Robust Design of Geographical Networks based on a Population against Failures and Attacks

Yukio Hayashi
Japan Advanced Institute of Science and Technology, Email: yhayashi@jaist.ac.jp

Abstract—Robust and efficient design of networks on a realistic geographical space is one of the important issues to realize dependable communication systems. In this paper, based on a percolation theory and a geometric graph property, we investigate it in the following viewpoints: 1) network evolution according to a spatially inhomogeneous population, 2) optimal modality of degrees for the tolerant connectivity against both failures and attacks, and 3) decentralized routing within short paths. Furthermore, we point out the weakened tolerance by geographical constraints than the theoretical prediction, and propose a practical strategy to improve it by adding shortcut links between randomly chosen nodes. These properties of the bounded distance of path, the efficient routing, and the small modality to be robust will be useful for constructing future ad-hoc networks in wide area communication.

I. INTRODUCTION

Point processes give useful theoretical insights for modeling of wireless communication networks. Some mathematical treatments have been developed [1], while studying other models than e.g. Poisson Voronoi tessellations or Gibbs point process for a decomposition into some territories is in an open and potential research field. As a different approach from the conventional one, we focus on the robustness of connectivity and the efficiency of routing inspired by the progress in computer science and complex network science [2], [3].

Real complex networks, such as power-grid, airline flight-connection, and the Internet, are embedded in a metric space, and long-range links are restricted [4], [5] for economical reasons. Moreover, there exist a common topological characteristics called scale-free (SF) that follows a power-law degree distribution $P(k) \sim k^{-\gamma}$, $2 < \gamma < 3$, which consists of many nodes with low degrees and a few hubs with high degrees. Here, the degree $k$ means the number of links connected to a node. The SF structure is quite different from the conventional simple regular lattices and random graphs (also from more complicated graphs, e.g. Voronoi or Delaunay diagram), and extremely vulnerable against the intentional attacks on hubs [6]. By removing only about a few % of high-degree nodes, a set of the ruinously fragmented nodes leads to the global malfunction of communication network [7]. Thus, the design of more robust networks than the SF structure such as in the Internet or a Peer-to-Peer system is important to reduce the threat of cyber terrorism and natural disaster for our communication infrastructure.

Recently, it has been analytically and numerically shown that [8] the robustness against both random and targeted removals of nodes is improved as the modality of degrees is fewer in the multimodal networks which include the SF structure as the largest modality. Although the bimodal network with only two types of degrees is the optimal in this class of networks, we can consider any other type. Even with a small modality, the best allocation of degrees to nodes is generally unknown and it may depend on the geographical positions of nodes. Clearly, in real communication networks, the position of node is not uniformly random [4] by the preference of crowded urbanism or geographical limitations on residence. Therefore the spatial distribution is non-Poisson.

In addition, we are motivated on some geometric constructions of spatially grown SF networks [9], [10], [11], in which newly added nodes and links are determined by the positions of already existing nodes. In computer science, other geometric network models (e.g. Gabriel and $\theta$-graphs in restricted power consumption) have been proposed. Most of the researches have been devoted to algorithmic and graph theoretical issues [12] about efficient routing and linking on a general (usually, uniform randomly distributed) positions of nodes, however the percolation analysis in statistical physics should be included in the discussion, because the designs of the robust infrastructure on a realistic space and the routing scheme are closely interlaced [13]. The positions of nodes and the linkings strongly affect the distances of optimal paths and the tolerance of connectivity. Thus, including issues of the routing and the robustness, we consider how to design a communication network in realistic positions of nodes as base stations and linkings between them on an inhomogeneously distributed population.

The organization is as follows. In Section II, we introduce a new network construction according to a given distribution of spatially inhomogeneous population. We consider an incremental design of communication network and a good property for decentralized routings. In Section III, we derive a theoretical prediction for the breaking of whole connectivity by the removals of nodes, from the percolation analysis under an assumption of tree-like structure. In Section IV, we numerically investigate the robustness of the geographical networks. In particular, we point out more weakened tolerance by geographical constraints on local cycles than the theoretical prediction, and propose a practical strategy to improve it by adding a small fraction of shortcuts. In Section V, we summarize these results and briefly discuss the further studies.
Let us consider the basic process of network evolution as follows. It is based on a point process for load balancing of incremental communication requests by stochastic subdivisions of triangle. Each node of triangle corresponds to a base station for transferring messages and the link between them corresponds to a wireless or wired communication line, however the technical details to distinguish them at physical device level is beyond our current scope of network modeling.

**Step 0:** Set an initial triangulation of any polygonal region which consists of equilateral triangles.

**Step 1:** At each time step, a triangle is chosen with a probability proportional to the population (by summing up the number of people) in the corresponding space.

**Step 2:** Then, as shown in Figs. 1 and 2(a), smaller four triangles are created by adding a few facility nodes at the intermediate points on communication links of the chosen triangle. This procedure is for a division of service area.

**Step 3:** Return to Step 1.

In the subdivision, we use a mesh data of population statistics \( (8^2 \times 10^2 \times 4 = 256000 \text{ blocks for } 80km^2 \text{ in Fukui-Kanazawa area provided by Japan Statistical Association}) \), and recalculate the mapping between the blocks and each triangle to count the number of people in the triangle space. It is natural that the amount of communication requests depends on the activities of people, therefore approximated to be proportional to the population in the area.

A network with narrow triangles gives rise to long paths, such construction is not suitable.

For a routing, because of the planarity of network, we can apply the efficient decentralized algorithm [15] that guarantees the delivery of message using only local information based on the positions of source, terminal, and the adjacencies of the current node on a path. In recent technologies, the positions are measured by using GPS or other methods, then the shortest path may be expected. Note that the smallest degree \( k_1 \) are fixed on the initial nodes at the existing probability \( p_1 = 3/N \) (or \( 6/N \) for the initial hexagon) in this rule of network generation.

On the other hand, as a good property known in computer science, the proposed network becomes the \( t \)-spanner [14] with the stretch factor \( t = 2 \times \frac{4 \sqrt{3}}{3} \approx 2.3094 \), since the equilateral property hold in spite of various sizes of triangle. In other words, the network consists of the most fat triangles, then the length of the shortest path between any nodes \( u \) and \( v \) is bounded by \( t \) times the direct Euclidean distance \( d(u, v) \).

A network with narrow triangles gives rise to long paths, such the construction is not suitable.
path can be found in a proper mixing of upper and lower chains extracted from edges of the faces as shown in Fig. 3. The next forwarding node from each node on the upper and lower chains is determined by only the positions of adjacent nodes and the distance from a given source. Obviously, other paths on the edges out from the faces become longer.

This routing algorithm [15] on a planar network acts in a decentralized manner, a global information such as the static routing table in the Internet’s TCP/IP protocol is not necessary. Moreover, the algorithm can be extended for adding shortcuts discussed in Section IV. When some shortcut links are added to an original planar network, they directly connect two nodes in the shortest distance, but may cross to other links. Thus, we search it in the ellipsoid whose chord is defined by the distance of the previous shortest path \( l_s \) on the edges of faces that intersect the straight line between the source and terminal as two focuses of the ellipsoid (see Fig. 4). Since an outer node from the ellipsoid is further position from the source and the terminal, the distance of a path through the outer node exceeds that of \( l_s \). The extended procedures are outlined as follows [16].

- Find the shortest path \( l_s \) on the original planar network without shortcuts.
- Then search shorter one including shortcuts in the ellipsoid with the additionally stored information of node sequences on the paths.
- Through backtracks from the terminal to the source in the above process, prune the outer nodes of the ellipsoid or the inner nodes on longer paths than \( l_s \) by using the positions.

We anticipate the additional steps for searching are not so much as visiting almost all nodes, when the fraction of shortcuts is low. Moreover, by adding shortcuts between randomly chosen nodes, the robustness of connectivity can be considerably improved as shown in later.

![Fig. 3. Efficient face routing on a planar network. The (red) shortest path between a source and terminal can be found from edges of the (blue) shaded faces that intersect the (green) dashed straight line. The two paths on the edges above and below the line are the upper and the lower chains, respectively.](image)

We emphasize that such geographical protocols are not only successively used for a message delivery in social friendships [17] but also very promising for constructing wireless networks with higher efficiency and scalability in dynamic environment [18].

III. ROBUSTNESS OF THE CONNECTIVITY

In this section, we prepare for studying the robustness in the geographical networks. We apply the percolation analysis [19], [20], [21] to a general trimodal network, and derive the critical fraction of removed nodes at which the network undergoes a sharp phase transition from a connected or percolating phase to a fragmented phase.

A. Random Failures

At the breaking point of the giant component (GC) for the fraction \( f_r \) of random failures of nodes in any network, the critical fraction is given by [19], [20]

\[
f_r = 1 - \frac{1}{\kappa + 1}, \tag{1}
\]

where \( \kappa \) is defined as \( \langle k^2 \rangle / \langle k \rangle \) and the bracket \( \langle \cdot \rangle \) denotes the expectation by a probability degree distribution \( P(k) \) of the network. Note that this analysis by using a generating function or Bayes rule neglects the presence of cycles.

In the case of a trimodal network, we have

\[
\langle k^2 \rangle = k_1^2 p_1 + k_2^2 p_2 + k_3^2 p_3 = p_1 (k_1^2 + \alpha k_2^2 + \beta k_3^2),
\]

and

\[
\langle k \rangle = k_1 p_1 + k_2 p_2 + k_3 p_3 = \frac{1}{1 + \alpha + \beta} (k_1 + \alpha k_2 + \beta k_3). \tag{2}
\]

In the right-hand sides, \( \sum_{i=1}^{3} p_i = 1, \) \( p_2 \) is defined as \( \alpha p_1, \) and \( p_3 \) is defined as \( \beta p_1 \) are used, where \( \alpha, \beta > 0, \) \( p_i \) denotes the existing probability of node with each degree \( k_i. \) We investigate the range of parameter which controls the existing probability of degree \( k_i. \) Remember that the allocation of degrees is crucial.

![Fig. 4. Illustration of the extended routing for adding shortcut links. The solid (red) line is the shortest path \( l_s \) on the edges of shaded faces in the original planar network, and the distance is equal to that of the dashed (red) chord between two focuses of the (blue) dashed ellipsoid. More short path including a shortcut link is obtained in the (blue line of) two hops.](image)
for the robustness [8]. On the \(\alpha\)-\(\beta\) coordinates, from the cross point of Eq.(2) and

\[
p_1 = \frac{1}{1 + \alpha + \beta} \geq 1/N,
\]

which means the existence of at least one node with the degree \(k_1\), we obtain the maximal value of \(\alpha\):

\[
\alpha_{\text{max}} = \frac{N(k_3 - \langle k \rangle)}{k_3 - k_2} = \frac{N(6 - \langle k \rangle) - 4}{N(6 - \langle k \rangle) - 3}.
\]

The two expressions in the right-hand side correspond to the geographical networks generated from the initial configurations of triangle and hexagon.

B. Intentional Attacks

By the intentional attacks on hubs, nodes are removed in the decreasing order of degrees. We assume that the degree of adjacent node to the hub is independent of the hub’s degree, then only the number of the adjacent nodes (the number of the outlinks from the hub) is discussed. Therefore, the damage is equivalent to the random removal of a fraction \(p_t\) of nodes where \(p_t\) is the ratio of the sum of the outlinks from the removed hubs by the fraction \(f_t\) of attacks and the total number of links before the removal. From Eq.(1) on the assumption of a tree-like structure, the critical fraction \(\tilde{p}_t\) follows [20], [21]

\[
1 - \tilde{p}_t = \frac{1}{k'} - 1,
\]

where \(k' \equiv \langle k'^2 \rangle / \langle k \rangle\), and these expectations \(\langle k'^2 \rangle\) and \(\langle k' \rangle\) are defined after the removal due to the remarkable change of degree distribution. In a trimodal network, we consider the following two cases that depend on the range of removal.

1) Case of \(f_t < p_3\): In this case, the nodes with degree \(k_3\) still remain in the fraction \(p_3 - f_t\) after removing of the fraction \(f_t\) of nodes in the decreasing order of degree, while the nodes with degrees \(k_1\) and \(k_2\) are almost all remained. Since the attacks are equivalent to the fraction \(\tilde{p}_t\) of removed outlinks, by substituting

\[
\langle k' \rangle = p_1 k_1 + p_2 k_2 + (p_3 - f_t) k_3 = \langle k \rangle - k_3 f_t,
\]

and

\[
\langle k'^2 \rangle = p_1 k_1^2 + p_2 k_2^2 + (p_3 - f_t) k_3^2 = \langle k^2 \rangle - k_3^2 f_t,
\]

into Eq. (3), for the quadratic equation

\[
Af_t^2 + Bf_t + C = 0,
\]

we have the solution

\[
f_t = \frac{-B \pm \sqrt{D}}{2A} > 0,
\]

because of \(D \equiv B^2 - 4AC = \left( \frac{\langle k \rangle}{k_2} A - \frac{k_3}{k_2} C \right)^2 \geq 0\), where

\[
A \equiv k_3^2 (k_3 - 1) > 0,
\]

\[
B \equiv k_3 (\langle k \rangle (1 - k_3) - \langle k^2 \rangle + 2 \langle k \rangle) = -\left( \frac{\langle k \rangle}{k_3} A + \frac{k_3}{k_2} C \right) < 0,
\]

and

\[
C \equiv \langle k \rangle (\langle k^2 \rangle - 2 \langle k \rangle) > 0.
\]

2) Case of \(p_3 \leq f_t < p_2\): In this case, the nodes with degree \(k_3\) are completely removed, and the nodes with degree \(k_2\) are also removed in the fraction of \(f_t - p_3\). The nodes with degree \(k_2\) are remained in the fraction \(p_2 - (f_t - p_3)\), while the nodes with degree \(k_1\) are almost all remained. Consequently, the number of the removed outlinks are \((k_2 (f_t - p_3) + k_3 p_3) N\). Thus, by substituting

\[
\tilde{p}_t = \frac{k_2 (f_t - p_3) + k_3 p_3}{\langle k \rangle},
\]

\[
\langle k' \rangle = p_1 k_1 + (p_2 + p_3 - f_t) k_2 = F - k_2 f_t,
\]

and

\[
\langle k'^2 \rangle = p_1 k_1^2 + (p_2 + p_3 - f_t) k_2^2
\]

into Eq.(3), we have the solution \(f_t\) from Eq.(4), where

\[
A \equiv k_2 (k_2 - 1) > 0,
\]

\[
B \equiv k_2 (F(1 - k_2) - E + \langle k \rangle) = -\left( \frac{F}{k_2} A + \frac{k_2}{F} C \right) < 0,
\]

\[
C \equiv F (E - \langle k \rangle) > 0,
\]

\[
D \equiv B^2 - 4AC = \left( \frac{F}{k_2} A - \frac{k_2}{F} C \right)^2 \geq 0,
\]

\[
E \equiv p_1 k_1^2 + (p_2 + p_3) k_2^2 - p_1 k_1 - (p_2 + p_3) k_2,
\]

and

\[
F \equiv p_1 k_1 + (p_2 + p_3) k_2 > 0.
\]

IV. NUMERICAL ANALYSIS

We aim to maximize \(f_t = f_r + f_o\) to be optimal network for both random failures and targeted attacks on nodes with large degrees as robust as possible on a spatially inhomogeneous distribution of population.

For the geographical networks defined in the preceding section, we numerically estimate the total numbers of nodes \(N(t)\) and links \(M(t)\) at time step \(t\) over 100 realizations. They are grown as

\[
N(t) = N(0) + 2.35t,
\]

and

\[
M(t) = M(0) + 5.34t,
\]

where \(N(0) = 6\) and \(M(0) = 9\) as the initial triangle (or \(N(0) = 7\) and \(M(0) = 12\) as the initial hexagon). Thus, in the subdivision, two or three nodes per one step are added with high frequency, and the average degree is \(\langle k \rangle = 2M(t)/N(t) \approx 4.54\). Th total \((k)N/2\) links are more than twice of \(N - 1\) links on a tree-like structure, there possibly exist many cycles.

On the fixed \(\langle k \rangle\) at a size \(N\), against the failures and the attacks, we estimate the tolerance of connectivity in trimodal
networks with a general allocation of the degrees \( k_1 = 2 \) (or 3 for the initial hexagon), \( k_2 = 4 \), and \( k_3 = 6 \). Note that a special allocation correspond to the geographical networks generated by the subdivisions of equilateral triangle.

Figure 5(a) shows the critical fractions calculated from Eqs. (1) and (4) for the allocations controlled in the range \( 0 < \alpha < \alpha_{\text{max}} \). These curves are almost saturated in \( \alpha > 0.5 \). In comparison with the Insets, the fraction \( f_T \) of attacks is dominant in the total \( f_T \), however there exists a trade-off between the curves of \( f_T \) and \( f_r \). Intuitively, the majority of small degree nodes contributes to be robust against intentional attacks because of the small damage of removed links, on the contrary, the minority of large degree nodes is so against random failures because of the maintaining of many links. Figure 5(b) shows that the number of nodes with degree \( k_2 \) is larger as increasing the value of \( \alpha \), because \( p_2/p_3 > 1 \) means the existing of relatively much more nodes with degree \( k_2 \) than that with degree \( k_3 \). Remember that the nodes with degree \( k_1 \) are fixed and \( p_2 = \alpha p_1 \) and \( p_3 = \beta p_1 \) denote the probabilities of nodes with degrees \( k_2 \) and \( k_3 \). Figure 5(c) shows the varying of each \( p_i \). Thus, the increase of nodes with the middle degree \( k_2 \) leads to slightly smaller \( f_r \) and larger \( f_T \), and consequently to the larger \( f_T \) as more robust connectivity.

In the geographical networks with constraints by the subdivision according to the population density, we obtain \( p_2/p_3 \approx 2.6 \) at \( \alpha/\alpha_{\text{max}} \approx 0.3 \) from computer simulations for actually constructed networks. Therefore, in Fig. 5(b) as the theoretically calculated prediction, it becomes nearly optimal with \( f_T \approx 1.32 \sim 1.35 \), which means the strong tolerant connectivity until both 60-70% of the removals by random failures and targeted attacks.

However, the simulation results for the removals as in Fig. 6 marked with blue + and magenta × show that the size \( S \) of the GC is rapidly decreased at the fraction \( f \approx 0.4 \) before the theoretically predicted critical points \( f_r \approx 0.7 \) and \( f_T \approx 0.6 \). We remark the difference between the simulation results in realistic networks with geographically constrained local cycles and the theoretical prediction under the assumption of a tree-like structure. A similar vulnerability caused by geographical constraints with local cycles has been found in a family of SF networks embedded in a planar space [22] and SF networks on a lattice [23].

To make the effect of constraints clearly, we investigate the non-geographical rewired versions under the same degree distribution (of course the same average degree \( \langle k \rangle \)). In the rewiring, each other’s node of randomly chosen links in all pairs is exchanged as in Fig. 7, therefore the geographical constraints are entirely reduced. In general, the rewired version of a network is the null model that depend on only the degree distribution ignoring the other topological structures: cycles, degree-degree correlation, fractal or hierarchical substructure, diameter of graph, etc. Figure 6 marked with red △ and green ◢ shows that the critical fraction at the peak of \( \langle s \rangle \) is increased to the theoretical levels \( f_r \approx 0.6 \) and \( f_T \approx 0.7 \). Such improvement of the robustness is consistent with the result for the geographical SF networks [22].

Fig. 5. Connectivity in a general trimodal network for varying the allocation of the degrees by the value of \( \alpha \). (a) The critical fraction of \( f_T \) or (b) the ratio of \( p_2/p_3 \) vs. the parameter \( \alpha \). Inset shows the trade-off between increasing and decreasing in the fraction \( f_T \) for the random failures at left or \( f_r \) for the intentional attacks at right. The solid (red), dashed (green), dotted (blue) lines correspond to the results for \( N = 100, 1000, 10000 \). (c) The varying of \( p_1 \) (solid red), \( p_2 \) (dashed green), and \( p_3 \) (dotted blue) at \( N = 100 \) for the parameter \( \alpha \). Inset shows the result at \( N = 1000 \).
Fig. 6. Comparison with the geographical and the non-geographical rewired networks. (a) Relative size $S/N$ of the GC after the removal of nodes. The marks of plus(blue) and cross(magenta) correspond to the results for the geographical networks against random failures and intentional attacks. The marks of upper triangle(red) and down triangle(green) correspond to ones for the non-geographical rewired versions against them, respectively. (b) The average size $\langle s \rangle$ of isolated clusters except the GC. The peak indicates the critical point for the breaking the whole connectivity. These results are obtained in 100 realizations at $N = 1000$ from the initial triangle.

Although the fully rewiring is better for the robustness, it completely ignores the positions of nodes and the distances of links. As another practical strategy, it is expected that adding a small fraction of shortcuts between randomly chosen nodes has a similar effect to the rewiring [16]. Indeed, Figs. 8 and 9 show the improvement of robustness against both random failures and intentional attacks. As increasing the shortcut rate marked from red $\bigcirc$ to black $\blacktriangledown$, a larger GC remains and the breaking point at the peak of $\langle s \rangle$ is shifted to a larger fraction $f$ of the removal. Only about 10% of the adding reaches the similar level of the theoretical predictions $f_r \approx 0.7$, $f_t \approx 0.6$, and of the non-geographical rewired version (compare with Fig. 6 marked with $\bigtriangleup$ and $\blacktriangleleft$). Since the shortcuts bridge isolated clusters by the removals, the connectivity is still maintained as illustrated in Fig. 10.

V. CONCLUSION

According to spatially distributed communication requests based on a given population density, we have proposed an evolutionary construction of geographical network by iterative division of equilateral triangles. The obtained results summarized as follows. We note that they are consistent with the results on other population densities than our example of Fig. 2(b) as mentioned in Appendix.

- The proposed networks have suitable properties of short paths as the $t$-spanner [14], efficient decentralized routing [15], and small modality of degrees to be robust connectivity [8] for wireless communication.
- By using the percolation theory [19], [20], [21], we have estimated the critical fractions of removed nodes for the breaking of the GC against random failures and intentional attacks in the decreasing order of degrees.
- Although the critical fractions almost coincide with the simulation results in the non-geographical rewired version, they become smaller in the geographical networks with locally restricted cycles.
- To improve the vulnerability caused by the geographical constraints [22], [23], we have considered a practical strategy by adding a small fraction of shortcuts between randomly chosen nodes [16], and numerically confirmed the effect.

These results will be useful for the design of ad-hoc networks with efficiency, scalability, and tolerance of connectivity in wide area communication.

Since our trimodal model has the maximum degree not extremely large to other degrees, instead of hub attacks, a spatial cutting into several dense parts by removing a small number of nodes on lines of large triangles may be considerable. Even in such cases, we expect that shortcuts effectively work to bridge them. More detail analysis will be a further study including how to find the structural vulnerable points.

ACKNOWLEDGMENT

The author would like to thank anonymous reviewers for valuable comments to revise the manuscript, and Yasumasa
**Fig. 8.** Damages by random failures in the geographical networks and the shortcut versions at $N = 1000$. (a) Relative size $S/N$ of the GC. The marks of open circle(red), plus(green), cross(cyan), upper triangle(orange), diamond(magenta), closed circle(yellow), and down triangle(black) correspond to the results for the shortcut rate of 0, 3, 5, 7, 10, 20, and 30%. (b) The average size $\langle s \rangle$ of isolated clusters except the GC. These results are obtained in 100 realizations from the initial triangle.

Ono in my laboratory for helping the simulation. This research is supported in part by Grant-in-Aide for Scientific Research in Japan, No.18500049.

**REFERENCES**


APPENDIX

We have also investigated the robustness of connectivity in the geographical networks generated by iterative choice of a triangle with the uniformly random probability, in which various distributions of population are included (as an example: Fig. 11). These results are shown in Figs. 12 - 13, and similar to ones in Figs. 6, 8 and 9. The marks are same as in them. We note that the case of the initial hexagon is quite similar.
Fig. 13. Damages by (a) random failures and (b) intentional attacks in the geographical networks and the shortcut versions on uniformly random choices of triangle. These results are obtained in 100 realizations at $N = 1000$ from the initial triangle.