1 The Small-World Phenomenon

Any two individuals, who do not know each other in the United States, are likely to be connected through less than six acquaintances.

2 Experiment by Milgram

Task: A source person receives a letter, which is to be delivered to a target person in Massachusetts. If the source person does not know the target person directly, then he forwards the letter to a friend who is more likely to know the target person. Repeat until the target person receives the letter.

3 Jon Kleinberg's Algorithm

The idea: Model the world as a 2-dimensional grid, in which each vertex is a person. There are two types of connections: local edges, in which a person knows his neighbors, represented by edges to adjacent vertices, and long distance edges, going from a vertex u to an arbitrary vertex v.
Start vertex: s
Current vertex: u
End vertex: t
Let's define the probability of a long distance edge existing as:
$Pr(u,v) \propto d^{-r}(u,v) = \frac{1}{d^r(u,v)}$, where r is parameter range in $[0,\infty)$.
It takes $\log n$ to find a short path.

Case $r = 0$:
End point pick uniformly at random.
Case $r < 2$:
Surely there exists short path but no local algorithm.
Case $r = 2$:
There exists efficient algorithm to find the short path.
Case $r > 2$:
Maybe no short path exists.
Case $r = \infty$:
No long distance edges exist.

u is in phase j when distance d(u,t) is in range $(2^j, 2^{j+1}]$. At each phase, finding short path takes $\log n$ time. Since we have $\log n$ phases, total expected time is $\log^2 n$.

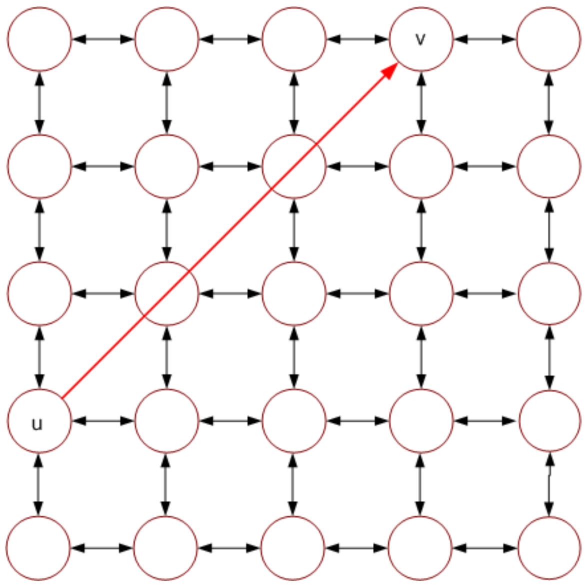

Figure 1: n by n grid, edge (u,v) represents long distance edge

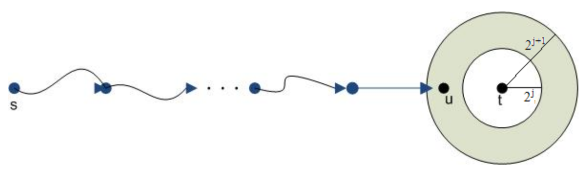

Figure 2: u is in phase j when distance d(u,t) is in range $(2^j, 2^{j+1}]$.

Lemma Select s and t uniformly at random on n x n grid, s and t will be at least n/4 apart with prob $\geq \frac{1}{2}$.

Proof

Consider case r=0:

There are order of n nodes in the circle. There are order n^2 nodes total.

Pr(long distance edge lands in circle) $= \frac{1}{n}$.

Pr(not in circle) $= 1 - \frac{1}{n} = 1 - \frac{1}{n}$.

Pr(none of $n^{1/2}$ long distance edges land in circle) $= \left(1 - \frac{1}{n}\right)^{n^{1/2}}$.

As $n \to \infty$,

$$\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^{n^{1/2}} = \left(1 - \frac{1}{n}\right)^{n * \frac{n^{1/2}}{n}} = \lim e^{-\frac{1}{n^{1/2}}} = 1.$$

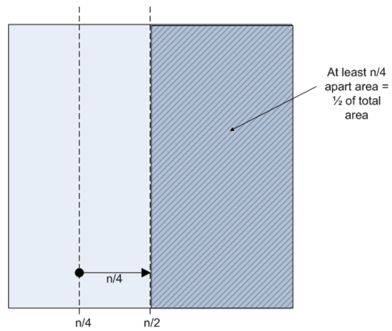

Figure 3: At least with probability 1/2, s and t are at least n/4 apart

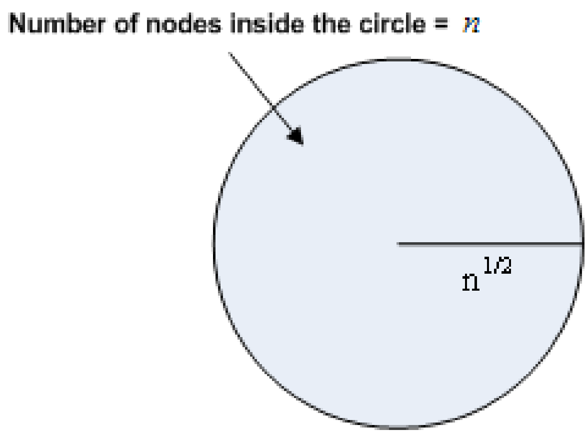

Figure 4: Number of nodes inside the circle.

4 Jon Kleinberg's algorithm for r=2

But why does the r=2 algorithm work?
Lemma
For r=2, there exists constant c such that probability of long distance edges from u to v is at least $c\,\frac{d^{-2}(u,v)}{\ln n}$.
Proof
Pr(long distance edge from u to v) = $\frac{d^{-2}(u,v)}{\sum\limits_{w \neq u} d^{-2}(u,w)}$
Since upper bound on denominator implies lower on proportionality,

$$\begin{aligned}
\sum_{w \neq u} d^{-2}(u,w) &\leq \sum_{i=1}^{2n-2} 4i \cdot \frac{1}{i^2} \\
&= 4 \sum \frac{1}{i} \\
&\leq c_1 \ln n
\end{aligned}$$