Problems on small world

November 29, 2014

1) A 3D grid is the graph consisting of vertices at points (x, y, z), where each of these are integers, say all between 0 and n, and two vertices are connected by an edge if ONE of their coordinates differ by 1 and the others are the same (so there is an edge between $(0, 0, 1)$ and $(0, 1, 1)$). This is the obvious 3D "cube wireframe" analogy to the 2D grids we considered. Suppose we add in one random edge to each vertex v which goes to another vertex u with probability proportional to $d(u, v)^{-q}$. Fix v and let R_d be the ring of vertices a distance at least d from v and at most $2d$ from v. What value should q have so that the probability that the random edge from v jumps into R can be bound by a number independent of d (and hence independent of the choice of ring R_d)? Justify your answer.

2) Go to wikipedia.com and try to find a short path from "Lynah Faithful" to "Zona Zoo" using only links on each page. It may be helpful to know that "Lynah Faithful" refers to hockey fans at Cornell University in Ithaca, NY, and "Zona Zoo" refers to University of Arizona student fans. What is your strategy?

3) Consider the complete graph K_n, the path with n vertices P_n, and the cycle with n vertices. Compute the characteristic path length and the clustering coefficient.

4) On the 2D grid, show that there are no more than $4i$ vertices a distance i from a given vertex, and no less than $i + 1$ vertices a distance i from a given vertex.

5) Show that there are no more than 2^{2j+2} vertices in B_j.

In the next few questions, we see why our argument for decentralized search does not work for $q \neq 2$. (Note that Kleinberg shows that you expect it to take longer, but we are only seeing why our argument did not show these have short paths.) We will consider the situation in the notes.

6) Show the estimates:

a)
$$\sum_{w \neq v} d(w, v)^{-q} \leq c$$

for some constant c if $q > 2$ (hint: p-series converge for $p > 1$!) and

b)
$$\sum_{w \neq v} d(w, v)^{-q} \leq \frac{1}{2 - q} (2n)^{2-q}$$

if $q < 2$.

7) If $q > 2$, show that the probability P_j that a message in stage j (i.e., the message is currently in B_{j+1} but not B_j) enters B_j after one step satisfies

$$P_j \geq \frac{1}{c2^{2q+1}2^{(q-2)j}}$$

for the c in problem 6.

8) If $q > 2$, show that our bound for the expected number of steps in phase j is les than or equal to

$$\sum_{i=1}^{\infty}\left(1 - \frac{1}{2^{(q-2)j}2^{2q+1}c}\right)^i \leq c2^{2q+1}2^{(q-2)j}$$

and so the expected number of steps E in the myopic search is

$$E \leq 2^{2q+1}c\sum_{j=1}^{\log_2 n}2^{(q-2)j}$$

$$= 2^{2q+1}c\left[\frac{2^{(q-2)(\log_2 n+1)} - 1}{2^{(q-2)} - 1} - 1\right]$$

$$\leq Cn^{q-2}$$

for some C depending only on q. (Technically, this is only if $\log_2 n$ is an integer, but this estimate is good enough.)So the expected number of steps is n to a positive power.

9) If $q < 2$, show that the probablility P_j that a message in stage j (i.e., the message is currently in B_{j+1} but not B_j) enters B_j after one step satisfies

$$P_j \geq \frac{(2 - q)\,2^{(2-q)j}}{2^{2q+1}\,(2n)^{2-q}} = C\frac{2^{(2-q)j}}{n^{2-q}}$$

for appropriate C depending on q but not on j or n.

10) If $q < 2$, show that our bound for the expected number of steps in phase j is less than or equal to

$$\sum_{i=1}^{\infty}\left(1 - C\frac{2^{(2-q)j}}{n^{2-q}}\right)^i \leq \frac{n^{2-q}}{C2^{(2-q)j}} - 1$$

and so the expected number of steps E in the myopic search is

$$E \leq \sum_{j=1}^{\log_2 n}\frac{n^{2-q}}{C2^{(2-q)j}}$$

$$\leq \frac{n^{2-q}}{C}\left[\sum_{j=1}^{\infty}2^{-(2-q)j}\right]$$

$$= \left(\frac{1}{C}\frac{2^{-(2-q)}}{1 - 2^{-(2-q)}}\right)n^{2-q}$$

$$= C'n^{2-q}$$

for some constant C' depending only on q. So again it grows like n to a power.