NETWORK GRAPH MODELS I

ADD. READING
MUNG CHIANG CH. 9-10

In previous lectures we have discussed routing algorithms that transport packets with a small # hops to the destination.
We have also discussed addressing that allows a huge # hosts to be connected.

Routing: performs navigation over a huge # nodes with only a few hops.
How does the routing graph look like?

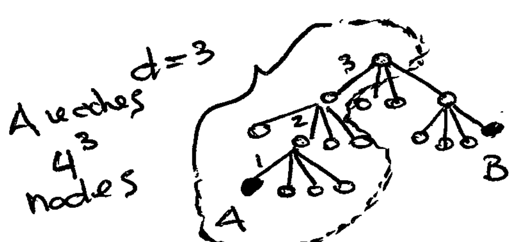
As you should know by clicking on the picture on our web page, this is not an easy question!
We will answer it in steps by looking at some "Toy generative models".

1. Let N be # number of nodes. We know this is a very large number
2. Let d be # hops (distance). We know this is a very small number compared to N.

For example $d \approx \log N$

> This means I can reach 1 Billion nodes with 9 hops. *Note*: the actual numbers for the internet are difficult to measure but let's assume this relationship is reasonably close to reality.

A simple way to have a network with small diameter d is to construct a "Tree" network



If every node has c children $\Rightarrow$ then
For every distance d a node can reach

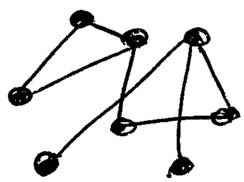
$$N = c^d \text{ other nodes}$$

$$\boxed{d = \frac{\log N}{\log c}}$$

This gives right relationship between d and N.
Do you think Internet can be so regular structure?
Of course NOT !

Let's look at the other extreme: a RANDOM-GRAPH
Erdos - Renyi '59 graph.

N nodes , $P_{ij} = p$ indep. $\forall$ edges. This provides a mathematical mode $G(n,p)$



QUESTION 1. What is the degree distribution?

Let k_i = # neighbors of node i

$$\forall i, \; P(k_i = k) = \underbrace{\binom{N-1}{k}}_{\text{all ways we can choose } k \text{ nodes out of } N-1} \underbrace{p^k}_{\text{Prob connect to the chosen } k} \underbrace{(1-p)^{N-1-k}}_{\text{prob not connect remaining } N-1-k}$$

This is Binomial distribution

QUESTION 2. What is the expected node degree?

E(# connections per node) $= (N-1) \cdot p \approx Np$

QUESTION 3 To maintain a constant expected node degree, we need $p \to 0$ as $N \to \infty$ so that $N \cdot p \to$ constant

Choose $p = p(N)$: $\boxed{\lambda = p(N) \cdot N}$

What happens now to degree distribution when N is large?

$$P(k_i = k) \longrightarrow e^{-Np} \frac{(Np)^k}{k!} = e^{-\lambda} \frac{\lambda^k}{k!}$$

This is Poisson distribution

EXERCISE: Prove convergence to Poisson

Let $N-1=n$

$$\lim_{n\to\infty} \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$= \lim_{n\to\infty} \frac{n!}{k!\,(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$= \lim_{n\to\infty} \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n \underbrace{\frac{n!}{(n-k)!\,n^k}}_{\to 1}\ \underbrace{\left(1-\frac{\lambda}{n}\right)^{-k}}_{\to 1}$$

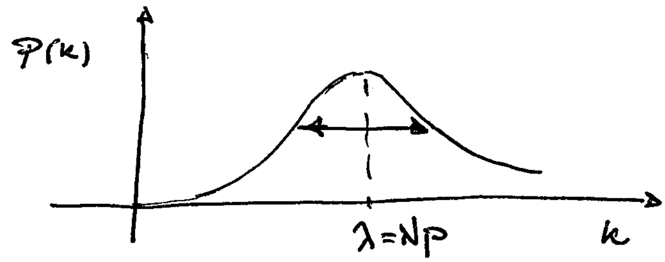
$$= \boxed{\frac{\lambda^k}{k!}\, e^{-\lambda}}$$

$\cancel{\left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda}}$

$$\left(1+\frac{x}{n}\right)^n \longrightarrow e^x$$

$$\left(1-\frac{\lambda}{n}\right)^{n-k} = \left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

How does node degree distr. look like ?



most nodes have a degree "around the average"
The probability of deviation far from average tends to zero exponentially fast.

EXERCISE Try to plot Poisson distr. for different values of λ
what happens when λ is very small or very large?

EXERCISE Compute standard deviation of Poisson distribution.

For this model of random graph, the average path length between any two nodes is:

$$\boxed{\mathbb{E}(d) = \frac{\log N}{\log \lambda}}$$

Another interesting phenomenon is the emergence of a "giant" component

$\mathbb{E}$ (node degree) $= np = \lambda > 1$ $\Rightarrow$ a giant connected cluster of size proportional to N appears

$np = \lambda < 1$ $\Rightarrow$ no giant cluster exists

$\lambda = \log n$ $\Rightarrow$ Fully connected network, giant cluster has size N.

The idea is that the random graph is not "so different" than the tree: with λ connections per node, we reach
$N = \lambda^d$ nodes at distance d

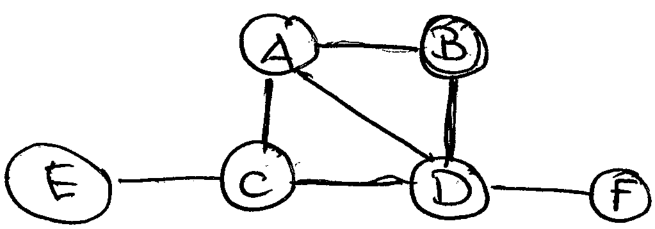
$$\boxed{d = \frac{\log N}{\log \lambda}}$$

Also, having more than 1 connection appears enough to "not get stuck" and have a giant component.

A lot of good properties... BUT...
Problem with random graph model.
It's too random!
It does not exhibit $\boxed{\text{local clustering}}$

Clustering coefficient
DEFINITION:

proportion of neighbors of a node A that are also neighbors of themselves.
Example



$$\text{Cluster}(A) = \frac{2}{3}$$

$$\sum_{A} C(A)\,\frac{1}{N} = \text{network average clustering coeff.}$$

For random graph the network clustering is

$$\frac{\mathbb{E}(\text{node degree})}{N} = \frac{PN}{N} = P$$

This means that probability two neighbors of a node are connected is the same as the prob any two nodes are connected.
This is NOT reflected in real networks!
Also it implies clustering $\to 0$ as $N \to \infty$.
This reinforces the view that random graph looks much like a tree as ~~the~~ $N \to \infty$.

Can we obtain a model that has both high clustering and small diameter?