A summary of
The diameter of random sparse graphs
by W. Aiello, Fan Chung and Lincoln Lu

Massive graphs that arise in the studies of the Internet share many similar aspects with random graphs, although there are significant differences (e.g., there can be vertices with large degrees in a sparse massive graph). Nevertheless, many of the methods and ideas [1, 3] that are used in modeling and analyzing massive graphs have been frequently traced to the seminal papers of Erdős and Rényi [9] in 1959. The paper [9] contains a masterful study of the sizes of the connected components as the graph evolves.

One topic of considerable interest is to examine the "shape" of the connected components in a random graph, in particular, when the graph is sparse. A natural question is to determine the diameter of a random sparse graph. Most known results on this topic involves graphs that are relatively dense. Namely, let $G(n,p)$ denote a random graph on n vertices in which a pair of vertices appears as an edge of $G(n,p)$ with probability p. Most work on the diameter of the random graph $G(n,p)$ covers the case that $pn/\log n \to \infty$ (see [5, 6, 7, 11]). The range of interest is for the average degree pn to range from 0 to $c \log n$ which includes the emergence of the unique giant component. Since there is a phase transition in connectivity at $p = \log n/n$, the problem of determining the diameter of $G(n,p)$ and its concentration seems to be difficult for certain ranges of p. In [8] we investigate this problem by identifying the ranges that results can be obtained as well as the ranges that the problems remain open.

For $\frac{np}{\log n} = c > 8$, we slightly improve Bollobás' result [6] by showing that the diameter of $G(n,p)$ is almost surely concentrated on at most two values around $\log n/\log np$. For $\frac{np}{\log n} = c > 2$, the diameter of $G(n,p)$ is almost surely concentrated on at most three values. For the range $2 \geq \frac{np}{\log n} = c > 1$, the diameter of $G(n,p)$ is almost surely concentrated on at most four values.

For the range $np < \log n$, the random graph $G(n,p)$ is almost surely disconnected. It is shown [8] that almost surely the diameter of $G(n,p)$ is $(1+o(1))\frac{\log n}{\log(np)}$ if $np \to \infty$. Moreover, if $\frac{np}{\log n} = c > c_0$ for any (small) constant c and c_0, then the diameter of $G(n,p)$ is almost surely concentrated on finitely many values, namely, no more than $2\lfloor \frac{1}{c_0} \rfloor + 4$ values.

In the range of $\frac{1}{n} < p < \frac{\log n}{n}$, the random graph $G(n,p)$ almost surely has a unique giant component. A tight upper bound of the sizes of its small components was obtained if p satisfies $np \geq c > 1$. Also, the diameter of

$G(n,p)$ almost surely equals the diameter of its giant component for the range $np > 3.513$. This problem was previously considered by Łuczak [10].

In [10] Łuczak examined the diameter of the random graph for the case of $np < 1$. Łuczak determined the limit distribution of the diameter of the random graph if $(1-np)n^{1/3} \to \infty$. The diameter of $G(n,p)$ almost surely either is equal to the diameter of its tree components or differs by 1.

Numerous questions on the diameter of sparse random graphs remain unanswered, several of which we mention here:

Problem 1: Is the diameter of $G(n,p)$ equal to the diameter of its giant component ?

By the result in [8], this question only concerns the range $1 < p \leq 3.5128$. There are numerous questions concerning the diameter in the evolution of the random graph. The classical paper of Erdős and Rényi [9] stated that all connected components are trees or unicyclic in this range. What is the the distribution of the diameters of all connected components? Is there any "jump" or "double jumps" as the connectivity [9] in the evolution of the random graphs during this range for p?

In [8] we proved that almost surely the diameter of $G(n,p)$ is $(1+o(1))\frac{\log n}{\log np}$ if $np \to \infty$. When $np = c$ for some constant $c > 1$, we can only show that the diameter is within a constant factor of is $\frac{\log n}{\log np}$. Can this be further improved?

Problem 2: Prove or disprove

$$diam(G(n, \frac{c}{n})) = (1+o(1))\frac{\log n}{\log c}$$

for constant $c > 1$.

Our method for bounding the diameter through estimating $|N_i(x)|$ does not seem to work for this range. This difficulty can perhaps be explained by the following observation. The probability that $|\Gamma_1(x)| = 1$ is approximately $\frac{c}{e^c}$, a constant. Hence, the probability that

$$|\Gamma_1(x)| = |\Gamma_2(x)| = \ldots = |\Gamma_l(x)| = 1$$

is about $(\frac{c}{e^c})^l$. For some l up to $(1-\varepsilon)\frac{\log n}{c - \log c}$, this probability is at least $n^{1-\varepsilon}$. So it is quite likely that this may happen for vertex x. In other words, there is a nontrivial probability that the random graph around x is just a path starting at x of length $c \log n$. The i-th neighborhood $N_i(x)$ of x, for $i = c\log n$, does not grow at all!

In citedia we consider the case of $p > \frac{c \log n}{n}$. Does the statements still hold for $p = \frac{c \log n}{n}$?

Problem 3: Is it true that the diameter of $G(n, p)$ is concentrated on $2k+3$ values if $p = \frac{\log n}{kn}$?

It is worth mentioning that the case $k = 1$ is of special interest.

For the range $np = 1 + n^{-c}$, Lemma 1 in [8] implies $diam(G(n, p)) \geq (\frac{1}{1-3c} + o(1))\frac{\log n}{\log(np)}$. Can one establish a similar upper bound?

Problem 4: Is it true that

$$diam(G(n,p)) = \Theta(\frac{\log n}{\log(np)})$$

for $np = 1 + n^{-c}$?

Łuczak [10] proved that the diameter of $G(n, p)$ is equal to the diameter of a tree component in the subcritical phase $(1-np)n^{1/3} \to \infty$. What can we say about the diameter of $G(n, p)$ when $(1 - np)n^{1/3} \to c$, for some constant c? The diameter problem seems to be hard in this case.

References

[1] W. Aiello, F. Chung and L. Lu, A random graph model for massive graphs, *Proceedings of the Thirtysecond Annual ACM Symposium on Theory of Computing* (2000), 171-180,

[2] W. Aiello, Fan Chung and Linyuan Lu, Random evolution in power law graphs, preprint.

[3] Réka Albert, Hawoong Jeong, and Albert-László Barabási, Diameter of the World Wide Web, *Nature* **401** (1999) 130-131.

[4] N. Alon and J. H. Spencer, *The Probabilistic Method*, Wiley and Sons, New York, 1992.

[5] B. Bollobás, The Diameter of Random Graphs. *IEEE Trans. Inform. Theory* **36** (1990), no. 2, 285-288.

[6] B. Bollobás, The evolution of sparse graphs, *Graph Theory and Combinatorics* (Cambridge 1983), 35-57, Academic Press, London-New York, 1984.

[7] J. D. Burtin, Extremal metric characteristics of a random graph I, *Theo. Verojatnost. i Primenen* **19** (1974), 740-754.

[8] Fan Chung and Linyuan Lu, The diameter of random sparse graphs,

[9] P. Erdős and A. Rényi, On random graphs. I, *Publ. Math. Debrecen* **6** (1959), 290-291.

[10] Tomasz Łuczak, Random Trees and Random Graphs, *Random structures & algorithms* **13**, pp. 485-500.

[11] V. Klee and D. Larman, Diameters of random graphs. *Canad. J. Math.* **33** (1981) 618-640.

[12] A. V. Kostochka, A class of constructions for Turán's (3,4)-problem. *Combinatorica* **2** (1982): 187-192.