MCS 590 – Foundations of Data Science

Spring 2015

Problem Set 1

Lev Reyzin

Due: 2/13/15 at the beginning of class

Instructions: Atop your problem set, please write your name and list your collaborators.

Problems

1. Let $p > 0$ be a constant independent of n. Give a polynomial time algorithm that finds cliques of expected size $\Omega(\log n)$ in $G(n, p)$. Argue that your algorithm is correct.

2. Find the sharp threshold for p for the existence of 4-cliques in $G(n, p)$. Prove your answer correct.

3. The example at the end of Section 4.1.1 (and done in class) showed that if the degrees in $G(n, \frac{1}{n})$ were independent, there would be a vertex of degree $\frac{\log n}{\log \log n}$ with constant positive probability. However, the degrees are not independent. Show how to overcome this difficulty.

4.[1] Search the Web for two real-world graphs in machine readable form. Plot the degree distribution of the graphs. Give the average degree and number of connected components of the graphs. Describe your findings.

[1]No collaboration is allowed on this problem.