PROCEEDINGS OF THE YEREVAN STATE UNIVERSITY

Physical and Mathematical Sciences 2018, **52**(1), p. 34–40

I n f o r m a t i c s

ON CONNECTED COMPONENT DISTRIBUTION OF RANDOM BLOCK-HIERARCHICAL NETWORKS USING THE AUTOMATED SYSTEM *XRANDNET*

A. G. KOCHARYAN *

Chair of Programming and Information Technologies YSU, Armenia

In the paper the result of a research done by using the automated system *xRandNet* is presented, which is designed and implemented for generating and analyzing the main topological properties of some hierarchical models of random networks. The research is related to the connected component distribution of random block-hierarchical networks, which are quite new objects in the random network theory.

MSC2010: 68N30.

***Keywords*:** random networks, block-hierarchical networks, connected component distribution, automated system.

Introduction. In various fields of science, the number of so-called complex systems, which as a rule, are unconventional both in structure and in the nature of changes, is increasing. Among them are biological systems (in particular, biopolymers, DNA, RNA, proteins), technical systems, social, economic ones and others [1]. In the early 1960s Paul Erdös and Alfréd Rényi came up with a suggestion to use probabilistic methods in the study of networks. This was applied in the modeling of complex systems, by representing the elements of the system as nodes and their interaction as links. The network is called random, if any connection in it appears due to a certain probabilistic rule. From henceforth we will only examine non-directed networks, which do not have self-references and multiple bonds. Thus, Erdös and Rényi came up with the model of random network for which the number of nodes N and probability p, defining the existence of connection between any 2 nodes of the network, are given [1].

Apart from the studies of classical random *Erdös-Rényi* networks, the investigation of new classes of networks has recently become increasingly interesting: these are *small-worlds*, for which the value of the average distance between the nodes is very small against the size of the network; *scale-free* networks, characterized by a

* E-mail: ankocharyan@gmail.com

power-law behavior of the node degree distribution and others. In the last decade it was introduced an essentially different class of random networks, called *block-hierarchical*. It turned out this class of networks can be applied in the modeling of different biological structures ranging from proteins to neuronal networks [2, 3].

The study of random networks supposes statistical research of different topological properties: average path length, diameter, average degree, average clustering coefficient, etc., as well as distributions: node degrees, clustering coefficients, connected components, etc. [1].

Automated system *Extended Random Networks (xRandNet)* [4] was developed based on the needs of random networks researches of both models: classical *Erdös-Rényi, Watts-Strogatz, Barabási-Albert*; block-hierarchical *Regular Block-Hierarchical, Non Regular Block-Hierarchical, HMN1, HMN2* [2–6], etc. Block-hierarchical models in general, and *Regular Block-Hierarchical* model in particular, are almost not investigated, and as *xRandNet* is mainly aimed to efficiently study them, it has allowed to obtain some results about the behavior of the main topological properties of block-hierarchical networks.

In the paper definitions of regular block-hierarchical networks and connected component distribution (*CC* distribution) are given, the extraction of the calculation formula of the connection probability in the regular block-hierarchical networks and the results of research on *CC* distribution are presented. A comparison of the obtained results for classical and scale-free networks well also be discussed. All the studies were conducted by using the automated system *xRandNet*.

Random Regularly Branching Block-Hierarchical Network.

Definition 1. Let b and Γ be natural numbers, $b > 1$. For the given b and Γ a class of regularly branching block-hierarchical networks $\Re_{b,\Gamma}$ is defined as follows. The number of nodes in the network $G_{b,\Gamma} \in \Re_{b,\Gamma}$ is b^{Γ}. The network is constructed by levels. On every new level γ, $0 \le \gamma \le \Gamma$, new clusters (subnets) are formed via merging the clusters formed on the previous level, and by introducing new connections between them via joining some of them (Fig. 1) [5]. Random block-hierarchical network is defined by the probability of the occurrence of the connections, which varies from level to level and is given by:

$$q_\gamma = b^{-\mu\gamma}, \tag{1}$$

where μ is the parameter of the model (network density) and γ is the cluster level for which the connections are formed.

Definition 2. In an undirected network nodes i and j are call connected, if there is a path between them. A network is connected, if all pairs of nodes in the network are connected. A connected component is a subset of nodes in a network, so that there is a path between any two nodes that belong to the component, but one cannot add any more nodes having the same property [7]. By the distribution of connected components (*CC* distribution) means the dependence of the average counts of connected components $\langle CC\rangle_s$ on the size of the connected component s.

In other words, it describes the distribution of connected components of various sizes in the network.

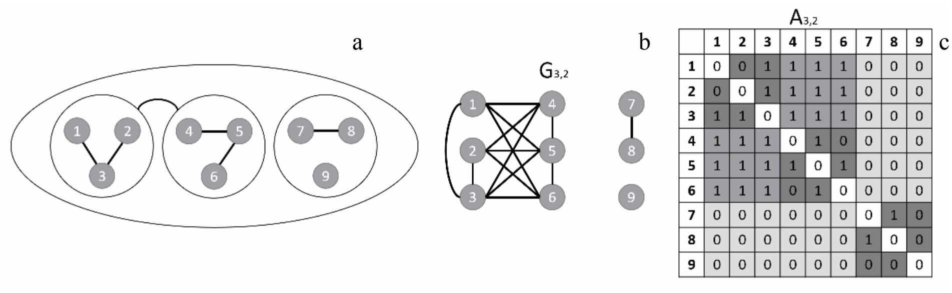

Fig. 1. Block-hierarchical network $G_{3,2}$: a) clusters view; b) network view; c) adjacency matrix $A_{3,2}$ view.

Probability of Connection between 2 Nodes in the Regular Block-Hierarchical Network. The definition of block-hierarchical network and Eq. (1) shows it clear that for this model parameter μ, network density, in a probabilistic one (determining the probability of connection). In order to compare the behavior of the topological properties of this model with the corresponding behavior of the classical model it is important to define how the parameter μ is mapped in terms of the connection probabilities between any 2 nodes in the network: parameter p in *Erdös-Rényi* model.

Consider the block-hierarchical network $G_{b,\Gamma}$. Suppose $V(G_{b,\Gamma}) = \{x_1, \ldots, x_N\}$ is a non-empty finite set of nodes, and $E(G_{b,\Gamma})$ is unordered pairs of different elements from $V(G_{b,\Gamma})$, set of connections. Let us also denote by $Join_\gamma, 1 \leq \gamma \leq \Gamma$, the set of pairs contained in by same cluster of the level γ, but not contained in a cluster of a level $\gamma - 1$. It is clear that the connections in the γ-level cluster may only occur between such pairs of nodes. We call the pair from the set $Join_\gamma$ "good" for the level γ. It is clear that $E(G_{b,\Gamma}) \subseteq \sum_{\gamma=1}^{\Gamma} Join_\gamma$. Denote by $QJoin_\gamma$ the probability of the pair of nodes from $V(G_{b,\Gamma})$ to be "good" for the level γ (Fig. 2).

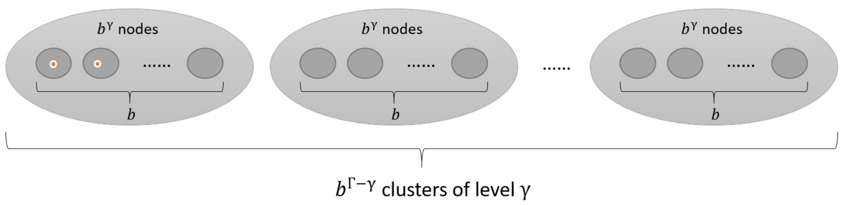

Fig. 2. Structure of the level γ, $1 \leq \gamma \leq \Gamma$, in block-hierarchical network.

To yield "good" pairs a cluster of the level γ is chosen in $b^{\Gamma-\gamma}$ ways. Out of the selected clusters 2 clusters of the level $\gamma - 1$ are chosen in C_b^2 ways, and out of the selected 2 clusters any nodes are chosen in $b^{\gamma-1}$ ways for every node. Thus

$$QJoin_\gamma = \frac{b^{\Gamma-\gamma} C_b^2 (b^{\gamma-1})(b^{\gamma-1})}{C_N^2}, \tag{2}$$

where C_N^2 is the number of all possible pairs of nodes.

Proposition . Suppose $G_{b,\Gamma}$ is a regular block-hierarchical network. Denote by Q_γ the probability of connections of 2 nodes of the network $G_{b,\Gamma}$ in the i-cluster of the level $\gamma, 1 \leq \gamma \leq \Gamma$. Then

$$Q_\gamma = \frac{(b-1)}{(N-1)b} \cdot \sum_{i=1}^{\gamma} (b^{1-\mu})^i. \tag{3}$$

Proof. The proof will be conducted by induction on the levels $\gamma, 1 \leq \gamma \leq \Gamma$.

Suppose $\gamma = 1$. On the first level N nodes are split into groups of b nodes, clusters of the first level. Inside each cluster of the first level the nodes are connected with a probability of q_1. Thus the probability of connection in the cluster of the first level equals $Q_1 = q_1 QJoin_1$.

Suppose $\gamma = 2$. On the second level the clusters of the first level are split into groups of b-clusters of the second level. Inside each cluster of the second level the clusters of the first level are connected with a probability of q_2. Thus the probability of connection in the cluster of the second level equals $Q_2 = q_2 QJoin_2 + Q_1$.

Likewise, for the level $\gamma > 2$ the probability of connection in the cluster of level γ equals $Q_\gamma = q_\gamma QJoin_\gamma + Q_{\gamma-1}$.

Solving the resulting recurrence relation, we find

$$Q_\gamma = \sum_{i=1}^{\gamma} q_i QJoin_i. \tag{4}$$

Using (1) and (2),

$$Q_{\gamma,i} = \sum_{i=1}^{\gamma} q_i QJoin_i = \sum_{i=1}^{\gamma} b^{-i\mu} \cdot \frac{b^{\Gamma-i} C_b^2 b^{i-1} b^{i-1}}{C_N^2} = \sum_{i=1}^{\gamma} \frac{C_b^2}{C_N^2} \cdot b^{i-i\mu+\Gamma-2} =$$

$$= \frac{C_b^2}{C_N^2} \cdot \sum_{i=1}^{\gamma} b^{\Gamma-2} b^{i-i\mu} = \frac{C_b^2}{C_N^2} \cdot b^{\Gamma-2} \sum_{i=1}^{\gamma} b^{i(1-\mu)} = \frac{C_b^2}{C_N^2} \cdot b^{\Gamma-2} \sum_{i=1}^{\gamma} (b^{1-\mu})^i =$$

$$= \frac{(b-1)}{(N-1)b} \cdot \sum_{i=1}^{\gamma} (b^{1-\mu})^i.$$

For the whole network (3) appears as follows

$$Q_\Gamma = \frac{(b-1)}{(N-1)b} \cdot \sum_{i=1}^{\Gamma} (b^{1-\mu})^i. \tag{5}$$

Connected Component Distribution of Regular Block-Hierarchical Networks. It is known for the connectedness in *Erdös-Rényi* networks that [7]:

– a connected component of size more than $O(\log N)$ does not appear in the network for $p < 1/N$ – subcritical phase;

– a connected component of size $O(N^{(2/3)})$ appears in the networks for $p = 1/N$ – critical point;

– the network is connected for $p > 1/N$ – supercritical phase.

Investigation on classical and scale-free networks in the subcritical phase shows power-law dependence of $\langle CC \rangle_s$ on the component size (Fig. 3), and one can see that the absolute value of slope decreases with the connection probability growth.

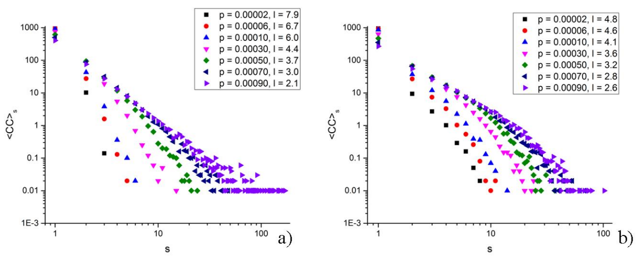

Fig. 3. Power-law behavior $\langle CC \rangle_s \sim s^{-l}$ in log-log plot for the networks of *Erdös-Rényi* model (a) and *Barabási-Albert* (scale-free) model (b). The results are obtained using the automated system *xRandNet* on the ensemble of 100 networks of size $N = 1024$ in the subcritical phase (for the N is $p \le 1/N \approx 0,00097$).

To identify the CC distribution behavior for block-hierarchical networks on the *xRandNet* system, the research was conducted for different branching indices ($b = \{2,3\}$) with different maximal levels (for $b = 2$, $\Gamma = \{10, 11, \ldots, 20\}$ is taken and for $b = 3$, $\Gamma = \{6, 7, \ldots, 14\}$). Ensembles of 10^3 network realizations are considered. The analysis of CC distribution behavior showed that for block-hierarchical networks 2 critical points, μ_1 and μ_2, can be identified which define the different behavior of CC distribution.

– For $\mu < \mu_1$, CC distribution is growing, wherein the behavior can be described as power-law (Fig. 4, a), namely $\langle CC \rangle_s \sim s^l$, where s is the size of the cluster.

– For $\mu_1 < \mu < \mu_2$ the CC distribution graph shows a peak (defining the size of the cluster, which occurs more frequently than all others (Fig. 4, b); the asterisks refer to the peaks).

– For $\mu > \mu_2$ CC distribution is starting to go down following the power-law rule as in the classical model of $\langle CC \rangle_s \sim s^{-l}$ (Fig. 4, c).

– In case of further increase $\mu > 10$ (for all networks this value did not exceed 10) the network has hardly any connections; it falls into single nodes.

– It is obvious, that if $\mu = 0$ (which corresponds to $p = 1$), the network is complete.

The change in the CC distribution behavior is tracked changing the density parameter with proper steps starting from 0. In case of a very small growth of the network density (which corresponds to the decrease in the probability of connection between the peak near 1 (this follows from (5)), even though the network remains dense, it immediately loses its connectivity. This behavior is drastically different from that of the classical network and is explained by the absence of connections on the last level in the block-hierarchical network, which already results in the emergence of discrete components. One can see from the graph Fig. 4, a, that for the values $\mu < \mu_1$ almost always (with a probability of near 1) there appears a gigantic connectivity component of the size of N order, and the number of other small components is near-zero. As closer μ to μ_1, as biggen the number of the connected small components. This is exactly why the slope is less sharp (that is to say, the value l is going down). Nevertheless, even if μ is near μ_1, the probability of the appearance of the gigantic component is more than all the others in sum.

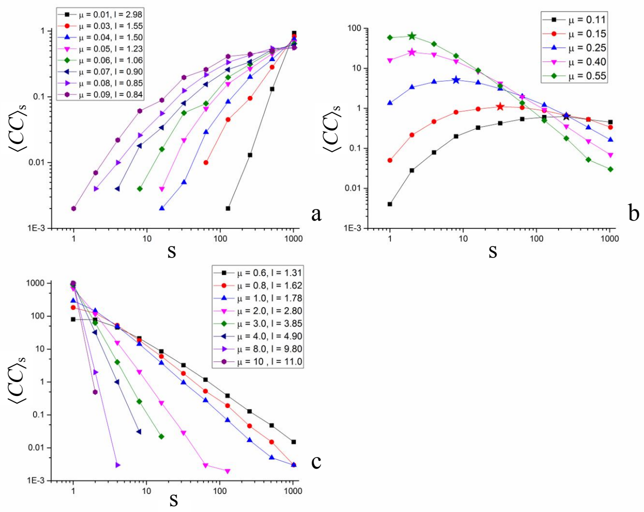

Fig. 4. *CC* distribution behavior in the log-log plot for the 3 ranges of the density parameter μ for a block-hierarchical network with the parameters $b = 2$, $\Gamma = 10$: a) exponential dependency of the growth $\langle CC\rangle_s \sim s^l$; b) emergence of the peak marked by asterisk; c) exponential dependency of the fall $\langle CC\rangle_s \sim s^{-l}$. The results are obtained using the *xRandNet* system on the ensemble of 1000 networks.

The picture changes when the value $\mu = \mu_1$ is attained: a transitional phase begins. The graph shows a peak (Fig. 4, b), that is the appearance of a component smaller in size than N is more probable than the appearance of a gigantic component. With each step of the change $\mu > \mu_1$ one can see that the peak appears for an increasingly smaller component. In other words, these are the connections appearing at a lower level that become decisive for the distribution of the connecting components. For example, on the graph of Fig. 4, b one can see that in the case of $\mu = 0.11$ the decisive connections appear on the 8$^{\text{th}}$ level (the biggest component has the size $2^8 = 256$) for $\mu = 0.15$ they emerge on the 5$^{\text{th}}$ level, and in case of $\mu = 0.55$ the connections are on the 1$^{\text{st}}$ level (that is, only 2s appear with a higher probability).

After the value μ_2 one can already see the full transition to the behavior typical to the classical model: there is less probability of the appearance of the gigantic component (near-zero) and many small components. With the growth of $\mu > \mu_2$ the appearance of the components of big sizes stops (on graph Fig. 4, c it is clear, that the components of the size 512, 256, 128, etc. are gradually lost). Already at the value of $\mu = 10$ the network comes to a situation where even the appearance of 2s has a near-zero probability, that is the network completely falls into isolated nodes.

Conclusion. The described behavior of *CC* distribution for the block-hierarchical networks is expectedly different from the corresponding behavior both for classical random networks and for scale-free networks:

- the state of connectivity for them is achieved when the probability of the connection of 2 nodes is close to 1 (μ is near-zero);

- on its way to such a state, *CC* distribution changes its behavior near the 2 values μ_1 and μ_2, which correspond to the probability of connections which for the classical random network would be in the supercritical phase.

μ_1, μ_2, p_1 and p_2 values calculated by Eq. (5) for block-hierarchical networks for $b = 2$

Γ	N	μ_1	p_1	μ_2	p_2	$1/N$
10	1 024	0.1	≈ 0.5	0.6	≈ 0.03	$\approx 9.7 \cdot 10^{-4}$
17	131 072	0.05	≈ 0.5	0.6	≈ 0.002	$\approx 7.6 \cdot 10^{-6}$
20	1 048 576	0.046	≈ 0.5	0.6	≈ 0.0005	$\approx 9.5 \cdot 10^{-7}$

In the course of research on *CC* distribution using the automated system *xRandNet* the values μ_1 and μ_2 were deduced for block-hierarchical networks with a branching index of 2 and different maximal levels ($\Gamma = 10, 17, 20$), which can be seen in Table. As well as this, the probabilities of the connection of the 2 nodes of these networks are calculated by formula (5).

Received 07.12.2017

REFERENCES

1. **Albert R., Barabási A.L.** Statistical Mechanics of Complex Networks. // Rev. Mod. Phys., 2002, v. 74, p. 47–97.
2. **Avetisov V.A., Bikulov A.Kh., Vasilyev O.A.** et al. Some Physical Applications of Random Hierarchical Matrices. // JETP, 2009, v. 109, № 3, p. 485–504.
3. **Moretti P., Muñoz M.A.** Griffiths Phases and the Stretching of Criticality in Brain Networks. // Nature. Communications, 2013, № 4, p. 2521–2547.
4. **Kocharyan A.G.** Application *xRandNet* for Studying the Topological Properties of Random Networks. // Proceedings of the NAS RA and SEUA. Ser. Technical Sciences, 2017, v. 70, № 4, p. 519–529 (in Russian).
5. **Avetisyan S., Harutyunyan A., Aslanyan D.** et al. Algorithms for Computation of Statistical Properties of Regular Block-hierarchical Networks. // Proceedings of the 6th Annual Scientific Conference. RAU, 2012, p. 108–121 (in Russian).
6. **Avetisyan S., Samvelyan M., Karapetyan M.** Random Irregular Block-Hierarchical Networks: Algorithms for Computation of Main Properties. // Proceedings of the 9th Annual Scientific Conference. RAU, 2015, p. 48–60 (in Russian).
7. **Barabási A.L.** Network Science. Cambridge University Press, 2016, 456 p.