On various statistics of a random web-graph in the Bollobás–Riordan model*

E.A. Grechnikov†, L.A. Ostroumova‡, A.M. Raigorodskii§, A.A. Ryabchenko¶, E.A. Samosvat‖, D.A. Shabanov**

During the last decade, many models of the random web-graph have appeared (see [1], [2]). One of them was proposed in 2002 by B. Bollobás and O. Riordan (see [3]). In this model, a random graph process $(G_k^t)_{t=1}^{\infty}$, for any fixed $k \in \mathbb{N}$, is defined. Let us recall the definition. First, we take the graph $G_1^1 = (\{1\}, \{(1,1)\})$ on the vertex 1 with a single loop. Further, given a graph $G_1^{t-1} = (\{1,\dots,t-1\}, E_1^{t-1})$ with $t-1$ edges or loops ($t \geq 2$) we transform it into G_1^t by adding one new vertex t and one new random edge (t,i), where $i \in \{1,\dots,t\}$ and $P(i=t) = \frac{1}{2t-1}$, $P(i=j) = \frac{\deg_{G_1^{t-1}} j}{2t-1}$, $j < t$. Finally, we take G_1^{kt} and identify blocks of k vertices: the vertices $1,\dots,k$ form a new vertex v_1, the vertices $k+1,\dots,2k$ form a new vertex v_2, etc.

For the random graph G_k^t, various results have been proved showing that this graph has many properties of the World Wide Web. For example, it is subject to a power law distribution of its vertex degrees (see [4] – [6], [7]). It also has small diameter (see [4] – [6], [3]).

In our work, we continue studying important statistics of the random web-graph in the described model. Below we briefly list our results.

1. Let X_{d_1,d_2} be the number of edges in G_k^t whose vertices have degrees d_1 and d_2 respectively. We calculate the expectation EX_{d_1,d_2} of this random variable getting an asymptotic formula with the exact values of the two first terms and a remainder of the form $O_{k,d_1,d_2}\left(\frac{1}{t}\right)$. As a simple corollary we re-obtain an asymptotic expression for the expectation of the number X_d of vertices having degree d: $EX_d = \frac{(2kt+1)(k+1)}{d(d+1)(d+2)} + O(d/t)$. It is worth noting that here we do not have any restriction on d (cf. [7]): one should only assume that $d = o(t)$. Also we prove tight concentration of X_{d_1,d_2} around its expectation.

2. We study the expectation of the "second indegree" of a vertex in G_k^t. By *the second indegree* of an i we mean the number of edges going to the neighbours of i. We find sharp formulas for this expectation and we also prove some tight concentration results.

3. We find simple formulas for the expectation of the numbers of given subgraphs H in the random graph G_k^t.

*This work was supported by Yandex Internet Company as well as by the grants of RFBR N 09-01-00294 and of the President of the Russian Federation N MD-8390.2010.1, NSH-691.2008.1.

†Moscow State University, Mechanics and Mathematics Faculty, Department of Mathematical Statistics and Random Processes; Yandex research laboratories.

‡Moscow State University, Mechanics and Mathematics Faculty, Department of Mathematical Statistics and Random Processes.

§Moscow State University, Mechanics and Mathematics Faculty, Department of Mathematical Statistics and Random Processes; Moscow Institute of Physics and Technology, Faculty of Innovations and High Technology, Department of Data Analysis; Yandex research laboratories.

¶Moscow Institute of Physics and Technology, Faculty of Innovations and High Technology, Department of Data Analysis; Yandex research laboratories.

‖Moscow Institute of Physics and Technology, Faculty of Innovations and High Technology, Department of Data Analysis; Yandex research laboratories.

**Moscow State University, Mechanics and Mathematics Faculty, Department of Mathematical Statistics and Random Processes; Moscow Institute of Physics and Technology, Faculty of Innovations and High Technology, Department of Data Analysis; Yandex research laboratories.

4. We use some of the obtained results in order to statistically classify “dense structures” (communities, link spam, etc.) in the World Wide Web. Such classification is important to improve quality of search engine rankings.

References

[1] B. Bollobás, O. Riordan, *Mathematical results on scale-free random graphs*, Handbook of graphs and networks, 1 - 34, Wiley-VCH, Weinheim, 2003.

[2] R. Durrett, *Random graph dynamics*, Cambridge Univ. Press, Cambridge, 2007.

[3] B. Bollobás, O. Riordan, *The diameter of a scale-free random graph*, Combinatorica, 24 (2004), N1, 5 - 34.

[4] L.-A. Barabási, R. Albert, *Emergence of scaling in random networks*, Science, 286 (1999), 509 - 512.

[5] L.-A. Barabási, R. Albert, H. Jeong, *Scale-free characteristics of random networks: the topology of the world-wide web*, Physica, A281 (2000), 69 - 77.

[6] R. Albert, H. Jeong, L.-A. Barabási, *Diameter of the world-wide web*, Nature, 401 (1999), 130 - 131.

[7] B. Bollobás, O. Riordan, J. Spencer, G. Tusnády, *The degree sequence of a scale-free random graph process*, Random Structures Algorithms, 18 (2001), N3, 279 - 290.