Graphs and Networks **out:** November 8, 2007

Problem Set 4

Lecturer: Daniel A. Spielman **due:** November 27, 2007

1 Introduction

There are 2 sections to this problem set. The first contains the 4 problems for the undergraduates. The second contains an additional problem for the graduate students. If you are doing a final project, you can skip this problem set.

2 Homework Policy

You may discuss the problems with other students. But, you must write your solutions independently, drawing on your own understanding. You should cite any sources that you use on the problem sets other than the textbook, TA and instructor. This means that you should list your collaborators.

You **may not** search the web for solutions to similar problems given out in other classes. If you think this policy needs any clarification, please let me know.

3 Corrections

1. In problem 2, I should have said "$p = (1+\epsilon)/k$", not $p = (1+\epsilon)k$. This is now fixed in the problem statement.

4 Undergraduate Problems

1. **Percolation on k-ary Trees**

   Consider the percolation problem on infinite k-ary trees with probability p of keeping each edge. Let $x_d(p)$ denote the probability that the root is connected to any leaf in the k-ary tree of depth d. In the base case, $d = 0$, the tree consists of a single vertex, and by assumption $x_0(p) = 1$.

   (a) Write an expression for $x_{d+1}(p)$ in terms of $x_d(p)$. In particular, find a function $f(x, p)$ such that

   $$x_{d+1}(p) = f(x_d(p), p).$$

(b) Show that for every ϵ there exists a constant c_ϵ such that if $p = (1+\epsilon)/k$ and $x < c_\epsilon$, then

$$f(x, p) > x.$$

You constant c_ϵ should be a function of ϵ alone, and in particular not depend on k.

It is easy to show that $x_{d+1}(p) \le x_d(p)$ for all d. So, if you've proved part (*b*), you may now conclude that $x_d \ge c_\epsilon$ for all d.

2. **FKG Inequality**

   Let $T_d^k(p)$ be the distribution on graphs obtained by keeping each edge of the depth-d complete k-ary tree with probability p.

   An increasing function on the space of graphs is a function $f(G)$ such that if H is a subgraph of G, then $f(H) \le f(G)$. The FKG Inequality says that if f and g are increasing functions then

$$\mathbf{E}_{G \leftarrow T_d^k(p)}\left[f(G)g(G)\right] \ge \mathbf{E}_{G \leftarrow T_d^k(p)}\left[f(G)\right] \mathbf{E}_{G \leftarrow T_d^k(p)}\left[g(G)\right].$$

   For $G \leftarrow T_d^k(p)$, let $A(G)$ be the event that G contains a path from the root to a leaf. Let $B(G)$ be the event that the root is connected to all of its k children.

   Prove that for $p = (1+\epsilon)/k$, there exists a constant c, independent of d but possibly depending on k and ϵ, such that

$$\Pr_{G \leftarrow T_d^k(p)}\left[A(G) \text{ and } B(G)\right] \ge c.$$

   Hint: If you use the result of the first problem, this problem is easy.

3. Let G be a random directed graph on n vertices in which each vertex has out-degree 1. That is, for every vertex i, we choose a j at random and add edge (i, j) to the graph. For a vertex v, let $R(v)$ denote the set of vertices reachable from v by a directed path.

   a. Prove that for a fixed vertex v, $\Pr\left[|R(v)| < \sqrt{n}/10\right] < 1/3$.
   b. Prove that for a fixed vertex v, $\Pr\left[|R(v)| > 10\sqrt{n}\right] < 1/3$.

   Hint: this is related to the "birthday paradox".

4. **Gossiping with an adversarial scheduler** In this problem, we are going to perform an analysis of the distributed averaging algorithm under the assumption that the communications that occur are being maliciously chosen, as opposed to occuring at random. In particular, we assume that a schedule of communications has been chosen in advance. This schedule consists of a sequence of pairs of vertices, $\left((u_i, v_i)\right)_i$, which indicates that in the ith step vertex u_i communicates with vertex v_i. When vertex u_i communicates with v_i, they average their values. In particular, if x_0 denotes the initial vector of values and x_i denotes the vector present at time i, then

$$x_{i+1}(u_i) = \frac{x_i(u_i) + x_i(v_i)}{2}, \text{ and}$$
$$x_{i+1}(v_i) = \frac{x_i(u_i) + x_i(v_i)}{2}.$$

Of course, if the schedule does not contain a set of edges that results in a connected graph, then the process need not converge. So, we will assume that there is some fixed number $\ell \geq n$ such that for every j, the graph

$$G = (\{1, \ldots, n\}, \{(u_i, v_i) : j \leq i \leq j + \ell\})$$

is connected (n is the number of vertices in the graph).

(a) Let M_j be the matrix such that
$$x_{j+\ell n} = M_j x_j.$$
Prove that every entry of M_j is at least $2^{-\ell n}$.
Hint: Let M_j^i denote the matrix such that $x_{j+\ell i} = M_j^i x_j$. For each u, consider the set of v such that $M_j^i(u, v) > 0$. Prove that this set has at least i elements.
(Note, if you can do part (b) without this part, you will get credit for this part too.)

(b) Consider the potential function $f(x) = \max_i x(i) - \min_i x(i)$. Prove that, for all j,
$$f(x_{j+\ell} n) \leq (1 - 2^{-\ell n}) f(x_j).$$

5 Graduate Problem

1. **Percolation on the Hexagonal Grid**

Consider percolation on the infinite hexagonal grid, of which a section is shown below.

(a) Prove that there is a constant $p_0 > 0$ such that if $p < p_0$ then the probability that the origin lies in an infinite component is zero.

(b) Prove that there are constants $p_1 < 1$ and $c > 0$ such that if $p > p_1$ then the probability that the origin lies in an infinite component is at least c.