Clustering of Random Networks under Topological Constraints

Minas, Hovhannisyan
Yerevan State University
Yerevan, Armenia
e-mail: minas.hovhannisyan@gmail.com

Svetlana, Avetisyan
Yerevan State University
Yerevan, Armenia
e-mail: avetisian@bk.ru

ABSTRACT
We study the directed evolution of random networks toward increasing local connectivity while conserving the degrees of nodes using computer simulations. We show that network evolution under "quenched topological disorder" manifests itself as a cascade of transitions that result in the splitting of random network into a set of inherently connected subnets weakly coupled to one another. We interpret this phenomenon as the formation of clustering.

Keywords
Computer simulations, random networks, topological characteristics, statistical analysis

1. INTRODUCTION
Complex networks have garnered the attention of the research community due to the discovery of non-trivial topological features in many real networks [1, 2, 3]. Network concepts, models and analytic techniques have been applied in many different disciplines such as sociology, biology, physics, mathematics, and computer science [4, 5]. In addition to general topological characteristics such as the average degree of nodes, information about the structure of a complex network can be obtained by investigating the local topology (e.g., the statistics of small subgraphs) [6]. One more aspect relates to the directed evolution of a network to the states that differ from purely random networks called the Erdős-Rényi networks [7]. Of particular interest is the evolution of networks with subgraphs consisting of only a few nodes. These subgraphs simulate, from a physical point of view, multi-particle local interactions. In this case, there is competition between the "energy gain," which is associated with an increase in the number of connected subgraphs, and a loss of entropy due to the decreased number of ways of reaching an energy-preferred state. As a result, the network follows the structural rearrangements, which may be interpreted in terms of phase transitions and critical phenomena.

An early study showed that growing networks to a maximal number of fully connected three-node subgraphs (cycles of length 3) possesses criticality similar to the first-order phase transition resulting in the formation of a "dense" core surrounded by a "vapor" of nodes [8]. This phenomenon was observed under conservation of global topological characteristics on average but not local distributions, which could have changed following the evolution. Tamm et al. (2014) recently reported that critical behavior such as a first-order phase transition occurs even when the distribution of the degree of nodes is quenched and remains unchanged over the course of the evolution [9]. We hypothesize that, while networks evolve given competition between energy gain and entropy cost, the local topological constraints may act as "quenched disorder". Therefore, the network structuring other than the one presented in [9] may occur. In this paper, we show that the directed evolution under local topological constraints indeed leads to a cascade of critical phenomena, namely, the subsequent splitting of an initially homogeneous random network into a set of highly connected subnets associated with "condensed clusters."

2. THE MODEL OF DIRECTED EVOLUTION OF RANDOM NETWORKS
We consider the ER random graph model. In this model, N identical pairs of vertices are connected by undirected edges; each edge appears with an independent probability p, forming on average pC_n^2 edges. The ER random graph model is given by $G = (V, E)$, where $V(G) = \{x_1, \dots, x_N\}$ is a non-empty finite set of vertices and $E(G)$ is a finite set of edges. The model possesses a symmetric adjacency matrix A with elements $a_{ij} = 1$ if $(x_i, x_j) \in E(G)$ and $a_{ij} = 0$, otherwise. The vertices of the graph G, representing a network, are referred to as nodes, and the edges are referred to as links.

Let $T(G)$ be the number of cycles of length 3 in network G. Now, consider the process of network evolution toward the growth of $T(G)$. The process of evolution is a sequence of transformations of graphs $\{G = G_0, G_1, G_2, \dots G_k\}, k \gg 1$, i.e., a random process in which each successive graph is obtained from the previous one by switching links, which preserves the degrees of nodes (see, e.g., [6]). Figure 1 illustrates such switching.

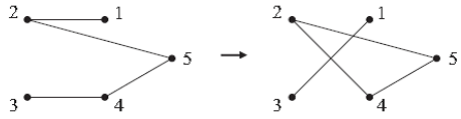


Figure 1. Switching of links, which preserves the degrees of nodes. Links 2->1, 3->4 are substituted with 1->3, 2->4, which results in the formation of a new 3-length cycle (2, 4, 5)

The value of $T(G)$ may change during switching. As a result, we let $\Delta_i = T(G_i) - T(G_{i-1})$, $1 \le i \le k$, which is a one-step change. To accept or reject network resulting from such a switch, we use the Metropolis-Hastings algorithm. Namely, the probability ω of acceptance or rejection of a switch is determined as follows:

$$\omega = \begin{cases} 1, & if\ \Delta_i \geq 0 \\ e^{\mu * \Delta_i}, & if\ \Delta_i < 0 \end{cases},$$

where μ is the parameter of model: higher μ values correspond to lower probabilities that the network is accepted with a lower value of $T(G)$. The evolution continues until stabilization, i.e., until the system reaches a state in which the number of cycles of length 3 does not tend to change over a sufficiently long series of switches, $\Phi(G, \mu)$, and the system only fluctuates around some value. It is natural to associate such a state with equilibrium. The evolution itself is represented by evolutionary trajectories, i.e., the variation of the number of cycles of length 3 along the number of elementary switches (time).

Below, we compare the evolutionary trajectories derived from one and the same initial random network but different values of the parameter μ. We proceed from the work of Tamm et al. (2014) [9]; these authors demonstrated the existence of a critical value $\mu = \mu_c$ at which the equilibrated amount of $T(G)$ increases sharply. In this paper, we study the formation of structure over the course of evolution before and after reaching the critical value $\mu = \mu_c$. In addition, for comparison, we reproduce the results of Park and Newman (2005) [8] for evolution without the degrees of nodes being conserved. In the last case, we realize elementary switching by removing one randomly selected link and forming one new one. The probability of acceptance or rejection of an elementary switch remains the same as specified above.

To study the directed evolution of random networks with and without conserving the degrees of nodes, we developed a software system that allows us to determine the critical points of the phase transitions in networks of various sizes; we are also able to explore the topological characteristics and structures of the intermediate and final networks.

3. CLUSTERING OF RANDOM NETWORKS THROGH DIRECTED EVOLUTION

The software that we developed was used to perform simulations of the process described above related to the directed evolution of the random networks of sizes 64, 128, 256 and 512. We also analyzed the topological properties and structures of these networks. Our computer experiments revealed that, above the critical point $\mu \geq \mu_c$, the evolution with conserved node degree leads to the subsequent formation of a set of strongly connected components. The formation of components is accompanied by an increase in the average clustering coefficient and the average distance between nodes. These topological changes, in combination, unambiguously point to the formation of clusters, which are well visualized.

Below, we illustrate our typical results of the splitting of an initially homogeneous network into clusters for a network of size 64. Similar results were observed for networks with 128, 256 and 512 nodes. For comparison, we show the structure of networks that are formed without topological constraints, i.e., without conserving the degree of nodes. In all of the examples below, the initial networks were generated randomly with probability $p = 0.3$ for the link formation, i.e., above the percolation threshold.

3.1. Evolution with conservation of the node degree

Figure 2a shows the evolutionary trajectories obtained for one and the same initial network G but for different values of μ above and below the critical point μ_c. Note that some trajectories possess stepwise transitions. Figure 2b shows the equilibrated amount of $T(G)$ as a function of μ. It is easy to see the sharp increase in $T(G)$ close to $\mu = 0.5$.

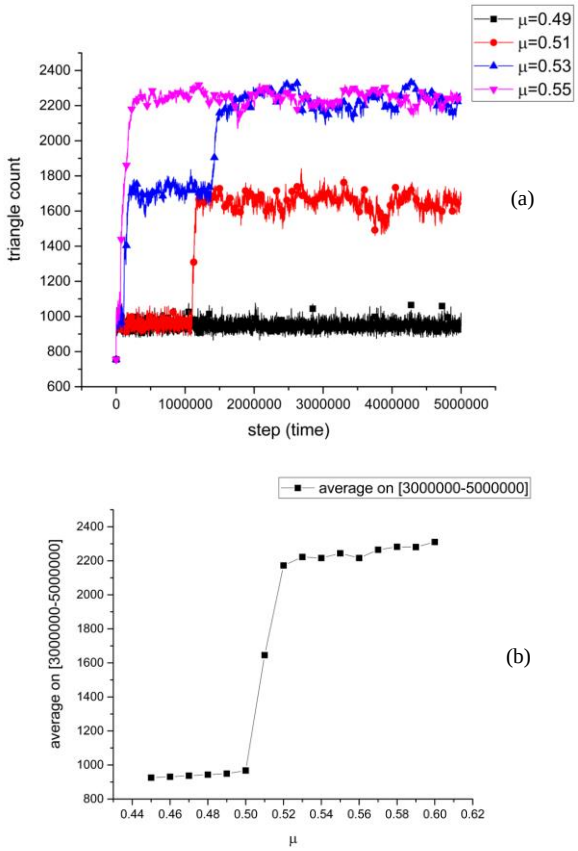
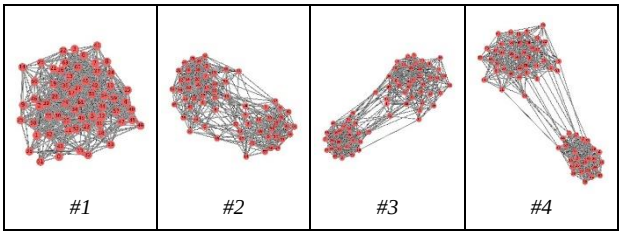


Figure 2. Evolution of random network ($N = 64, p = 0.3$) with a conserved number of node degrees. (a) – Trajectories for $\mu = 0.49$, $\mu = 0.51$, $\mu = 0.53$, $\mu = 0.55$. The transition is at $\mu = 0.51$ and $\mu = 0.53$. (b) – Equilibrated amount of $T(G)$ as a function of μ.

The trajectories show that above the critical point, for instance, at $\mu = 0.53$, a set of transitions occurs with an increase in the number of 3-node cycles. This finding is explained by the structure analysis of the networks generated in the course of evolution (Table 1). The structuring of the network above the critical point is related to the splitting of the network into strongly connected components. In the illustrated example, the network is first divided into two clusters, one of which is smaller but denser and contains nodes with the highest degrees (the empirical probability of hitting such nodes into the dense cluster exceeds 0.7). In the next stage of evolution, the denser cluster undergoes minor changes but the larger second cluster splits into two smaller ones. Additional "condensation" of nodes is not accompanied by the formation of new clusters.



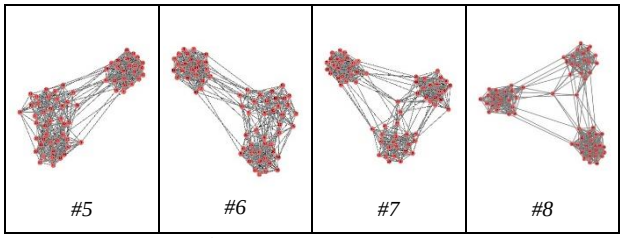
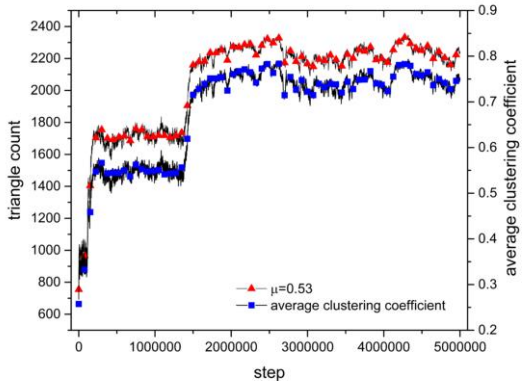


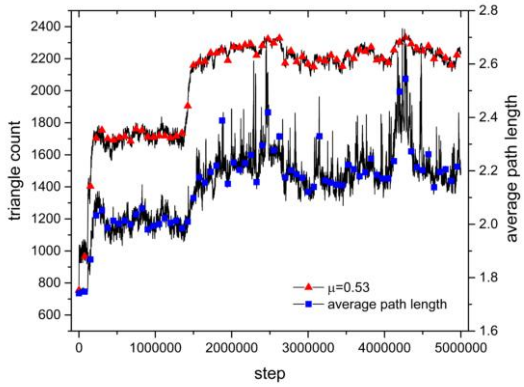
Table 1. Splitting of the random network into clusters given conservation of node degrees

In the illustrated example, the final structure represents three dense, small clusters linked to each other much weaker (on the order of magnitude less) than within the clusters.

In addition to the trajectories we also investigated the changes in the topological characteristics of the network over the course of the cluster formation. Below, we provide the data that correspond to the same evolutionary trajectory for which structures were visualized in Table 1. Figure 3 compares the change in the number of cycles of length 3 along the time with appropriate modifications of the average distance between the nodes and the clustering coefficient of the network.



(a)



(b)

Figure 3. Comparison of different topological properties in the case of network evolution given conservation of node degrees. The top curves in both panels relate to the number of 3-nodes cycles; the bottom curves relate to the average clustering coefficient (Figure 3a) and the average path length (Figure 3b).

It can be seen that the number of cycles of length 3 clearly correlates with the clustering coefficient and the average distance between nodes. All three characteristics behave in a synchronous manner.

3.2. Evolution without conservation of the node degree

For comparison, Figure 4 shows the results of the evolution of the same initial network G without conservation of the node degree.

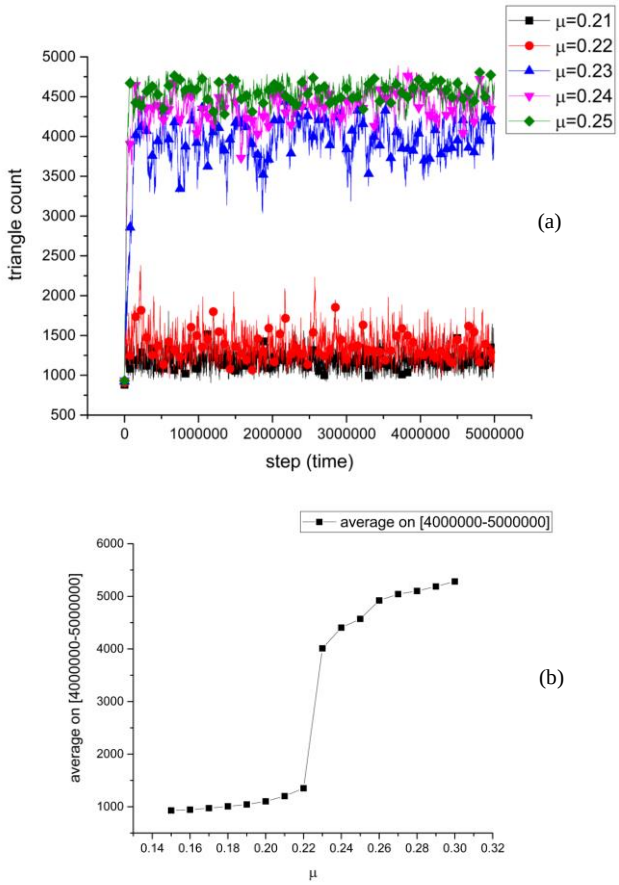


Figure 4. Evolution of a random network ($N = 64, p = 0.3$) without conservation of node degree. (a) – Trajectories for $\mu = 0.21$, $\mu = 0.22$, $\mu = 0.23$, $\mu = 0.24$ and $\mu = 0.25$. (b) – The equilibrated amount of $T(G)$ as a function of μ. The critical point is close to $\mu = 0.23$.

Here we recover a significantly different picture. The structures presented in Table 2 illustrate the fact that there is always one strongly connected component surrounded by a "vapor" of nodes.

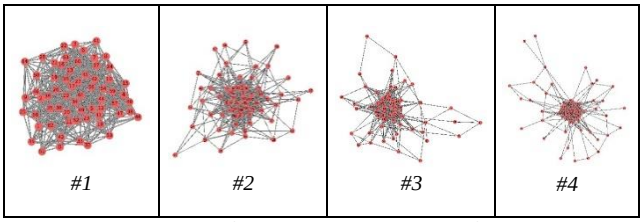
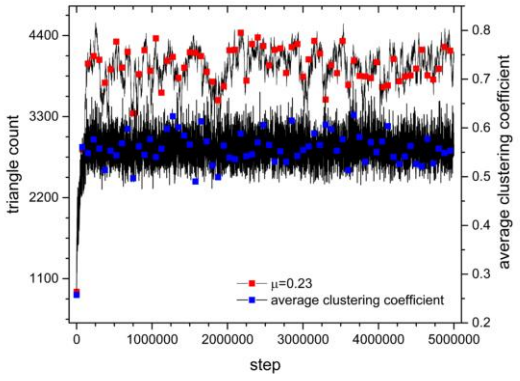


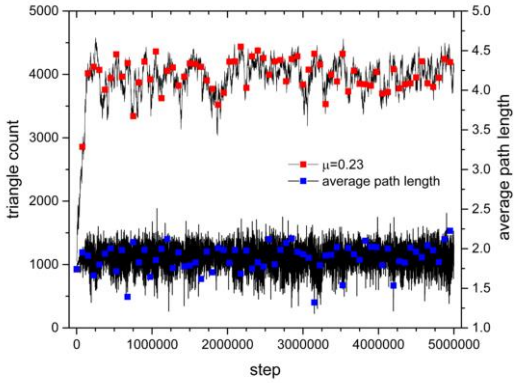
Table 2. Typical picture of the network condensation if the node degree is not conserved.

The following data correspond to the same evolutionary trajectory for which structures were visualized in Table 2. Figure 5 compares the changes in the number of cycles of

length 3 as a function of time with modifications to the average distance between the nodes and the clustering coefficient of the network.



(a)



(b)

Figure 5. Comparison of different topological properties in the case of network evolution without conservation of the node degree. The top curves in both panels represent the number of 3-nodes cycles; the bottom curves relate to the average clustering coefficient (Figure 5a) and the average path length (Figure 5b).

In this case, an increase in the average distance between nodes is not observed.

4. CONCLUSION

Our computer simulation (with developed software) studies of the evolution of networks with quenched topological disorder have shown that such evolution can be accompanied by the formation of internally connected components such as structural or functional modules. Our observations are able to shed light on the inhomogeneity in natural networks, which derives from selection or directed evolution.

5. ACKNOWLEDGEMENTS

We express our gratitude to V. Avetisov and S. Nechaev for their useful discussions and valuable comments.

REFERENCES

[1] R. Albert, A. L. Barabasi, “Statistical mechanics of complex networks”, *Reviews of Modern Physics* 74, pp. 47-97, 2002.

[2] S. N. Dorogovtsev, J. F. F. Mendes, “Evolution of networks”, *Advances in Physics* 51, pp. 1079-1187, 2002.

[3] M. E. J. Newman, “The structure and function of complex networks”, *SIAM Review* 45, pp. 167-256, 2003.

[4] D. J. Watts, S. H. Strogatz, “Collective dynamics of 'small-world' networks”, *Nature* 393, pp. 440-442, 1998.

[5] K.-I. Goh, B. Kahng, D. Kim, “Spectra and eigenvectors of scale-free networks”, *Phys. Rev. E* 64, 051903, 2001.

[6] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, U. Alon, “Superfamilies of evolved and designed networks”, *Science* 303(5663), pp. 1538–1542, 2004.

[7] P. Erdős, A. Rényi, “On the evolution of random graphs”, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5, pp. 17-61, 1960.

[8] J. Park, M. E. J. Newman, “Solution for the properties of a clustered network,” *Phys. Rev. E* 72, 026136, 2005.

[9] M. V. Tamm, A. B. Shkarin, V. A. Avetisov, O. V. Valba, S. K. Nechaev, “Islands of stability in motif distributions of random networks”, *Phys. Rev. Lett.*, 113(9), 095701, 2014.