Percolation and Network Resilience

Percolation is a process of removing some fraction of network’s nodes with adjacent edges. (more precisely site/link/cluster percolation)

- models real-life phenomena such as router failure, immunization of people, and disasters
- the process is parameterized by occupation probability $\phi$
- **Percolation transition**: when $\phi$ is large there is a giant component but as $\phi$ is decreased then gc breaks into many small components or clusters (similar to phase transition in Poisson random graphs with gc$\rightarrow$sc)

\[
\phi = 1 \quad \text{functional vertices} \quad \phi = \frac{6}{8}
\]
Percolation and Configuration Model

Consider a configuration model with
- degree distribution $p_k$
- occupation probability $\phi$

consider node $i$ which

- can belong to gc, i.e., connected to it through some $j \in N(i)$
- is not in gc (i.e., not connected to it via any of $N(i)$) - define the avg probability of it $u^k$, where $k = \text{deg}(i)$ and $u$ is the same prob for one particular neighbor
- avg probability of not being in gc

$$\sum_k p_k u^k = g_0(u), \text{ where } g_0(z) = \sum_k p_k z^k \text{ or } \Pr[i \in \text{gc}] = 1 - g_0(u)$$

- total fraction of nodes in gc when percolation is running

$$S = \phi(1 - g_0(u))$$
Let us calculate $u$, the probability that $i$ is not connected to gc via a particular neighbor. Two cases:

- $i$ is connected to $j$ which is removed with prob $1 - \phi$
- or $j$ is not removed with prob $\phi$ but it is not in gc

\[
\Pr[i \notin gc \text{ via } j] = 1 - \phi + \phi u^k
\]

Node $j$ is reached by following an edge, so average probability

\[
u = \sum_{k=0}^{\infty} q_k (1 - \phi + \phi u^k) = 1 - \phi + \phi + \sum q_k u^k = 1 - \phi + \phi g_1(u)
\]

(see handout for graphical solution of the equation)

[Cohen, Erez, Ben-Avraham, Halvin]: $\phi_c = \frac{1}{g_1'(1)} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$

Reminder: excess degree is the number of edges attached to a vertex other than the edge we arrived along.

\[
q_k = \frac{(k + 1) p_{k+1}}{\langle k \rangle}, \quad \sum q_k = 1, \quad g_1(z) = \sum_{k=0}^{\infty} q_k z^k
\]
Introduction to Network Analysis

[Cohen, Erez, Ben-Avraham, Halvin]: \( \phi_c = \frac{1}{g_1'(1)} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} \)

- In configuration model fixing low \( \phi_c \) leads to gc, for example \( \langle k^2 \rangle > \langle k \rangle \)
- Example: given Poisson degree distribution with \( c \) - mean degree

\[
p_k = e^{-c} \frac{c^k}{k!} \Rightarrow \langle k \rangle = c, \langle k^2 \rangle = c(c + 1) \Rightarrow \phi_c = \frac{1}{c}
\]

i.e., \( c = 4 \) means that \( \frac{3}{4} \) vertices will fail before gc disappears.

- Example: power laws with \( 2 < \alpha < 3 \) (Internet, etc)

\[\langle k \rangle \text{ is final, } \langle k^2 \rangle \text{ diverges } \Rightarrow \phi_c \to 0\]

i.e., remove many vertices form the network \( \Rightarrow \) gc will be there.

- Opposite Example: Epidemiological networks. Small \( \phi_c \) are bad! The fewer individuals we need to vaccinate to destroy gc the better.
Introduction to Network Analysis

[Cohen, Erez, Ben-Avraham, Halvin]: \( \phi_c = \frac{1}{g'_1(1)} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} \)

- Example: Exponential degree distribution \( p_k = (1 - e^{-\lambda})e^{-\lambda k}, \lambda > 0 \)

\[
g_0(z) = \frac{e^\lambda - 1}{e^\lambda - z}, \quad g_1(z) = \left( \frac{e^\lambda - 1}{e^\lambda - z} \right)^2 \Rightarrow
\]

\[
u(e^\lambda - u)^2 - (1 - \phi)(e^\lambda - u)^2 - \phi(e^\lambda - 1)^2 = 0 \Rightarrow
\]

\[
u = e^\lambda - \frac{1}{2}\phi - \sqrt{\frac{1}{4}\phi^2 + \phi(e^\lambda - 1)} \Rightarrow
\]

\[
S = \frac{3}{2}\phi - \sqrt{\frac{1}{4}\phi^2 + \phi(e^\lambda - 1)} \Rightarrow
\]

\[
\phi_c = \frac{1}{2}(e^\lambda - 1)
\]

Size of gc

Percolation threshold

\( u=1 \) is always a solution and (u-1) is always a factor
Figure 16.4: Size of the giant cluster for site percolation in the configuration model. The curve indicates the size of the giant cluster for a configuration model with an exponential degree distribution of the form (16.12) with \( \lambda = \frac{1}{2} \), as given by Eq. (16.18). The dotted line indicates the position of the percolation transition, Eq. (16.20).
Figure 16.5: Size of the giant cluster for a network with power-law degree distribution. The size of the giant cluster for a scale-free configuration model network with exponent $\alpha = 2.5$, a typical value for real-world networks. Note the non-linear form of the curve near $\phi = 0$, which means that $S$, while technically non-zero, becomes very small in this regime. Contrast this figure with Fig. 16.4 for the giant cluster size in a network with an exponential degree distribution.
Geometric graph: percolation example
Non-uniform Removal of Vertices

Figure 16.7: Removal of the highest-degree vertices in a scale-free network. (a) The size of the giant cluster in a configuration model network with a power-law degree distribution as vertices are removed in order of their degree, starting with the highest-degree vertices. Only a small fraction of the vertices need be removed to destroy the giant cluster completely. (b) The fraction of vertices that must be removed to destroy the giant cluster as a function of the exponent $\alpha$ of the power-law distribution. For no value of $\alpha$ does the fraction required exceed 3%.
Figure 16.6: Size of the giant percolation cluster as the highest degree vertices in a network are removed. (a) The size of the giant cluster in a network with an exponential degree distribution $p_k \sim e^{-\lambda k}$ with $\lambda = \frac{1}{2}$ as vertices are removed in order of degree, starting from those with the highest degree. The curve is shown as a function of the degree $k_0$ of the highest-degree vertex remaining in the network. Technically, since $k_0$ must be an integer, the plot is only valid at the integer points marked by the circles; the curves are just an aid to the eye. (b) The same data plotted now as a function of the fraction $\phi$ of vertices remaining in the network.