CSIT Conference 2019, Yerevan, Armenia, September 23-27

Recognition of Block-Hierarchical Structure of Random Networks Using Machine Learning Techniques

Ani, Kocharyan
YSU
Yerevan, Armenia
e-mail: ankocharyan@gmail.com

Khachatur, Khechoyan
YSU
Yerevan, Armenia
e-mail: khachatur1998@gmail.com

ABSTRACT
The active study of random block-hierarchical networks in various fields of science highlights the problem of recognition and retrieval of the block-hierarchical structure in real networks, which are represented in an adjacency matrix form. This paper presents the usage of machine learning methods for analyzing structures of adjacency matrices of random networks for that purpose.

Keywords
Random networks, block-hierarchical networks, adjacency matrix, machine learning.

1. INTRODUCTION
In various fields of science, the number of so-called complex systems is increasing. Among them are biological systems (in particular, biopolymers, DNA, RNA, proteins), technical systems, social, economic ones and others. In the 1960s P. Erdős and A. Rényi came up with a suggestion to use probabilistic methods in the modeling of complex systems, by representing them as networks, in which the elements of the system are nodes and their interactions are links. The network is called random, if any connection in it appears due to a certain probabilistic rule. Henceforth we will examine only non-directed networks, which do not have self-references and multiple bonds. Erdős and Rényi suggested a model of random network with two parameters: the number of nodes N and probability p, which defines the existence of connection between any 2 nodes of the network. Apart from the studies of classical random *Erdős-Rényi* networks, the investigation of other classes of networks has become interesting – “small-world” and “scale-free” networks with their own special characteristics [1][2].

In the last decade a quite different class of random networks was introduced, called block-hierarchical. It turned out this class of networks can be applied to the modeling of different biological structures ranging from proteins to networks of neurons [3][4]. Different subclasses of random block-hierarchical networks were defined – *regular*, *irregular*, *HMN1*, *HMN2* and others [4-6].

In the paper the block-hierarchical structure recognition problem is formulated, regular and irregular block-hierarchical networks are defined, the special structure of their adjacency matrices are described and the use of machine learning techniques for recognition problem is represented. In addition, there are some analysis of suggested solutions and results of their testing.

2. BLOCK-HIERARCHICAL STRUCTURE RECOGNITION PROBLEM
Adjacency matrices of block-hierarchical random networks have a special structure due to the way of their construction. That structure highlights the fact that nodes of these networks are grouped in blocks (clusters, subnets), which form a hierarchy. It is an interesting problem to identify the block-hierarchical structure in adjacency matrices of real complex systems, which are modeled as random networks, of course, if they have it. The mentioned problem can be split into two parts:

- *Recognition Problem*. Identify if the given adjacency matrix has a block-hierarchical structure, either regular or irregular.
- *Clearness Problem*. If not, then measure how much it differs from a possible block-hierarchical structure. Also, find so-called “bad” links – ones that do not correspond to the mentioned structure.

If we find a way to recognize the block-hierarchical structure in a real network, then we can apply the results of studies of block-hierarchical random networks to them. Such studies include an analysis of their topological properties - node degree distribution, node-node distance distribution, the distribution of the clustering coefficients of the nodes, cycle length distribution, length distribution of the connected subnetworks, network diameter, etc. The automated system Extended Random Network (*xRandNet*) is developed based on the needs of random networks research and is aimed at efficient study, especially block-hierarchical ones [7][8]. Now there are published results of several studies of Regular Block-Hierarchical and Irregular Block-Hierarchical models done using xRandNet system [9][10].

3. RANDOM BLOCK-HIERARCHICAL NETWORKS
3.1. Regular block-hierarchical networks
Let us start from the definition of *random regular or regularly branching block-hierarchical networks* [3][5]. Let b and Γ be natural numbers, $b > 1$. For the given b and Γ a class of regularly branching block-hierarchical networks $\mathfrak{R}$ is defined as follows. The number of nodes in the network $G_{b,\Gamma} \in \mathfrak{R}_{b,\Gamma}$ is b^{Γ}. The network is constructed by levels. On every new level $\gamma, 0 \leq \gamma \leq \Gamma$ new clusters (subnets) are formed by connecting some clusters formed on the previous level. Probability of connection of clusters is . Connection of clusters, of course, causes addition of new links in the network $G_{b,\Gamma}$. In this paper “all to all” connection type is discussed, which means that connection of clusters formed on the previous level causes addition of links between all nodes from both clusters. Let be the adjacency matrix of $G_{b,\Gamma}$ network. One can see that in matrix $A_{b,\Gamma}$ there are $b(b-1)/2$ matrix blocks with 1 or 0 values. In Figure 1 regular block-hierarchical network and its adjacency matrix $A_{3,2}$ are shown.

3.2. Irregular block-hierarchic networks

Now *random irregular block-hierarchic networks* can be described as a generalized case of regular ones. For an irregular case the construction of a network differs from the regular case by

forming on each level not exactly b clusters, but a number of 1 to b, which is chosen in a probabilistic way [6].

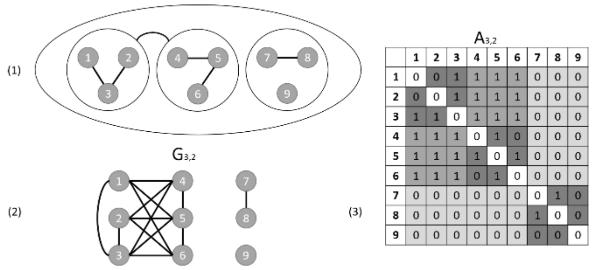

Fig. 1. Regular block-hierarchical network with its (1) clustered view, (2) nodes and links view, (3) adjacency matrix A view.

It is obvious, that for irregular block-hierarchical networks, the parameters b and Γ do not exactly define the number of nodes and branching pattern. So, in this case network is defined by giving count of nodes N, one of b and Γ parameters and also branching information. In Figure 2, the irregular block-hierarchical network and its adjacency matrix are shown.

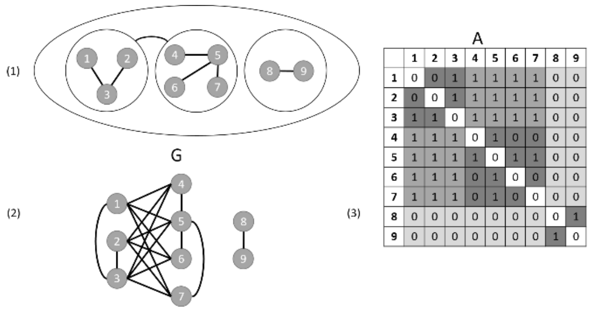

Fig. 2. Irregular block-hierarchical network G with its (1) clustered view, (2) nodes and links view, (3) adjacency matrix A view.

4. RECOGNITION OF BLOCK-HIERARCHICAL STRUCTURE

Since one can see that the network is block-hierarchical looking at its adjacency matrix, this paper treats the latter as an image, which allows the use of known image handling techniques.

4.1. Recognition Problem

Recognition Problem in machine learning terminology can be formulated as an image classification problem with 3 classes:

- “regular” class (for regular block-hierarchical network),
- “irregular” class (for irregular block-hierarchical network),
- “other”.

For the mentioned problem Convolutional Neural Networks (CNNs) are used [11]. In Figure 3 the architecture of the model is shown.

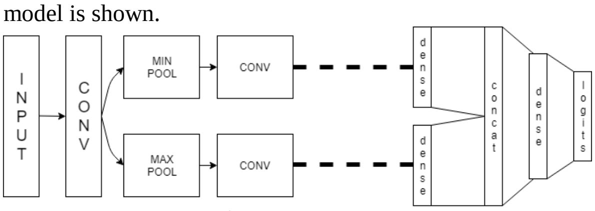

Fig. 3. The architecture of the model.

Unlike the traditional CNN, this architecture uses 2 branches: one for only min poolings and the other for max poolings. The intuition behind it is to avoid losing an amount of information using only one branch with min or max pooling.

Described model’s training quality depends on having well-formed training data for each class – “regular”, “irregular” and “other”. The automated system *xRandNet* makes it possible to effectively generate regular and irregular block-hierarchical networks with different b, Γ and μ parameters and save them to files as adjacency matrices [7][8][10]. It is obvious, that the latter matrices will cover training data for the first two classes – “regular” and “irregular”.

Training data for the “other” class is formed in several ways:

(1) by randomly adding “bad” connection to the generated block-hierarchical networks,

(2) by removing some connections from the generated block-hierarchical networks,

(3) by generating random networks of classical models - *Erdős-Rényi*, *Watts-Strogatz*, *Barabási-Albert*, etc., using the automated system *xRandNet*.

4.2. Clearness Problem

Clearness Problem can be solved with unsupervised learning technique called denoising autoencoder [12]. As it is known, autoencoder consists of encoder and decoder parts, where the encoder is responsible for reducing the dimensionality of input data by encoding it to a lower dimension, and the decoder - for decoding the output of the encoder to the same dimension as the input data. In formal:

$$\text{in}\|X - (\psi \circ \phi)X\|$$

where function conforms to the encoder part, and – to the decoder part. Therefore, the goal is to find such ϕ and that the output of autoencoder is very close to the input.

In denoising autoencoders it is necessary to add “noise” to the input data and choose such ϕ and ψ that the difference of the “clear” data and the output of the model on “noisy” data is minimum.

For the *Clearness Problem* the “bad” connections are treated as “noise”, which lets us train a denoising autoencoder on the (1) and (2) adjacency matrices mentioned above, which are generated to meet “other” class in *Recognition Problem*.

5. ANALYSYS OF THE SOLUTION

Some testing results of the described techniques for block-hierarchical networks are represented in this section. Experiments were performed on a machine with an Intel Core i7 4500U processor running at 2.40 GHz, 8 GB of RAM.

One can see network parameters, training and testing data information and appeared accuracy on the test data in Table 1. Let us mention, that:

- Training data for the *Recognition Problem* contains an equal number of data points for each class - regular and irregular.
- Training data for the *Clearness Problem* contains matrices with “noise” for each class - regular and irregular.

Table 1. Network parameters are b for "regular class" and N for "irregular class", training and testing data points and appeared accuracy.

	Training data	Testing data	Accuracy (%)
Recognition Problem	84 000	36 000	93.6
Clearness Problem	28 000	12 000	96.1

Looking at Figure 4, accuracy growth is obvious during the training process for both – *Recognition Problem* and *Clearness Problem.*

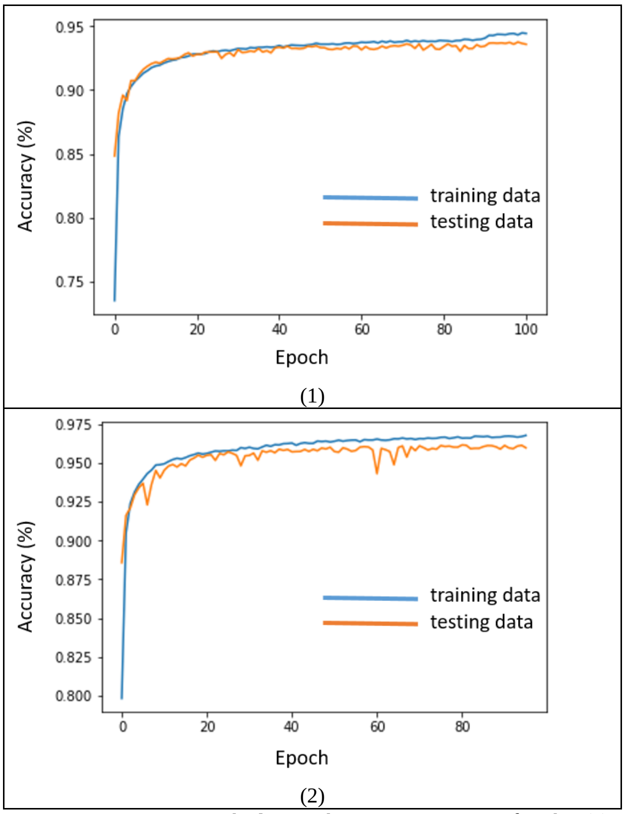

Fig. 4. Accuracy growth during the training process for the (1) Recognition Problem and (2) Clearness Problem.

Two examples of networks with and without noise, model output and model input-output differences are shown in Figure 5.

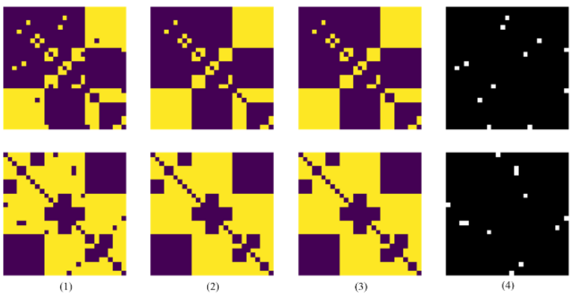

Fig. 5. Two examples of networks from "regular" class (1) with added "noise" (input matrices), (2) without the "noise", (3) output of the model for input matrices, (4) differences between input and output.

6. FUTURE GOALS

In some cases, a real network's adjacency matrix at first sight does not have a block-hierarchical structure, therefore, the suggested method will classify the latter matrix as "other". But simple re-enumeration of nodes can form an isomorphic matrix, in which the block-hierarchical structure will appear. Problem stands to find such an algorithm.

So, in addition to *Recognition Problem* and *Clearness Problem*, *Isomorphism Problem* can be formulated: identify if the given adjacency matrix or **an isomorphic one** has a block-hierarchical structure, either regular or irregular.

REFERENCES

[1] R. Albert, A. Barabási "Statistical Mechanics of Complex Networks", *Rev. Mod. Phys., v. 74, 2*, pp. 47–97, 2002.

[2] A. Barabási "Network Science", *Cambridge University Press*, 456 p, 2016.

[3] V. Avetisov, A. Bikulov et al. "Some Physical Applications of Random Hierarchical Matrices", *JETP, v. 109, № 3*, pp. 485–504, 2009.

[4] P. Moretti, M. Muñoz "Griffiths Phases and the Stretching of Criticality in Brain Networks", *Nature, Communications, № 4,* pp. 2521–2547, 2013.

[5] S. Avetisyan, A. Harutyunyan et al. "Algorithms for Computation of Statistical Properties of Regular Block-hierarchical Networks", *Proceedings of the 6th Annual Scientific Conference, RAU*, pp. 108–121, 2012 (in Russian).

[6] S. Avetisyan, M. Samvelyan et al. "Random Irregular Block-Hierarchical Networks: Algorithms for Computation of Main Properties", *Proceedings of the 9th Annual Scientific Conference, RAU*, pp. 48–60, 2015 (in Russian).

[7] S. Avetisyan, A. Kocharyan "The System of Generation of Random Networks and Computation of Their Topological Properties", *Proceedings of the Conference CSIT-2013, Publishing House of NAS of RA, Yerevan*, pp. 381-384, 2013.

[8] A. Kocharyan "Application xRandNet for Studying the Topological Properties of Random Networks", *Proceedings of the NAS RA and SEUA. Ser. Technical Sciences, v. 70, № 4*, pp. 519–529, 2017 (in Russian).

[9] A. Kocharyan "On Connected Component Distribution of Random Block-Hierarchical Networks Using The Automated System xRandNet", *Proceedings of the YSU: Physical and Mathematical Sciences, 52 (1)*, pp. 36-42, 2018.

[10] A. Kocharyan "Development of the system of automated research on topological properties of block-hierarchical networks", *doctoral dissertation, (http://etd.asj-oa.am/id/eprint/7660)*, 2018.

[11] Y. LeCun, P. Haffner, L. Bottou et al., "Object Recognition with Gradient-Based Learning", *Shape, Contour and Grouping in Computer Vision. Lecture Notes in Computer Science, vol 1681*, pp. 319-345, 1999.

[12] P. Vincent, H. Larochelle, Y. Bengio et al. "Extracting and composing robust features with denoising autoencoders", *Proceedings of the 25th International Conference on Machine Learning*, pp. 1096-1103, 2008.