Model of complex networks with rejection

Igor Zarvanskii, Andrei Snarskii

A modification of the preferential attachment rule, the attachment with rejection, is proposed and applied to network models constructed according to the Barabási-Albert algorithm and to (u, v)-flowers. The results of numerical simulation for the considered models are reported, and various parameters of simulated networks are analyzed. The parameters of obtained networks are shown to behave similarly to the phase transitions of the second kind. The threshold value of the rejection parameter, at which the phase transition takes place, has been calculated.

1. INTRODUCTION

A considerable number of real complex networks are scale-free, i.e. the degrees of network nodes are distributed according to the power law. Such networks include WWW networks, metabolic network, food webs, social networks, and many others [1].

Nowadays, the properties of such scale-free networks have been studied in detail, and the network parameters (the average node degree, the shortest average path length, the clustering coefficient, and so forth) have been determined [2]. It should be noticed that complex networks constructed according to work [3] are an idealization of real networks, and the parameters of the latter sometimes can considerably differ from perfect ones [2]. Nevertheless, the power-law dependence for node degrees in real complex networks is rather widespread, especially for the networks that are formed by processes evolving in time [4, 5].

One of such processes is the redistribution of "wealth" (as such, one may take money, investments, real estate, and so on) among people. Its study began long before the concept of complex network appeared. V. Pareto formulated the power-law distribution of wealth (the so-called Pareto law) [6]: the fraction ν of population is the power-law function of the fraction μ of wealth owned by them, $\nu \sim \mu^{\gamma}$. At $\gamma = 0.86$, we obtain that 20% of population own 80% of wealth, so that the Pareto law is often called the 80/20.rule.

In work [7], an algorithm for the formation of a complex network with the power law of node degree distribution (the so-called Barabási-Albert algorithm) was found. It is based on two essentially important points:

(i) the network is growing; starting from a definite node number m_0, at every timestep there appear a certain number of new nodes with n edges;
(ii) the probability for the edges of a new node to be attached to the existing ones is proportional to the node degree.
Briefly speaking, the Barabási-Albert model is a growing network with the preferential attachment.

Later, a lot of modifications of the Barabási-Albert algorithm were proposed. About 20 of them are indicated in work [3]. All those modifications give rise to scale-free networks with various indices of node distribution over the node degree. At first sight, the growing network with various types of preferential attachment inevitably has to transform into a scale-free network.

In this work, we demonstrate that a slight and, at first sight, insignificant modification of the law of preferential attachment is possible, at which the power-law distribution given by the Barabási-Albert model becomes violated. In this case, the distribution function has a discontinuity, which means that there are no network nodes which degrees fall within a certain interval of values. Detailed researches showed that the introduced parameter r, which describes the modification of the preferential attachment law, has a threshold value r_c, so that the network remains scale-free at $r < r_c$, whereas a discontinuity in the node degree distribution function emerges at $r \geq r_c$. The magnitude of discontinuity is a power-law function of the distance between the parameter r and its threshold value, which gives a ground to talk about an analogy with the order parameter behavior at the phase transition of the second kind.

We also verified the modification of the attachment law using deterministic hierarchical scale-free networks, the so-called (u, v)-flowers [8]. In this case, the violation of the power-law distribution function was observed, the behavior of which was similar to that of the order parameter at the phase transitions of the second kind.

2. BARABÁSI-ALBERT MODEL WITH REJECTION

2.1. Barabási-Albert algorithm

Consider a growing network. In the standard variant of Barabási-Albert model [7], there exist m_0 nodes at the first timestep, which are connected with each other. At every next step, there arises m new nodes, each with q edges. The quantity

$$p_i = \frac{k_i}{\sum\limits_j k_j} \tag{1}$$

characterizes the probability of attachment (the creation of an edge between the nodes) of a new node to the already existing i-th node; it is proportional to the i-th node degree k_i (the number of its edges), and summation is carried out over all "old" nodes.

When the number of timesteps is large, this algorithm gives rise to a power-law distribution function of node degrees $P(k)$:

$$P(k) \sim k^{-\gamma}, \tag{2}$$

with the exponent $\gamma = 3$ [7]. In work [3], a good many other modifications of the preferential attachment rule are quoted, which bring about various values of exponent γ. However, the power-law dependence (2) survives.

2.2. Modification of Barabási-Albert algorithm

In this work, we propose to generalize the model based on the preferential attachment rule introduced by Barabási and Albert. The new criterion will be referred to as the rule of preferential attachment with rejection. In the framework of this model, a new preferential attachment rejection parameter r, which acquires values in the interval $[0, 1]$, is introduced. In particular, when node i with the degree k_i is chosen to form an edge with the new node, a new attachment occurs with probability (1), i.e. in the case if the condition

$$k_i \geq r\langle k \rangle \tag{3}$$

is satisfied, where $\langle k \rangle = \frac{1}{N}\sum_j k_j$ is the average value of node degrees in the network at the moment immediately before the probable attachment. In other words, only attachments

to "rich" nodes, which degrees are not less than $r\langle k\rangle$, are allowed. Extra condition (3) deactivates some nodes in the course of network growth, i.e. they cannot form edges with the new node at the given timestep. It should be noticed that, if a certain node does not satisfy condition (3) at the given moment, it does not mean that new nodes cannot be attached to it later, because the $\langle k\rangle$-value changes in time.

2.3. Distribution function for node degrees

If the rejection parameter $r = 0$, the proposed model transforms into the standard Barabási-Albert model, because k_i is always larger than zero. Unexpectedly, the rejection parameter was found to have a threshold behavior. If the rejection parameter is less than a certain threshold value r_c, i.e. at $r < r_c$, the distribution function of node degrees $P(k)$ remains to be power-law and, therefore, the network itself remains scale-free. At the rejection parameter values larger than the threshold one ($r \geq r_c$), the network changes its structure; namely, nodes with "intermediate" node numbers disappear, which can be seen in Fig. 2.

Let us determine the threshold value of rejection parameter. For this purpose, we calculated a network with the initial node number $m_0 = 20$. At every step, there emerged one node with $q = 3$ edges. After 80 steps, we obtained a network with $N = 100$ nodes. Repeating this procedure many times for various r's , we found an r-value at which the distribution of p_i for the specific network ceased to be power-law, i.e. when there appeared a break in the network. The results of calculations demonstrated that $r_c(100) = 0.62$ at $N = 100$. As N increased from 100 to 2000, the value of r_c decreased and saturated at $r_c = 0.51 \pm 0.04$. This value of r_c was considered to be the threshold value of rejection parameter for large (infinite) networks. The threshold value of rejection parameter did not change if the initial parameters m_0 and q are varied. In further calculations, the indicated r_c-value was used.

Let us introduce a new characteristic of the network, the gap magnitude η (see Fig. 2), i.e. the smallest magnitude of the difference between the node degree values to the left and to the right of the discontinuity in the figure. The gap interval marks the node degree values (the ordinate axis) that are absent in the network.

The behavior of the parameter η is similar to that of the order parameter in the theory of phase transitions of the second kind [13]. It is known that, when the temperature approaches

the critical value T_c, the order parameter η—e.g., it may be the magnetization—diminishes following the power law $\eta \sim (T - T_c)^\beta$, where β is the critical index.

For the numerical experiment, the following parameters were chosen: the final node number $N = 5000$, the initial node number $m_0 = 20$, and the edge number for every new node $m = 3$. In Fig. 3, the obtained dependence $\eta = A\,(r - r_c)^\beta$, where $\beta = 1.15 \pm 0.05$, is shown.

2.4. Clustering and assortativity coefficients

The appearance of gap η in the node degree distribution function $P(k)$ testifies to a substantial variation in the network structure, which has to manifest itself in the network parameters. Let us consider the behavior of the clustering, C, and assortativity, A, coefficients as functions of the rejection coefficient r in the vicinity of $r = r_c$. For a network with $N = 5000$ nodes, the both do not depend on r at $r < r_c$; they equal $C_0 \approx 0.01$ and $A_0 \approx -0.096$, and coincide with the values calculated in works [3, 14]. For r growing from r_c to $r_c + 0.01$ with an increment of 0.001, the clustering coefficient increases from 0.04 to 0.14, and the assortativity one decreases from -0.3 to -0.6 (see Fig. 4). At $r \geq r_c$, the dependences of the coefficients C and A on r appear. The both turn out power-law; namely, $C \sim (r - r_c)^\alpha$, where $\alpha = 0.46 \pm 0.04$, and $A \sim (r - r_c)^\gamma$, where $\gamma = 0.26 \pm 0.04$ (see Fig. 5).

2.5. Adjacency matrix for the network with rejection

Consider the adjacency matrix A_{ij} for a network with rejection. For convenience, the nodes in the adjacency matrix are enumerated in the order when the number of node edges decreases. This means that $k_i = \sum_j A_{ij}$ decreases as i grows.

The change in the network structure at $r \geq r_c$ affects the form of the adjacency matrix as well. For a network with $N = 5000$, two adjacency matrices were constructed: for $r < r_c$ (Fig. 6) and $r > r_c$ (Fig. 7). The both matrices were ranked, i.e. the network nodes were enumerated in the decreasing order of the edge number k_i. From Fig. 7, one can see that, in the case $r > r_c$, a considerable square region filled with zeros, which correspond to unconnected node pairs, appears in the bottom right corner of the adjacency matrix. This region depends proportionally on the r-value.

Hence, such network parameters as the clustering and assortativity coefficients behave similarly to the order parameter η.

3. HIERARCHICAL (u, v)-FLOWER NETWORKS WITH REJECTION

Besides random scale-free networks constructed following the Barabási-Albert algorithm and its generalizations, another class of simple deterministic networks, which are also scale-free, is known [9]. These are the so-called (u, v)-flowers (Fig. 8).

In this Section, similarly to what was done for the Barabási-Albert model, we generalize the (u, v)-flower model by introducing the rejection parameter r and the stochasticity into the rule of network growth. In this case, the behavior of network parameters also turns out similar to that of the order parameter at phase transitions of the second kind.

3.1. Algorithm for generating (u, v)-flowers

Growing deterministic scale-free networks, which are called (u, v)-flower and (u, v)-trees, were proposed and studied in works [8–10]. Generally speaking, the enumeration of nodes in the (u, v)-flower can be arbitrary. However, in the example concerned, the nodes can be enumerated in such a way (Fig. 8) that the corresponding adjacency matrix A_{ij} acquires the simplest view. A simple view means in this case such a structure of A_{ij} matrix that the largest number of the largest square $N \times N$ regions in it remain empty.

Figure 9 illustrates an $N \times N$ adjacency matrix, where black cells correspond to matrix members $A_{ij} = 1$. At a first step, the adjacency matrix $\hat{A}$ consists of 3×3 elements (from A_{11} to A_{33}, the step $t = 0$ in Fig. 8). At a second step, new elements are added to it, and the matrix consists of 6×6 elements (the step $t = 1$ in Fig. 8). At a third step, new elements are added, and the matrix consists of 15×15 elements (the step $t = 2$ in Fig. 8). One can see from Fig. 9 that the bottom right square of the first step is free of edges and consists of one element. A bottom right square consisting of 3×3 elements is added at the second step, and a 9×9 square at the third one. In Fig. 9, the $A_{ij} = 0$ members of adjacency matrix are white-colored.

While comparing our procedure of adjacency matrix generation with that used in work [8], the node enumeration order selected in this work allowed us to obtain additional regions

K_1 and K_2, in which $A_{ij} = 0$ as well.

At every step t, there are N_t nodes and L_t edges [10] described by the formulas

$$N_t = (u+v) \cdot N_{t-1} - (u+v), \quad L_t = (u+v)^t. \tag{4}$$

At every step t, there appear $N_t - N_{t-1}$ nodes and $L_t - L_{t-1}$ edges. For instance (see Fig. 8), 3 nodes and 6 edges appear at the step $t = 1$. According to work [10], let us introduce the notation $w = u + v$ and expand recurrent formulas (4). We obtain

$$N_t = \frac{w-2}{w-1} \cdot w^t + \frac{w}{w-1}, \quad L_t = w^t. \tag{5}$$

At every step t, there are Ω_t cells which can be filled (see Fig. 9). Their number equals

$$\Omega_t = (N_t - N_{t-1}) \cdot N_t - N_{t-2}^2 = \frac{w^3 - w^2 - 1}{w^4} N_t^2 + \frac{w^3 - 2w^2 - 2w - 2}{w^3} N_t - \frac{2w^2 + 2w + 1}{w^2} =$$
$$= \frac{w^{2t+4} - 5w^{2t+3} + 8w^{2t+2} - 5w^{2t+1} + 4w^{2t} + w^{t+5} - 3w^{t+4} + 4w^{t+2} - w^5}{w^3 \cdot (w-1)^2}. \tag{6}$$

Hence, the probability for a matrix cell to be filled equals

$$W_t = \frac{L_t - L_{t-1}}{\Omega_t} = \frac{w^{t+5} - 3w^{t+4} + 3w^{t+3} - w^{t+2}}{w^{2t+4} - 5w^{2t+3} + 8w^{2t+2} - 5w^{2t+1} + 4w^{2t} + w^{t+5} - 3w^{t+4} + 4w^{t+2} - w^5}. \tag{7}$$

This quantity becomes smaller as the step number grows, so that the adjacency matrix gets more and more sparse. In work [8], it was shown that (u, v)-flowers are scale-free networks. For the (1,2)-flower exhibited in Fig. 8, the node degree distribution function looks like $P(k) \sim k^{-(1+\ln 3/\ln 2)}$. In the general case, for (u, v)-flowers we obtain [10]

$$P(k) \sim k^{-\alpha}, \quad \alpha = 1 + \frac{\ln(u+v)}{\ln 2}. \tag{8}$$

3.2. Modification of the (u, v)-flower generation algorithm

Consider the case when not all edges are realized when an (u, v)-flower is generated. Let us also suppose that the lesser are the degrees of nodes, the higher is the probability that the edge connecting them is absent.

In the case of Barabási-Albert model, the rule of attachment with rejection resulted in the appearance of empty region (the region free of edges) in the bottom right corner of the

adjacency matrix (see Fig. 7). In the case of (u, v)-flower model, such a region already exists. Therefore, empty regions E_1, E_2, etc. will be created in the bottom right corners of already filled regions Ω_1, Ω_2, and so on.

The network is generated by filling the regions $\Omega_t - E_t$ with the edge number L_t in the adjacency matrix. For the random creation of edges to correspond to the (u, v)-flower, the law of node degree distribution $p \sim k^{-(1+\ln(u+v)/\ln 2)}$ [9] has to be preserved. For this purpose, a partitioning of $\Omega_t - E_t$ regions into 4 equal parts is used. The probability for the t-th part $(t = 1 \div 4)$ to be filled equals

$$\psi_t = \left(\frac{1}{t}\right)^{-(1+\ln(u+v)/\ln 2)} - \psi_{t-1}.$$

Below, we deal with $(1, 2)$-flowers, and the filling probabilities for the $\Omega_t - E_t$ regions are 0.37, 0.3, 0.22, and 0.11. Hence, the rejection manifests itself in the adjacency matrix as E_t regions, i.e. empty bottom right corners of Ω regions (Fig. 10).

At simulation, empty E_t regions proportional to Ω_t ones were used:

$$E_t = [r \cdot N_{t-1}] \cdot [r \cdot (N_t - N_{t-1})], \tag{9}$$

where $r\,N_{t-1}$ is the number of rows, and $r\,(N_t - N_{t-1})$ the number of columns in the E_t region (Fig. 10). The filling probability for the matrix cell equals

$$W_t = \psi_t \cdot \frac{L_t - L_{t-1}}{\Omega_t - E_t}, \tag{10}$$

where ψ_t is the filling probability for the $\Omega_t - E_t$ region.

3.3. Node degree distribution function

If the rejection parameter $r = 0$, the proposed model transforms into the standard (u, v)-flower one. Similarly to the Barabási-Albert model, the (u, v)-flower also contains the threshold value of rejection parameter, r_c. If the rejection parameter is less than the threshold value, $r < r_c$, the node degree distribution function $P(k)$ remains to be power-law, and the network itself remains to be scale-free. But at rejection parameter values larger than the threshold, $r \geq r_c$, the network changes its structure.

Let us determine the threshold value of rejection parameter. For this purpose, we calculated the threshold value for all generations of $(1, 2)$-flower from the first to the 14-th one.

Analogously to the Barabási-Albert model with rejection, the threshold value also saturated. Now, it occurred at the 8-th generation, and we obtained $r_c = 0.75 \pm 0.04$.

In order to determine the gap magnitude (the parameter η), we have to find the bend points (see Fig. 11). The sough points correspond to the minima of the radius of curvature $k = (1 + y'^2)^{3/2} / |y''|$ [11], where $y(x) = a + bx + \dot{c} \arctan(x) + d \arctan(\alpha x + \beta)$ is the function approximating the dependence of the node degree y on the ordinal node number x [12] (see Fig. 11).

The parameter η behaves similarly to the order parameter in the theory of phase transitions of the second kind [13]. When approaching the critical rejection value, it decreases following the power law, $\eta \sim (r - r_c)^{\beta}$, where β is the critical index. $(1, 2)$-flower networks of the eighth generation containing $N = 3282$ nodes were simulated. In Fig. 12, the calculated dependence $\eta = A\,(r - r_c)^{\beta}$, where $\beta = 0.28 \pm 0.05$, is shown.

3.4. Clustering and assortativity coefficients

Again, let us consider the behavior of the clustering, C, and assortativity, A, coefficients as functions of the rejection parameter r. At $r \geq r_c$, for the eighth generation of the $(1, 2)$-flower network with $N = 3282$ nodes, the both do not depend on r and equal $C_0 \approx 0.02$ and $A_0 \approx -0.18$, which expectedly coincides with the results of calculations in works [9, 10]. At $r \geq r_c$, this dependence appears (as r grows, the clustering coefficient increases and the assortativity one decreases) and turns out to be power-law; namely, $C \sim (r - r_c)^{\alpha}$, where $\alpha = 0.11 \pm 0.04$, and $A \sim (r - r_c)^{\gamma}$, where $\gamma = 0.08 \pm 0.04$ (Fig. 13).

4. CONCLUSIONS

A modified rule of preferential attachment, namely, the attachment with rejection, was proposed for the classes of scale-free networks, the Barabási-Albert and (u, v)-flower models. The modification of the preferential attachment rule consists in the introduction of the rejection parameter, which rejects some nodes in the course of network growth, i.e. disables their ability to form new edges with new nodes at the present moment.

The results of numerical simulation showed that the networks belonging to the considered classes undergo substantial structural changes. The introduction of a new network param-

eter, the gap magnitude, and the calculation of already known characteristics, such as the clustering and assortativity coefficients and the average least distance between the nodes, allowed a conclusion to be drawn that the simulated classes of networks undergo the phase transition of the second kind. The corresponding threshold values of the rejection parameter, at which the phase transition takes place, were calculated. The introduced parameter "gap magnitude" was found to be proportional to the order parameter. The appearance of the gap η in the ranked distribution of network nodes at $r > r_c$ (Fig. 1) may be of interest for economic models analyzing the redistribution of wealth [15].

Let us consider the following model describing the distribution of income. Let every node represent an enterprise. Suppose that the amount of wealth for the enterprise is proportional to the number of its links with other enterprises, i.e. to the node degree. Every new node (an enterprise) makes an attachment (forms a contact) with other, already existing nodes. If the probability of this contact is proportional to the wealth magnitude (the node degree) of the enterprise with which the contact is established, the distribution of enterprises over the magnitudes of their wealth turns out the Pareto distribution [15, 16], which is observed in a good many real cases [16].

However, if the rejection parameter is introduced, a distribution with a discontinuity (Fig. 1) rather than the Pareto one is observed. In terms of the considered economic model (the node degree vs. the wealth of the enterprise and/or its people), this means that the so-called middle class disappears. In Fig. 1, one can see that there are almost no enterprises with the wealth in the range of η (these are nodes with the degrees approximately from 4 to 70). In other words, there are only very rich enterprises/people (nodes with large degrees) and poor ones (nodes with small degrees).

The authors express their gratitude to I.V.Bezsudnov and D.V. Lande for numerous useful discussions.

[1] 1

[2] 1

[3] 1

[4] 1

[5] 1

[6] 1

[7] 1

[8] 1

[9] 1

[10] 1

[11] 1

[12] 1

[13] 1

[14] 1

[15] 1

[16] 1

Fig. 1. (a) Ranked distribution of nodes in the log-log scale for a network with $N = 1000$ nodes at $r = 0.6$. The node ordinal numbers are reckoned along the abscissa axis, and the node degrees along the ordinate one. (b) The gap magnitude η as a function of the normalized rejection parameter $(r - r_c)/r_c$ in the interval $r_c < r < r_c + 0.01$ in the log-log scale.

Fig. 2. (a) Dependences of the clustering, C, and assortativity, A, coefficients on the rejection parameter r within the interval $0.4 < r < 0.7$. (b) Dependences of C and A on the normalized rejection parameter $(r - r_c)/r_c$ in the interval $r_c < r < r_c + 0.06$ in the log-log scale.

Fig. 3. Adjacency matrices for a network with $N = 5000$ nodes at $r = 0$ (a) and 0.6 (b). Black cells mark elements $A_{ij} \neq 0$.

Fig. 4. Illustration of the (1,2)-flower generation algorithm at the steps $t = 0$, 1, and 2. Nodes appearing at the current step are marked bold (red online)

Fig. 5. Adjacency matrices for the 3-rd step of (1,2)-flower generation: (a) adjacency matrix of the network with the selected node enumeration; (b) its schematic diagram, where K_1, K_2, and so on are empty regions

Fig. 6. Schematic diagram of adjacency matrix for the 3-rd step of (1,2)-flower generation.

Fig. 7. (a) Ranked distribution of network nodes in the log-log scale at $r = 0.8$. The node ordinal numbers are reckoned along the abscissa axis, and the node degrees along the ordinate one. (b) The gap magnitude η as a function of the normalized rejection parameter $(r - r_c)/r_c$ in the interval $r_c < r < r_c + 0.08$ in the log-log scale.

Fig. 8. Dependences of the clustering, C, and assortativity, A, coefficients on the normalized rejection parameter $(r - r_c)/r_c$ in the interval $r_c < r < r_c + 0.01$ in the log-log scale.

FIG. 1: fig:rankDistribution

FIG. 2: fig:rankDistribution-rank

FIG. 4: fig:baCharacteristic-raw

FIG. 5: fig:rankDistribution-log

FIG. 6: fig:baRankedMatrix-0

FIG. 7: fig:baRankedMatrix-06

FIG. 8: fig:flowerGraph

FIG. 9: fig:flowerMatrix

FIG. 10: fig:flowerMatrixExceptive

FIG. 11: fig:flowerRank-08

FIG. 12: fig:flowerRank-gap

FIG. 13: fig:flowerParam-log

FIG. 3: fig:rankDistribution-gap