically.

A power law of cluster sizes distribution is also observed very near the critical point for IPRT process (see Fig. 4 (f)). It implies that the percolation transition is continuous for IPRT process. In contrast, n_s has almost uniform distribution if compared with the scope of the cluster sizes for PRT process (see the inset of Fig. 4 (c)). For a uniform distribution, *i.e.*, $n_s = b, s =$

$1, 2, \cdots, C_{\max}$ (b is a constant), one will have $\sum_{s=1}^{C_{\max}} bs = N$. This means that the clusters number $\sum_{s=1}^{C_{\max}} n_s$ can be approximated by $\sqrt{2bN}$, which is a number sublinear in N. This indicates that the addition of a small fraction of edges leads to the merger of most of such small clusters. That is to say, only a number sublinear in N of edges are needed to add to the graph, can the largest connected component grow from miniscale to macroscale. It is a discontinuous transition process [19]. That is to say, the evolution process for PRT undergoes a discontinuous phase transition but the fragmentation process does a continuous one. Furthermore, we check carefully the cluster sizes distribution on various system sizes of PRT process and find that n_s always observes the similar distribution.

It is notable that the distribution in PR process is power law (see, Fig. 4 (d)), and percolation transition for PR process on random tree model is continuous, which is the same as on regular lattice and scale free network [28]. However, the percolation transition for a PRT process (specific tree model) is discontinuous. It is obvious that the curves for the sizes of the largest connected components in PRT and IPRT processes form a loop. This loop, however, is quite different with the typical hysteresis loop in the traditional discontinuous phase transition. Here, one process in the loop is responding to a discontinuous transition but another is not.

4. Conclusion

In this paper, we have studied numerically in detail the evolution process and fragmentation process of a specific tree, including two kinds of candidate edges strategies. The explicit way for the determination of the critical point of the phase transition has been given by the treatment of the number of the connected components with different sizes. The feature of phase transition has been clarified by calculating the distribution of the clusters with respect to the size at the critical point. The results show that for PRT process, there existed two types of phase transitions, one is discontinuous in the evolution process, but the other is continuous in the fragmentation process.

REFERENCES

[1] S.R. Broadbent, J.M. Hammersley, *Math. Proc. Cambridge* **53**, 629 (1957).

[2] D. Stauffer, A. Aharony, *Introduction to Percolation Theory*, Taylor and Francis, 2003.

[3] H.A. Abdusalam, *Int. J. Theor. Phys.* **37**, 1587 (1998).

[4] H.A. Abdusalam, *Chaos, Solitons Fract.* **11**, 1407 (2000).

[5] C. Marzouk, *Stoch. Proc. Appl.* **126**, 265 (2016).

[6] G.H. Pollack, *Cells, Gels and the Engines of Life*, Ebner and Sons, 2001.
[7] E. Ahmed, H.N. Agiza, *Int. J. Theor. Phys.* **35**, 2483 (1996).
[8] R. Albert, H. Jeong, A.-L. Barabási, *Nature* **406**, 378 (2000).
[9] L. Wang *et al.*, *Solid State Commun.* **152**, 1641 (2012).
[10] B.J. Simons *et al.*, *Chem. Phys. Lett.* **605**, 121 (2014).
[11] Z. Burda, J.D. Correia, A. Krzywicki, *Phys. Rev. E* **64**, 046118 (2001).
[12] R. Bundschuh, R. Bruinsma, *Phys. Rev. Lett.* **100**, 148101 (2008).
[13] R. Kürsten, *Phys. Rev. E* **93**, 032111 (2016).
[14] D. Khoshnevisan, *Probab. Theory Rel. Fields* **140**, 169 (2008).
[15] J. Bertoin, *Electron. J. Probab.* **19**, 15 (2014).
[16] E. Arias-Castro, *Stat. Probab. Lett.* **81**, 302 (2011).
[17] T. Nogawa, T. Hasegawa, *J. Phys. A* **42**, 145001 (2009).
[18] D. Achlioptas, R.M. D'Souza, J. Spencer, *Science* **323**, 1453 (2009).
[19] F. Radicchi, S. Fortunato, *Phys. Rev. E* **81**, 036110 (2010).
[20] N. Bastas, K. Kosmidis, P. Argyrakis, *Phys. Rev. E* **84**, 066112 (2011).
[21] N. Bastas, P. Giazitzidis, M. Maragakis, K. Kosmidis, *Physica A* **407**, 54 (2014).
[22] P.L. Krapivsky, S.N. Majumdar, *Phys. Rev. Lett.* **85**, 5492 (2000).
[23] Z. Kalay, E. Ben-Naim, *J. Phys. A* **48**, 045001 (2015).
[24] S.H. Yook, H. Jeong, A.-L. Barabási, Y. Tu, *Phys. Rev. Lett.* **86**, 5835 (2001).
[25] A. Barrat, M. Barthélemy, A. Vespignani, *Phys. Rev. E* **70**, 066149 (2004).
[26] N.A.M. Araújo, H.J. Herrmann, *Phys. Rev. Lett.* **105**, 035701 (2010).
[27] P. Grassberger *et al.*, *Phys. Rev. Lett.* **106**, 225701 (2011).
[28] O. Riordan, L. Warnke, *Science* **333**, 322 (2011).