On the diameter of power-law configuration graphs

M. M. Leri [1],

[1] Karelian Research Centre of the Russian Academy of Sciences,
Petrozavodsk, Russia, leri@krc.karelia.ru

We consider configuration graphs introduced by Bollobas [1] consisting of N nodes with vertex degrees $\xi_1, \ldots, \xi_N$ which are independent identically distributed random variables following the power-law distribution (1) (see e.g. Reittu and Norros [2]):

$$\mathbf{P}\{\xi = k\} = k^{-\tau} - (k+1)^{-\tau}, \qquad k = 1, 2, \ldots \tag{1}$$

where $\tau > 1$ is a parameter of the node degree distribution. Graph construction starts by forming incident stubs (see Reittu and Norros [2]) for each node the number of which is equal to the node degree. All the stubs are numbered in an arbitrary order. To form graph links all stubs are equiprobably joined pairwise. Such a construction requires the sum of node degrees to be even, so if otherwise one stub is added to an equiprobably chosen node to form a lacking link. Obviously, configuration graphs may have loops, cycles and multiple links.

Having considered power-law configuration graphs with the parameter $1 < \tau < 2$ Reittu and Norros [2] proved that in such graph asymptotically almost surely (a.a.s.) exists one unique connected component of the size proportional to N. This component was called "the giant" and it was shown that the sizes of other "smaller" components are infinitely small in comparison to the size of the giant one. Furthermore, Leri [3] showed that the closer the value of the parameter τ to 1 the higher is the probability that all N graph nodes are connected into one component. On the other hand, graphs with $2 \leq \tau \leq 3$ still may contain a component with the largest number of nodes in comparison to other "smaller" components. As for the number of components, it grows with the growth of N.

One of the important numerical characteristics of random graphs is its diameter (see e.g. Bollobas [4]). The graph diameter is the length of the "longest shortest path" between any two graph nodes. "Shortest path" is a distance between two graph nodes. Therefore, to find a graph diameter is enough to find distances between each pair of graph nodes, and maximum of all these distances will be a graph diameter. In configuration graphs loops are excluded, multiple links are considered as one and paths between nodes lying in different connected components do not count. To find the shortest paths, Dijkstra's algorithm [5] is used.

The study was carried out using simulation methods with subsequent statistical data processing. We considered graphs of $1000 \leq N \leq 10000$ sizes with a step 500. The values of the parameter of node degree distribution (1) varied in the interval $(1, 3]$ with a step 0.01 for $1.01 \le \tau < 1.1$ and 0.1 for $1.1 \le \tau \le 3$. To form statistical data for each pair (N, τ) there were generated 100 configuration graphs. A regression dependence of the diameter D on the graph size N and the parameter of node degree distribution τ was obtained to be the following:

$$D = -5.27 + (1.48\tau + 0.05\tau^5 - 0.02\tau^6)\ln N, \quad (R^2 = 0.93). \tag{2}$$

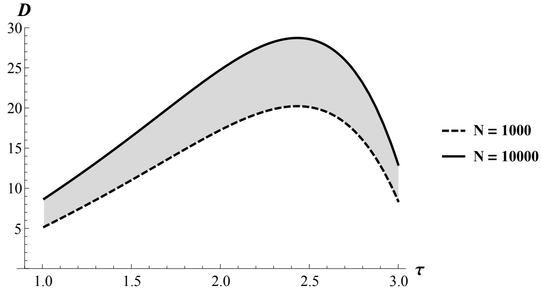

Figure 1: Regression dependence of D on N and τ.

For each pair (N, τ), the average values of the diameter D (out of 100 experiments) were calculated. Further, for each N, $D_{\max}$ and the corresponding values of τ at which this maximum D is reached, were found, as well as the values of $D_{\min}$ and the corresponding values of τ at which this minimum is attained. Regression dependencies of $D_{\max}$ on N and $D_{\min}$ on N are modeled by equations (3) and (4) with $R^2 = 0.99$ in both cases, respectively (see Figure 2).

$$D_{\max} = 6.97 \ln N - 30.59, \tag{3}$$

$$D_{\min} = 0.47 \ln N + 3.83. \tag{4}$$

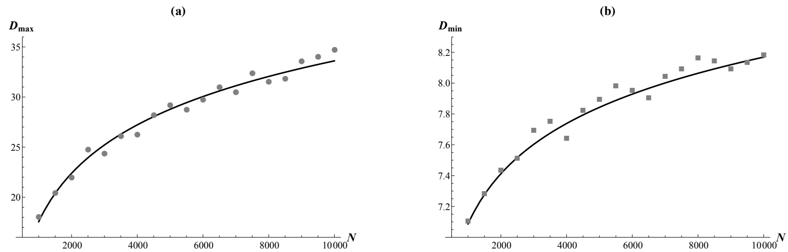

Figure 2: Dependences of $D_{\max}$ on N (a) and $D_{\min}$ on N (b). Solid line shows regression dependence, grey dots mark sample values $D_{\max}$ and $D_{\min}$, respectively.

Furthermore, there were considered the sizes η_1 of the "giant" (or the largest) graph component under condition of maximum or minimum of its diameter. Dependencies of experimental averages $\eta_1(D_{\max})$ and $\eta_1(D_{\max})$ on N were analyzed.

Thus, in power-law configuration graphs the graph diameter reaches its maximum within the values of the parameter $2.3 \leq \tau \leq 2.4$ with the increase of the graph size N. However, in small graphs ($N < 2000$) maximum diameter is observed for minor values of τ, though no less than 2.1. This diameter increase is apparently associated with the appearance of longer simple chains of nodes. In addition, regardless of the graph size, the closer the value of the parameter τ to 1, the smaller is the diameter of such graph. This may be due to the fact that the graphs in this case are more connected, i.e. the giant component includes more than 95% of the nodes, however the distances between the nodes in this component are small, whence it follows small size of the diameter. This also aligns with to the so-called idea of "six degrees of separation".

References

1. B. Bollobas, A probabilistic proof of an asymptotic formula of the number of labelled regular graphs, *European J. Combin.* **1**:4 (1980) 311–316.

2. H. Reittu, I. Norros, On the power-law random graph model of massive data networks, *Performance Evaluation* **55**:1-2 (2004) 3–23.

3. M. M. Leri, On conditions of configuration graphs' connectivity, *Trans. of KRC RAS: Mathematical Modeling and Information Technologies* **7** (2020) 71–75.

4. B. Bollobas, O. M. Riordan, The diameter of a scale-free random graph, *Combinatorica* **24**:1 (2004) 5–34.

5. E. W. Dijkstra, A note on two problems in connexion with graphs, *Numer. Math. / F. Brezzi Springer-Verlag* **1**:1 (1959) 269–271.