Hybrid Percolation Transition in Cluster Merging Processes:
Continuously Varying Exponents

Y.S. Cho,1 J.S. Lee,2 H.J. Herrmann,3 and B. Kahng1

1CCSS, CTP and Department of Physics and Astronomy, Seoul National University, Seoul 08826, Korea
2School of Physics, Korea Institute for Advanced Study, Seoul 02455, Korea
3Computational Physics for Engineering Materials,
Institute for Building Materials, ETH Zürich, 8093 Zürich, Switzerland

Consider growing a network, in which every new connection is made between two disconnected
nodes. At least one node is chosen randomly from a subset consisting of g fraction of the entire
population in the smallest clusters. Here we show that this simple strategy for improving connection
exhibits a phase transition barely studied before, namely a hybrid percolation transition exhibiting
the properties of both first-order and second-order phase transitions. The cluster size distribution of
finite clusters at a transition point exhibits power-law behavior with a continuously varying exponent
\( \tau \) in the range \( 2 < \tau (g) \leq 2.5 \). This pattern reveals a necessary condition for a hybrid transition in
cluster aggregation processes, which is comparable to the power-law behavior of the avalanche size
distribution arising in models with link-deleting processes in interdependent networks.

PACS numbers: 64.60.De,64.60.ah,89.75.Da

Transport or communication systems grow by adding new connections. Often certain constraints are imposed
by society and if these constraint involve global knowledge about connectivity, the transition to a percolating
system can become first order, as happens for instance when suppressing the spanning cluster, when imposing a
cluster size \([1]\) or when favoring the most disconnected sites \([2]\). Typically this effect is accompanied by the loss
of critical scaling making these abrupt transitions less predictable and thus more dangerous. We will show here,
that for a specific case, namely a variant of the model introduced in Ref. \([3]\), critical fluctuations and power-law
distributions can prevail and for the first time identify a hybrid transition in explosive percolation.

Hybrid phase transitions have been observed recently in many complex network systems \([4, 5]\); in these
transitions, the order parameter \(m(t)\) exhibits behaviors of both first-order and second-order transitions simultane-
ously as

\[
m(t) = \begin{cases}
0 & \text{for } t < t_c, \\
n_0 + r(t - t_c)^{\beta} & \text{for } t \geq t_c,
\end{cases}
\]

where \(n_0\) and \(r\) are constants and \(\beta\) is the critical exponent of the order parameter, and \(t\) is a control parameter.
Examples of such behavior include \(k\)-core percolation \([6, 7]\), the cascading failure model on interdependent
complex networks \([8, 9]\), and the Kuramoto synchronization model with a correlation between the natural fre-
cuencies and degrees of each node on complex networks \([10, 11]\), etc. For the models in \([8, 9]\), a critical behavior
appears as nodes or links are deleted from a percolating cluster above the percolation threshold until reaching
a transition point \(t_c\). As \(t\) is decreased infinitesimally further as \(1/N\) beyond \(t_c\) in finite systems, the order
parameter decreases suddenly to zero and a first-order phase transition occurs. Thus, a hybrid phase transition
occurs at \(t = t_c\) in the thermodynamic limit.

Next we recall discontinuous percolation transitions occurring in generalized contagion models \([12, 14]\). Re-
cent studies \([13]\) of a generalized epidemic model \([13]\) revealed that the discontinuous percolation transition turns
out to be a hybrid percolation transition (HPT) represented by \([1]\). For this case, a HPT is induced by cluster
merging processes. However, the critical behavior arising in a HPT in a cluster merging process has not been yet
studied at all, even though one may guess that its nature can differ from that of the HPT in link-deleting pro-
cesses \([6, 9]\). This Letter aims to understand the nature of an HPT in a cluster merging process and identify the
similarities and differences in the phase transition compared with those of HPTs in link-deleting processes. Our
study is based on a simple stochastic model introduced later, from which we could obtain analytic solutions for
diverse properties of the critical behavior.

The model we study is defined as follows: We start with a system consisting of \(N\) isolated nodes. At each
time step, one node is selected uniformly at random from the entire system and the other node is selected from a
restricted set consisting of approximately \(g\) fraction of the entire nodes, as defined below. The two nodes are
connected by an edge unless they are already connected. The time step \(t\) is defined as the number of edges added to
the system per node, which is a control parameter. The restricted set is defined as follows: First, we rank clusters
by ascending order of size at each time step. When more than one cluster of the same size exists, they are sorted
randomly. At this stage, a restricted set of clusters \(R(t)\) is defined as the subset consisting of a certain number of
the smallest clusters (say \(k\) clusters), denoted as \(R(t) = \{c_1, c_2, \cdots, c_k\}\). Further, \(k\) is determined as the value
satisfying the inequalities, \(N_{k-1}(t) < \lfloor gN \rfloor \leq N_k(t)\) for a given model parameter \(g\in (0, 1]\). \(N_k(t) \equiv \sum_{\ell=1}^{k} s_\ell(t),\)
The merged cluster is still in \( R \) in the system exceeds \( (1 - s) \) referred to as \( S \) with ranking \( S \) for \( s > s^* \) or equal to \( s \). This behavior is similar to that obtained in the colation threshold of the ordinary ER model as shown in Fig. 2. 

Fig. 2. (Color online) Plot of \( m(t) \) vs \( t \) for different model parameter values \( g = 0.2, 0.4, 0.5, 0.6, 0.8 \) and 1.0 from right to left. Inset: Plot of the susceptibility \( \chi(t) \) vs \( t \) for different system sizes \( N/10^4 = 8, 64, 512 \) and 4096 for a fixed \( g = 0.5 \).

We examine the cluster size distribution \( n_s(t) \) for several time steps as shown in Fig. 3 where \( n_s(t) \) denotes the number of clusters of size \( s \) divided by \( N \). i) When \( t \ll t_c \), \( n_s(t) \) decays exponentially with respect to \( s \). ii) When \( t = t_c \), \( n_s(t) \) exhibits power-law decay in the small-cluster region but contains a bump in the large-cluster region. iii) When \( t > t_c \), at which the order parameter becomes \( m_0 \), \( n_s(t) \) exhibits power-law decay with the exponent \( \tau \) for finite clusters, at an \( s \)-distance from a giant cluster positioned separately at \( s = m_0 N \). Note that \( \tau \) depends on the model parameter \( g \). iv) When \( t > t_c \), the size distribution of finite clusters exhibits crossover behavior: it undergoes power-law decay for \( s < s^* \), but exponential decay for \( s > s^* \). Thus, \( n_s(t) \sim s^{-\tau} e^{-s/s^*} \), where \( s^* \sim (t - t_c)^{-1/\sigma} \) according to the conventional scaling theory.

We consider the evolution of the cluster size at an arbitrary time step \( t \) using the Smoluchowski equations for
Note that the number of clusters of size $S_R$ can be greater than one. In this case, some of them are randomly chosen to belong to $R(t)$ and the others belong to $R^{(c)}$ to satisfy the inequalities $N_{k-1}(t) < [gN] \leq N_k(t)$ presented previously. Thus, we need to consider the case $s = S_R$ separately. In the above equations, we used the approximation that the total number of nodes in the set $R$ is $N_R \approx gN$ for $t < t_g$, where $t_g$ is the time step at which the size of the largest cluster in the system exceeds $(1 - g)N$ for the first time. Thus, $g$ is taken as unity for $t < t_g$. Further, $t_g$ is located between $t_1$ and $t_c$. The derivations of each term for each case of $s$ are explained in the SI. Moreover, performing direct numerical integration of the Smoluchowski equations, we successfully reproduce the behaviors of $m(t)$ in Fig. 2 and $n_s$ in Fig. 3, which are shown in the SI.

$$\frac{dn_s(t)}{dt} = \sum_{i,j=1}^{\infty} \frac{in_jn_j}{g} \delta_{i+j,s} - (1 + \frac{1}{g})sn_s$$

for $s < S_R$, (2)

$$\frac{dn_s(t)}{dt} = \sum_{i,j=1}^{\infty} \frac{in_jn_j}{g} \delta_{i+j,s} - sn_s - \left(1 - \sum_{k=1}^{S_R-1} \frac{in_k}{g}\right)$$

for $s = S_R$, (3)

$$\frac{dn_s(t)}{dt} = \sum_{j=1}^{\infty} \delta_{i+j,s}jn_j (1 - \sum_{i=1}^{S_R-1} \frac{in_i}{g}) - sn_s$$

for $s > S_R$. (4)

We focus on the rate equation for $n_1(t)$ as follows: At early initial time steps, monomers are abundant and can be in both $R$ and $R^{(c)}$. However, as time passes, their number decreases and at a certain time step denoted as $t_1$, all the monomers are in the set $R$. The rate equation for monomers is written separately as $\frac{dn_1}{dt} = -n_1 - 1$ for $t < t_1$ using Eq. (3), and $\frac{dn_1}{dt} = - (1 + \frac{1}{g})n_1(t)$ for $t > t_1$ using Eq. (2). Then

$$n_1(t) = \begin{cases} \frac{2e^{-t} - 1}{g(\frac{1}{g}+1)e^{-(1+\frac{1}{g})t}} & \text{if } t < t_1, \\ \frac{2e^{-t} - 1}{g(\frac{1}{g}+1)e^{-(1+\frac{1}{g})t}} & \text{if } t > t_1, \end{cases}$$

(5)

and $t_1$ is obtained as $t_1 = \ln(\frac{2}{g+1})$.

Next, we use the fact that $n_s(t_c)$ follows a power law, i.e., $n_s(t_c) = a_0 s^{-\tau}$ for $1 \leq s \leq S_0$, where $S_0$ is the size of the largest cluster among the finite clusters and $a_0$ can be determined in terms of $g$ and $t_c$ from Eq. (5). We can also use the relation $\sum_{s=1}^{S_0} n_s(t_c) = 1 - t_c$, because up to $t_c$, links are likely to be added between clusters, which is checked numerically in the SI. We also confirm that $d\sum_{s=1}^{\infty} n_s(t)/dt = -1$ from the Smoluchowski equations (2). Moreover, the relation $\sum_{s=1}^{S_0} n_s(t_c) \approx \sum_{s=1}^{\infty} n_s(t_c) = 1 - m_0$ can be used in the limit $N \to \infty$. Taking account of those facts, we obtain the self-consistency equation for the exponent $\tau$ as

$$\zeta(\tau) = 1 - m_0 \left(1 - \ln \left[ \frac{g\zeta(\tau-1) + 1}{g\zeta(\tau-1) + 1 - m_0} \right] \right)$$

(6)

To obtain $\tau$, one needs $m_0$ values for given $g$ values. We use the numerically estimated values of $m_0$ in Table I. The obtained values denoted as $\tau$ range between $2 < \tau(g) \leq 2.5$, as listed in Table I. We note that as $g$ decreases, $\tau$ becomes smaller, and the jump in the order parameter becomes larger. This result implies that the growth of large clusters is more strongly suppressed at smaller $g$.

To obtain the exponent $\sigma$ analytically, we take into account two facts: First, the dynamics in regime iv) is of the ER type: Because $m_0 > 1 - g$, the restricted set covers the entire system. Next we take $t_c$ as an ad hoc time origin, and the cluster size distribution at $t_c$ is used as an initial condition of the ER dynamics. Under this transformation, the solution for the evolution of the ER

![Graphs](image-url)
model remains the same [20]. Then using the solution for the cluster size distribution of the ER model [21], we obtain that \( s_n(\hat{t}) = s^{1−\gamma}(f(s^{1/\sigma})) \), where \( f(x) \) is a scaling function and \( \hat{t} = t − t_c \) for \( t > t_c \). Using the property that \( f(x) \) is an analytic function [22], we obtain \( \sigma = 1 \) independent of \( g \). The detailed derivation is presented in the SI.

The order parameter \( m(t) \) increases continuously with time when \( t > t_c \). To study the criticality for \( t > t_c \), we again take the transition point \( t_c \) as an ad hoc time origin. Next, we use the formalism for the ordinary ER model with an arbitrary initial cluster size distribution as presented in [20,22]. For instance, the order parameter can be obtained via the self-consistency equation, \( m(t) = H(2tm(\hat{t})) \), where \( H(\mu) \equiv 1 − \sum s_n(0)e^{-\mu s} \) [20,21]. Using \( n_i(\hat{t} = 0) = a_0s^{−\gamma} \) with \( a_0 = (1 − m_0)/(\zeta(\zeta − 1)) \), we obtain that \( m(\hat{t} = t − t_c) = m_0 + \beta(t − t_c)^{−\gamma} \), where \( \beta = \frac{\Gamma(2−\gamma)}{(\zeta−1)}(1 − m_0)(2m_0)^{−\gamma} \). This gives \( \beta = \tau − 2 \). The detailed derivation of the exponent \( \beta \) is presented in the SI. Note that we used \( m(\hat{t}) = m_0 + \hat{t}^{\beta/\gamma} \) instead of \( m(\hat{t}) \sim \hat{t}^{\beta/\gamma} \), relevant to the case \( m_0 = 0 \). We have obtained \( \beta = (\tau − 2)/(3 − \tau) \). For the ER model, \( \tau_{ER} = 5/2 \) and \( \beta_{ER} = 1 \). Accordingly, we say that the discontinuity of the order parameter \( m_0 \neq 0 \) at \( t_c \) changes the criticality for \( t > t_c \).

The susceptibility \( \chi^+ = \sum_{s=1} s^2n_s(\hat{t}) \) for \( t > t_c \), where the prime indicates the exclusion of an infinite-size cluster, can be obtained using \( \chi^+ = H′(2tm(\hat{t}))/[1 − 2H′(2tm(\hat{t}))] \) with \( \hat{t} = t − t_c \), where the prime in \( H(\mu) \) indicates the derivative with respect to its argument \( \mu \). When \( 2 < \tau < 3 \), \( H′(\mu) \) has a non-integer singularity of \( \mu^{−3} \). Plugging \( \mu = 2tm_0 \) into the formula for \( \chi^+(\hat{t}) \), we obtain that \( \chi^+(\hat{t}) \sim (t − t_c)^{−\gamma} \) with \( \gamma = 3 − \tau \). Note that the denominator is finite at \( \hat{t} = 0 \).

Obtaining the analytical result of the susceptibility \( \chi^−(\hat{t}) \equiv \sum_{s=1} s^2n_s(t) \) for \( t < t_c \) is intriguing. Using the Smoluchowski equation for the r-ER model, we obtain the following relation \( \partial\chi^−/\partial t = 2\chi^−(\chi_R)(t) \), where \( \chi_R \equiv (1/g)\sum_{s=1}^s s^2n_s(\hat{t}) \) and we used the approximation \( \sum_{s=1}^s s^2n_s \approx g \) which is valid near \( t_c \). If we assume \( \chi^−(\hat{t}) \) diverges algebraically as \( t \to t_c \), i.e., \( \chi^− \sim (t_c − t)^{−\gamma} \), which is confirmed numerically, then we could obtain \( \chi_R = (\gamma/2)(t_c − t)^{−1} \). However, the exponent \( \gamma \) was not analytically determined. In fact, \( \chi_R \) is another form of the susceptibility defined via the correlation function [23] as \( \chi_R \equiv (1/N)\sum_{i,j=1}^N C(i,j) \), where \( C(i,j) \) is the probability that two nodes \( i \) and \( j \) belong to the same finite cluster. Thus, \( C(i,j) \) can be obtained as the probability that two nodes \( i \) and \( j \) are selected from the same cluster. Then, \( \chi_c = (N\sum_{s=1}^{S_R}(s/n)(s/gN)) \), where \( \alpha = 1 \) is the index of cluster. This formula reduces to \( \chi_R \) using \( N_R/N \approx \sum_{s=1}^{S_R} s_n \approx g \) near \( t_c \).

A critical behavior at \( t_c \) appears in the form of the power-law-type size distribution of finite clusters, which is necessary to generate an HPT, with exponent \( \tau \) ranging between \( 2 < \tau < 2.5 \). This pattern is analogous to the power-law behavior of the avalanche size distribution in the cascading failure model [8].

The r-ER model is a simple model exhibiting an HPT in a cluster merging process. As the classical ER model has served as a basic model for understanding the evolution of social networks, the r-ER model should be similarly useful but under a certain constraint that the least connected members are preferred to connect to others, for instance, as in the merging of the lowest financial companies under government control in financial crisis. Moreover, the theoretical framework obtained from the r-ER model can be used for understanding other HPTs in cluster merging processes, for instance, in generalized epidemic models [13,15] and synchronization models [10,11]. Our preliminary results suggest that indeed HPTs in cluster merging processes could be found from those models.

We have shown here, that in a model, in which the most disconnected agents are preferentially connected, spanning occurs abruptly, while astonishingly critical fluctuations and power-law distributions prevail after the transition within the critical region. These post-transition critical mergers of rather big clusters into a rather meager spanning one represent a rather transparent geometrical interpretation of the recently discovered hybrid phase transitions.

This work was supported by NRF grants (No. 2010-0015066 and No. 2014R1A3A2069005), the exchange program between the Korean NRF and ERC, ERC Advanced Grant No. FP7-319968-FlowCCS and the Global Frontier Program (YSC).
Several different system sizes up to $N = 4 \times 10^7$. The details of how to measure them are presented in the SI.

\[\text{Reference} \] R.M. D’Souza and J. Nagler, Nat. Phys. \textbf{11}, 531 (2015).

| $g$    | $t_c$  | $m_0$  | $\tau^*$ | $\tau$  | $1/\sigma$ | $\beta$  | $\gamma'$ | $\gamma$ |
|--------|--------|--------|-----------|---------|------------|----------|-----------|----------|
| 0.1    | 0.9984±0.0002 | 0.98±0.05 | 2.06      | 2.03±0.04 | 1.00±0.05 | 0.04±0.05 | 1.00±0.02 | 0.97±0.05 |
| 0.2    | 0.9836±0.0002 | 0.90±0.05 | 2.13      | 2.08±0.04 | 1.01±0.05 | 0.09±0.05 | 1.03±0.02 | 0.91±0.05 |
| 0.3    | 0.9570±0.0002 | 0.81±0.05 | 2.18      | 2.12±0.04 | 1.01±0.05 | 0.13±0.05 | 1.06±0.03 | 0.88±0.05 |
| 0.4    | 0.9227±0.0002 | 0.72±0.05 | 2.23      | 2.16±0.04 | 1.02±0.05 | 0.17±0.04 | 1.08±0.03 | 0.85±0.05 |
| 0.5    | 0.8826±0.0002 | 0.62±0.05 | 2.26      | 2.18±0.04 | 1.04±0.05 | 0.21±0.05 | 1.10±0.03 | 0.83±0.05 |
| 0.6    | 0.8370±0.0002 | 0.52±0.06 | 2.29      | 2.20±0.04 | 1.05±0.05 | 0.23±0.04 | 1.11±0.04 | 0.82±0.05 |
| 0.7    | 0.7852±0.0003 | 0.41±0.05 | 2.32      | 2.22±0.04 | 1.08±0.05 | 0.28±0.05 | 1.13±0.05 | 0.81±0.05 |
| 0.8    | 0.7250±0.0003 | 0.29±0.05 | 2.35      | 2.25±0.04 | 1.12±0.05 | 0.32±0.05 | 1.15±0.05 | 0.81±0.05 |
| 0.9    | 0.650±0.002  | 0.15±0.05 | 2.38      | 2.28±0.04 | 1.20±0.10 | 0.39±0.08 | 1.15±0.05 | 0.80±0.05 |

Moreno, Phys. Rev. Lett. \textbf{106}, 128701 (2011).
[11] B.C. Coutinho, A.V. Goltsev, S.N. Dorogovtsev, and J.F.F. Mendes, Phys. Rev. E \textbf{87}, 032106 (2013).
[12] G. Bizhani, M. Paczuski and P. Grassberger, Phys. Rev. E \textbf{86}, 011128 (2012).
[13] H.-K. Janssen, M. Müller, and O. Stenull, Phys. Rev. E \textbf{70}, 026114 (2004).
[14] P. S. Dodds and D.J. Watts, Phys. Rev. Lett. \textbf{92}, 218701 (2004).
[15] K. Chung, Y. Baek, D. Kim, M. Ha and H. Jeong, Phys. Rev. E \textbf{89}, 052811 (2014).
[16] D. Achlioptas, R. M. D’Souza, and J. Spencer, Science \textbf{323}, 1453 (2009).
[17] R.M. D’Souza and J. Nagler, Nat. Phys. \textbf{11}, 531 (2015).
[18] T. Bohman, A. Frieze, and N.C. Wormald, Random Struct. Algorithms \textbf{25}, 432 (2004).
[19] $t_c$ can be taken as the intercept with the $t$-axis of the tangential line of $m(t)$ at the inflection point of $m(t)$. $t_c$ differs from $t_c$ in finite systems, but it reduces to $t_c$ in the thermodynamic limit.
[20] Y. S. Cho, B. Kahng, and D. Kim, Phys. Rev. E \textbf{81}, 030103(R) (2010).
[21] R.M. Ziff, E.M. Hendriks, and M. H. Ernst, J. Phys. A \textbf{16}, 2293-2320 (1983).
[22] R.A. da Costa, S.N. Dorogovtsev, A.V. Goltsev and J.F.F. Mendes, Phys. Rev. E \textbf{91}, 032140 (2015).
[23] R.A. da Costa, S.N. Dorogovtsev, A.V. Goltsev and J.F.F. Mendes, Phys. Rev. E \textbf{90}, 022145 (2014).