Social and Information Networks

Tutorial #5: Analyzing Decentralized Search

University of Toronto CSC303

Winter/Spring 2024

Ian Berlot-Attwell

Week 6: Feb 12-16

Today's agenda

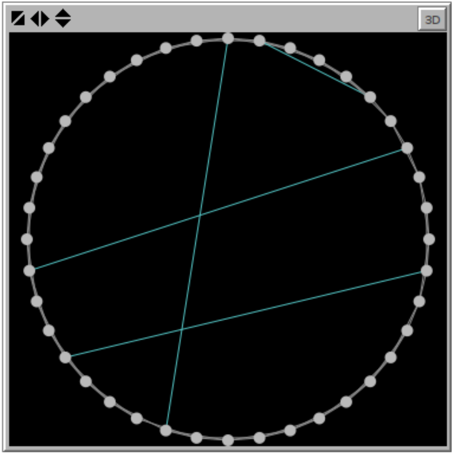
In lecture we've covered Chapter 20 of the textbook looking at Small Worlds and decentralized search.

Today:

- Questions from Lecture
- NetLogo small worlds demo
- Analysis of Decentralized Search in Small Worlds (Ch 20.7A of E&K)
- Quercus Quiz

Questions?

NetLogo Small Worlds Demo



Decentralized Search in Small Worlds

- From class we know that efficient decentralized search in the Watts-Strogatz model require selecting the endpoint of weak links with probability $\propto 1/rank$, where *rank* is the number of closer endpoints
- Recall $rank \approx d^2$ in 2D, and $rank \approx d$ in 1D
- Consider the 1D model where nodes are arranged in a ring, have a strong link to their immediate neighbours, and have a weak link to some other node

Decentralized Search in Small Worlds

Theorem (Efficient Decentralized Search in 1D)

Consider n nodes arranged in a ring. Each node gets 1 long-distance connection, chosen with probability proportional to $d(v, u)^{-1}$. Then expected decentralized search length is $O((\log_2 n)^2)$

Decentralized Search in Small Worlds

Let the random variable X be the number of steps until we reach our target. Then

$$X = \sum_{i=0}^{\log_2 n} X_i$$

Where X_i is the time spent with a distance in $[2^i, 2^{i+1})$. Let these be referred to as "phases" of the decentralized search.

Thus $E[X] = \sum_{i=0}^{\log_2 n} E[X_i]$

Decentralized Search in Small Worlds

Now, recall that the probability that the weak link for v is w is inversely proportional to $d(v, w)$, therefore it's equal to $\frac{1}{Zd(v,w)}$, for normalizing constant Z.

$$
\begin{aligned}
Z &= \sum_{v \in V, v \neq w} \frac{1}{d(v, w)} \\
&\leq 2 \sum_{d=1}^{\lfloor n/2 \rfloor} \frac{1}{d} = 2 + 2 \sum_{d=2}^{\lfloor n/2 \rfloor} \frac{1}{d} \\
&\leq 2 + 2 \int_{1}^{\lfloor n/2 \rfloor} \frac{1}{x} dx = 2 + 2 \ln(\lfloor n/2 \rfloor) \\
&\leq 2 + 2 \log_2(n/2) = 2 + 2 \log_2(n) - 2 \log_2(2) = 2 \log_2(n)
\end{aligned}
$$

Therefore, the probability of v having a weak link to w is $\frac{1}{Z} d(v, w)^{-1} \geq \frac{1}{2 \log_2(n)} d(v, w)^{-1}$

Decentralized Search in Small Worlds

Suppose that we are currently executing a decentralized search, and we are presently at distance d from our target.

Recall that we split the search into "phases", such that the ith phase is when we are at a distance $[2^i, 2^{i+1})$.

Therefore, the current "phase" will definitely end once we reach the distance of $d/2$ or less. There are $d + 1$ nodes at a distance of $d/2$ or less of the target. Let I be this set. Note that the node in I furthest from us must be at a distance of $d + d/2 = 3d/2$.

Decentralized Search in Small Worlds

I is the set of $d + 1$ nodes at a distance of $d/2$ or less of the target. Therefore if we are currently at node u:

$$
\begin{aligned}
P(u \text{ has a weak tie to } I) &= \sum_{w \in I} \frac{1}{Z} d(u, w)^{-1} = \sum_{x=d/2}^{3d/2} \frac{1}{Z} x^{-1} \\
&\geq \sum_{x=d/2}^{3d/2} \frac{1}{Z} (3d/2)^{-1} \\
&\geq \frac{d}{Z} (3d/2)^{-1} = \frac{2}{3Z} \\
&\geq \frac{1}{3 \log_2 n}
\end{aligned}
$$

Decentralized Search in Small Worlds

Ignoring the possibility of moving out of the phase through local connections, then each node has a probability of at least $\frac{1}{3\log_2 n}$ of exiting the phase.
Therefore the probability of staying in the phase for i steps is at most:

$$(1 - \frac{1}{3\log_2 n})^{i-1}$$

Decentralized Search in Small Worlds

Now, note:

$$
\begin{aligned}
E[X_j] &= \sum_{i=1}^{\infty} iP(X_j = i) \\
&= \sum_{i=1}^{\infty} P(X_j \geq i) \\
&\leq \sum_{i=1}^{\infty} (1 - \frac{1}{3\log_2 n})^{i-1} \\
&= \frac{1}{1 - (1 - \frac{1}{3\log_2 n})} \\
&= 3\log_2 n
\end{aligned}
$$

Decentralized Search in Small Worlds

Therefore, given $E[X_j] \leq 3\log_2 n$ we can conclude:

$$E[X] = \sum_{i=0}^{\log_2 n} E[X_i] \leq (1 + \log_2 n) \times 3\log_2 n \in O((\log_2 n)^2)$$

Quercus Quiz