Enhancing robustness and immunization in geographical networks

Liang Huang,1,3,* Kongqing Yang,2,3 and Lei Yang1,3,4†

1Institute of Modern Physics, Chinese Academy of Science, Lanzhou 730000, China
2Center for Complex Systems, Jimei University, Xiamen 361021, China
3Department of Physics, Lanzhou University, Lanzhou 730000, China
4The Beijing-Hong Kong-Singapore Joint Centre for Nonlinear and Complex Systems (Hong Kong), Hong Kong Baptist University, Hong Kong, China

(Received 11 April 2005; revised manuscript received 6 September 2006; published 1 March 2007)

We find that different geographical structures of networks lead to varied percolation thresholds, although these networks may have similar abstract topological structures. Thus, strategies for enhancing robustness and immunization of a geographical network are proposed. Using the generating function formalism, we obtain an explicit form of the percolation threshold \( q_e \) for networks containing arbitrary order cycles. For three-cycles, the dependence of \( q_e \) on the clustering coefficients is ascertained. The analysis substantiates the validity of the strategies with analytical evidence.

DOI: 10.1103/PhysRevE.75.036101 PACS numbers: 89.75.Hc, 84.35.+i, 05.70.Fh, 87.23.Ge

Complex networks (see reviews [1]) provide powerful tools to investigate complex systems in nature and society. The properties of complex systems are affected by the geographical distribution of the components. For example, routers of the Internet [2] and transport networks [3] lay on the two-dimensional surface of the globe, worldwide airport networks are confined by the geography [4], and neuronal networks in brains [5] occupy three-dimensional space. Thus it is helpful to study geographical complex networks [6–11].

From the abstract geometrical point of view, an abstract set can describe a general system. When it is equipped with some geometrical structures, the set can further describe a specified system. So an abstract topological network is an abstract set that consists of nodes and links. A metric can be added to the abstract topological network and the metric can be arbitrary, not necessarily the Euclidean metric [6]. Embedding a particular abstract topological network into a suitable metric space provides a method to add a metric.

In this paper, the problem of percolation thresholds in geographical networks is studied. Three types of geographical networks are investigated: the normal model [7,8], the hollow model, and the concentrated model. By extensive numerical simulations, we found that the percolation threshold \( q_e \) (the point that a spanning connected cluster emerges) for these models satisfies \( q_e \) (concentrated) > \( q_e \) (normal) > \( q_e \) (hollow). Based on these results, we suggest a strategy (hollowing) for enhancing robustness and a strategy (concentrating) for the immunization of geographical networks. The different geographical networks have different distributions of cycles. The geographical dependence of the percolation threshold is investigated by the generating function process in abstract networks containing cycles.

Based on the lattice embedded model [7], the networks are generated as follows: each node in an \( L \times L \) lattice with periodic boundary conditions is assigned a degree quota \( k \), drawn from the prescribed degree distribution. Here, scale-free \( P(k) \sim k^{-\lambda} \), \( m \leq k \) and exponential \( P(k) \sim e^{-k/k_0}, \ m \leq k \) degree distributions are considered, where \( m \) is the minimum degree a node can have. (a) The normal case [lattice embedded scale-free (LESF) or lattice embedded exponential (LEE)]: A node \( i \) connects to its closest neighbors until its degree \( k_i \) is realized, or up to a cutoff distance \( A \sqrt{k_i} \), where \( A \) is large enough to ensure that almost all the degree quotas can be fulfilled. (b) The hollow case [hollow LESF (HLESF) or hollow LEE (HLEE)]: similar to the normal case, except that a node \( i \) has probability \( p \) to be forbidden to connect its first \( n \) nearest neighbors. (c) The concentrated case: \( A \) is set smaller than that in the normal case. The process is repeated throughout all the nodes in the lattice. For the hollow case, when the network degenerates into a lattice and \( n \) is small, it becomes similar to the tunneling effect in Euclidean lattices [12]. In the following simulations, we choose \( A=7, n=8, \) and \( p=1 \); network size \( N=10^6 \), minimum degree \( m \) is 4 for scale-free networks and 1 for exponential networks, and all the data are averaged over 1000 ensembles, unless otherwise specified.

The algorithms of Newman and Ziff [13] are performed to calculate the threshold \( q_e \), which is defined as the point where the differential of the size of the largest cluster as a function of occupying probability \( q \) maximizes. The definition is equivalent to the usual definition by the emergence of a spanning cluster for the large-network-size limit [14]. Figure 1 shows a clear drop of the percolation threshold \( q_e \) in the hollow networks than in the normal networks. So the robustness can be enhanced by the model. The size effect (scale-free degree distribution) is demonstrated in Fig. 2. The LESF and even the HLESF networks also have nonzero percolation thresholds for \( \lambda \in (2,3) \); the results are consistent with Ref. [14]. Again, the drop in \( q_e \) for hollow networks is apparent.

The amendment in hollow networks is small in the physical space. As Fig. 3 shows, the average spatial length \( \langle l \rangle \) of the edges for the hollow networks dose not increase much compared with the normal networks, for small \( \lambda \) or large \( k_0 \). As the average degree \( \langle k \rangle \) becomes larger; the difference

*Present address: Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287, USA.
†Corresponding author.
between \( \langle l \rangle \) decreases. In the limit case, the difference goes to 0. While the cost of constructing hollow networks still remains low, they are more robust than normal models. Under the same conditions, hollow networks have significant lower percolation thresholds.

Based on the above observations, we propose a hollowing strategy to enhance the robustness of geographical networks. For each node in a geographical networks, we introduce a probability \( p \) to cut down the edges that linked to its first \( n \) nearest neighbors, then to reconnect further nodes in the geographical distance. In this case, the degree distribution \( P(k) \) deviates a little for small \( k \)—i.e., around about 10. This only causes a variation in percolation thresholds of a much smaller magnitude. Figure 4 demonstrates the efficiency of the hollowing strategy for LESF networks. The percolation threshold drops about 0.1 with \( n=8 \); namely, it needs 10% less nodes of the network to maintain a spanning cluster. This can have significant effect to prevent the network from breakdown when the network undergoes a serious crisis and is losing its global function.

For concentrated cases, we study the concentrated scale-free networks since the problem is mainly concerned with immunization strategies [15]. The long-range links (for large \( A \)) in the LESF model are different from those in the small-world model [16]. In the LESF model, a node \( i \) first tries to connect to its nearest neighbors; if they have already fulfilled their degree quotas, this node has to try further nodes until up to the cutoff distance \( A \sqrt{k_i} \). The long-range links in the LESF model are still somewhat local. As far as \( A \) is large enough to ensure that almost all the nodes' degree quotas are fulfilled, the network properties depend little on \( A \), as shown in Fig. 5(a). However, as \( A \) decreases, more and more long-range links are prohibited [7]. The network becomes more localized; thus, the clustering coefficient increases while the average degree decreases a little. As \( A \) is approximately larger than 0.5, most of the nodes’ degree quotas are still fulfilled except the nodes with large degree quotas. Thus the network structure remains mainly unchanged, and the percolation threshold remains almost the same. As \( A \) decreases further (\( A<0.5 \)), a large portion of the degree quotas of nodes cannot be fulfilled. Figure 5(b) shows a serious degree cutoff (\( A=0.3 \)), where the steps in the degree distribution come from the symmetry of the two-dimensional (2D) lattice. Thus the network structure becomes seriously deteriorated: the average degree decreases sharply, and the clustering coefficient drops; the percolation threshold increases rapidly. These effects indicate that the concentrated networks are much easier to be immunized. It is indeed the case that,
However, if an outgoing edge is not independent—i.e., it is percolation problem. A real network is usually clustered and contains a cluster size distribution that diverges or, equivalently, the average cluster size diverges during an epidemic, most people stay home to be secluded from infection.

To better understand the numerical results, we employ the generating function method to determine the percolation threshold for networks with different clustering properties. Here the clustering properties can be simply depicted by the clustering coefficient, which counts for the triangles (three-cycles) in the network, and it can also be represented by the number of rectangles (four-cycles), and generally by the number of L-cycles. In the following, a general relation of the percolation threshold \( q_c \) and the number of L-cycles for a random network with arbitrary degree distribution is determined. As an example, the dependence of \( q_c \) on the clustering coefficient is obtained.

For uniform occupations (or random failures), the percolation threshold of random tree-like networks is \( q_c = \langle k \rangle / \langle k(k-1) \rangle \) \([17,18]\). It could be obtained by the condition that the average cluster size diverges or, equivalently, the average size of clusters that reached by following an edge diverges. A real network is usually clustered and contains a certain amount of cycles. If the number of cycles is small (e.g., each node belongs to at most one cycle), the generating function process can be extended to cope with the random percolation problem.

For a uniform occupation probability \( q \), the generating function for the probability of the number of outgoing edges of a target node reached by following a randomly chosen edge on a clustered network remains the same as that of the random tree-like networks \([18,19]\): \( F_1(x) = \frac{q}{\langle k \rangle} \sum k P(k) x^{k-1} \). However, if an outgoing edge is not independent—i.e., it is terminated to a node that has already been visited (having no contribution for reaching new nodes)—the node with degree \( k \) will only have \( k-2 \) independent outgoing edges. Thus it is convenient to define

\[
F^{(1)}_1(x) = \frac{q}{\langle k \rangle} \sum k P(k) x^{k-2} = x^{-1} F_1(x)
\]

as the generating function of the number of outgoing edges reaching new nodes for such a target node.

Let \( H_i(x) \) be the generating function of the size distribution of the cluster that is reached by following an edge; by following two independent edges, the cluster size distribution is generated by \( H_i(x)^2 \). But if the two edges originate from a common node and belong to an L-cycle, the generating function should be \( H_i^{(1)}(x) H_i^{(l-1)}(x) \), where \( l_i = \lfloor (L-1)/2 \rfloor \) and \( l = (L-1)/2 \). \([q]\) is Gauss’ function which returns the integer part of \( g \). \( H_i^{(0)}(x) \) is the generating function for the size distribution of the clusters that is reached by an edge and has one edge terminated after \( l \) steps, and satisfies an iterative relation

\[
H_i^{(0)}(x) = 1 - F_1(1) + x F^{(1)}_1(H_i(1)) H_i^{(l-1)}(x),
\]

where the terminal condition \( H_i^{(1)}(x) \) is

\[
H_i^{(1)}(x) = 1 - F_1(1) + x F^{(1)}_1(H_i(1)).
\]

Thus

\[
H_i^{(0)}(x) = [1 - F_1(1)] \left[ 1 - \left[ x F^{(1)}_1(H_i(1)) \right]^2 / \left[ 1 - x F^{(1)}_1(H_i(1)) \right] + [x F^{(1)}_1(H_i(1))]^2 \right].
\]

In general, we may assume that a node \( i \) with degree \( k_i \) on average belongs to \( n_i(k) \) L-cycles, where \( n_i(k) < 1 \) or \( \sim 1 \). So \( H_i(x) \) satisfies the self-consistent equation

\[
H_i(x) = 1 - F_1(1) + \frac{q x}{\langle k \rangle} \sum k P(k) \times [H_i^{(1)}(x) H_i^{(l-1)}(x)]^{q[k-2]} H_i^{(l-1)}(x)^{k-1-2n_i}.
\]

The average cluster size reached by an edge is

\[
\langle x \rangle = H_i^{(1)}(1) = q + \frac{q}{\langle k \rangle} \sum k P(k) n_i(k) [H_i^{(1)}(1) + H_i^{(l-1)}(1)] + [k - 2 n_i(k)] H_i^{(1)}(1),
\]

where

\[
H_i^{(m)}(1) = 1 - q^m / 1 - q [x F^{(1)}_1(H_i)](1) = 1 - q^m / 1 - q H_i^{(1)}(1)
\]

and

\[
H_i^{(1)}(1) = q + \frac{q}{\langle k \rangle} (k-2) H_i^{(1)}(1).
\]
A simple substitution yields

\[
\langle \bar{s} \rangle = \frac{q + \frac{2 - q^4 - q^5 q^2}{1 - q} \langle k \rangle (kn_l)}{1 - \frac{q}{\langle k \rangle} \left( \frac{k(k-1)}{2} - 2\langle kn_l \rangle \left( 1 - \frac{2 - q^4 - q^5 q^2 q(k(k-2))}{2(1-q) \langle k \rangle} \right) \right)}.
\]

Thus the percolation threshold \( q_c \), given by the divergence of \( \langle \bar{s} \rangle \), is

\[
q_c = \frac{\langle k \rangle}{\langle k(k-1) - 2\langle kn_l \rangle (1 - \frac{2 - q^4 - q^5 q^2 q(k(k-2))}{2(1-q) \langle k \rangle})}. \tag{1}
\]

When \( n_L=0 \), this result degenerates to the known result of the treelike networks: \( q_c = \langle k \rangle / \langle k(k-1) \rangle \). In general, \( q_c \to 1 \), as \( L \to \infty \), \( q_c^1 \) and \( q_c^2 \) are limit to 0, and Eq. (1) reduces to a second-order equation for \( q_c \). The root with physical meaning is \( q_c = \langle k \rangle / \langle k(k-1) \rangle \), which is just the percolation threshold of the treelike networks. This means that cycles of infinite length do not affect the percolation thresholds. The numerical tests for several typical data values show that, for the same number of cycles, \( q_c \) is an increasing function of \( q_c^1 + q_c^2 \). The higher the cycle order \( L \) is, the less influence it will be. Thus the most influential cycles are three-cycles, which could be expressed by the clustering coefficients.

For \( L=3 \), \( l_1=l_2=1 \). If the clustering coefficient \( C = \langle C(k) \rangle \) [20] is small enough, we may assume that two triangles could only have at most one common node. A node with degree \( k \) reached by an edge will belong to on average

\[
n_s(k) = \frac{\langle C(k)k(k-1)(k-2) + C(k)(k-1) \rangle}{2} = C(k)(k-1)^2 / 2 \text{ triangles.}
\]

Equation (1) will reduce to

\[
q_c = \frac{\langle k \rangle}{\langle k(k-1) - \left( 1 - q_c \frac{k(k-2)}{\langle k \rangle} \right) C(k)(k-1)^2 \rangle} \text{.} \tag{2}
\]

The percolation threshold \( q_c \) increases monotonically with \( \langle C(k)k(k-1)^2 \rangle \). It is straightforward that when \( C(k) \) is limited to 0, \( n_L \to 0 \), and \( q_c \) returns to \( \langle k \rangle / \langle k(k-1) \rangle \). On the other hand, if \( \langle C(k)k(k-1)^2 \rangle \) diverges, \( q_c \) maximizes to \( \langle k \rangle / \langle k(k-2) \rangle \). Figure 6 shows the dependence of the percolation threshold \( q_c \) on the clustering coefficient \( C \). The curves are from Eq. (2) and the symbols are from numerical simulations. For simulations, first a random network with prescribed degree distributions is generated; then, a rewiring process [21] is applied to achieve the required value of clustering coefficient, while keeping the degree distribution unchanged. Percolation is performed using Newman and Ziff’s algorithm [13].

Although Eq. (2) holds only for the case of small clustering coefficient \( C \), the analysis above indicates that for large \( C \), or if the network has higher-order cycles, the fraction of edges that interconnect existing nodes in a local cluster will further increase. Thus the number of efficient edges connecting to “new” nodes (in comparison to the nodes within the local cluster) may be even smaller, which results in a higher percolation threshold. Thus when a network is more clustered, or has more cycles—not only of order 3—it will be less robust.

Equation (1) shows that the most influential cycles on percolation threshold are triangles; furthermore, for the same number of edges forming cycles, there will be more triangles than higher-order cycles. Thus the decrease of the number of triangles (clustering coefficients) in the same kind of networks will be always accompanied by the drop of percola-
local clustering coefficient is much smaller for nodes with small degrees, while it follows the same scaling relation as that for normal networks when the degree is large. Thus the influence of the spectrum of clustering coefficients is not crucial. Therefore the decrease in $C$ for hollow networks is consistent with the drop of the percolation thresholds (Figs. 1 and 2).

In short, we have studied how the geographical structure affects the percolation behavior of complex networks and provided analytical understandings by generating function formalism on networks with cycles. Our study gives a general suggestion on constructing more robust real functional networks, such as the Internet, power grid networks, etc., which arrange the edges to connect neighboring nodes as far as possible. Although it may cost a little more, it will stand a much reduced risk in the case of node failures. Also, the hollowing strategy could be useful to maintain the global functions of real-world networks during some emergencies, such as epidemic occurrences, eruptions of electronical virus, cascade failures of power stations, etc.

L.Y. thanks the 100 Person Project of the Chinese Academy of Sciences, the Hong Kong Research Grants Council (RGC), and the Hong Kong Baptist University Faculty Research Grant (FRG), for their support. K.Y. thanks the Institute of Geology and Geophysics, CAS for their support. The work is supported by the China National Natural Sciences Foundation with Grant No. 49894190 of a major project and the Chinese Academy of Science with Grant No. KZCX1-sw-18 of a major project of knowledge innovation engineering.


