Percolation I: Two Dimensions and Mean Field

Given a connected host graph H (infinite or finite), we can construct a random subgraph $G = G(p)$ by independently placing each edge into the subgraph with probability p. Percolation theory involves the study of the connected components of these graphs. Physicists often consider H to be the d-dimensional Euclidean Lattice, or the infinite d-regular tree (i.e. the tree in which every vertex has d neighbours).

One of the most fundamental questions is to examine the behaviour of (the random variable which is) the size of the largest component or the size of the component containing a particular vertex v. If G is infinite, we want to know the probability that there is an infinite component or the probability that v is in an infinite component. We denote this latter probability by $\Theta(p, v)$. For graphs like the Euclidean Lattice, for which symmetry tells us that this value is the same for every vertex, we simply use $\Theta(p)$. We restrict our attention to such graphs when discussing Θ in the future.

We remark that we must be careful when considering such infinite product spaces, as the probability of any particular graph being $G(p)$ will be 0. It is easy to determine the probability of events which are determined by a finite number of edges as we can just restrict our attention to the edges under consideration. Thus, in the 2-dimensional square lattice, the probability the origin is a single vertex component of $G(p)$ is $(1-p)^4$. In the same way we can determine the probability that the origin is in a component of size n for any integer n by considering a finite subset of the lattice. The probability that the origin is in a finite component is just the sum of these probabilities, and the probability that it is in an infinite component is obtained by subtracting this sum from 1.

Alternatively, we can consider a complete graph (all pairs of vertices connected by an edge) on n vertices and examine the size of the largest component. This latter model is known as $G_{n,p}$ (or the Erdos-Renyi model of random graphs) or the Mean Field Model.

Obviously, if $p = 1$ then $G = H$. If $p = 0$ then every component of G has one vertex. Further, an easy coupling argument (in which we assign independent random variables uniformily chosen from $[0..1]$ to each edge and let $G(p)$ be the graph of those edges whose variable is at most p), shows that $\Theta(p)$ is a monotonic function of p. A key question is to determine, given an infinite graph G, the critical value $p_c = p_c(G)$ such that $\Theta(p)$ is zero for $p < p_c$ and positive for $p > p_c$. A second question is to determine whether $\Theta(p_c)$ is positive or zero.

The theory of branching processes tells us immediately that if H is the infinite tree in which every node has $d-1$ children then the probability that the root is in an infinite component is 0 if $p < \frac{1}{d-1}$ and is positive if $p > \frac{1}{d-1}$. Since the root r of the d-regular infinite tree T_d is in an infinite component precisely if at least one of its children is in an infinite component of the subtree of $T_d - r$ rooted at it, this implies immediately that $p_c(T_d)$ is $\frac{1}{d-1}$.

In bounding p_c for the 2-dimensional lattice, we used the fact that the number of paths or cycles of the lattice of length l containing a given vertex is at most $4(3^{l-1})$. Thus, the probability p_l that the origin is in a path of length l of $G(P)$ is at most $4(3p)^{l-1}$ which goes to zero as l goes to infinity if $p < \frac{1}{3}$. It is easy to see that if the origin is in an infinite component then it is in a path of length l for every l. So the probability that this is happens is the infemum of the p_l which is 0 if $p < \frac{1}{3}$. Thus $p_c \geq \frac{1}{3}$.

We also saw that the critical probability for the square lattice was at most $2/3$. In doing so, we used the fact that our process for building $G(p)$ at the same time formed a subgraph $D(p)$ of the dual lattice, where each edge was present independently with probability $1-p$ (the edges of the dual lattice in $D(p)$ corresponded exactly to the edges of the lattice not in $G(p)$). We noted that the origin was in a finite component of the lattice precisely if it was surrounded by a cycle of $D(p)$. We then showed that for any $p > \frac{2}{3}$, the probability that this occurred was less than one.

The key to doing so was to note that any such cycle of length l would have to use one of the dual edges corresponding to an edge of the lattice joining two vertices with both x-coordinate and y-coordinate at most $\frac{l}{2}$. So, our bounds above tell us that the expected cycles of length l in $D(P)$ surrounding the origin is $O(l^2(3i(1-p)^{l-1})$. Taking an infinite sum, we see that the expected number of cycles of length more than l surrounding the origin goes to zero as l goes to infinity. Hence we can choose an l^* such that expected number of such cycles is less than $\frac{1}{2}$.

We then noted that the event E that the entire set S of edges joining two vertices with both x-coordinate and y-coordinate at most $\frac{l^*}{2}$ is in $G(p)$ satisfies $Prob(E) = p^{|S|} > 0$ and conditioned on this being true.

The conditional expected number of cycles of $D(p)$ of length at least l^* given E is at most the unconditioned number, as cycles intersecting S cannot occur in $D(p)$ given the conditioning and the other cycles occurr with the same probability irrespective of conditioning. Since given E, no cycle of $D(p)$ of length less than l can surround the origin, it follows that the conditional number of cycles of $D(p)$ surrounding the origin is less than $\frac{1}{2}$ and hence the conditional probability the origin is in an infinite component exceeds $\frac{1}{2}$. Thus, $\Theta(p) > \frac{Prob(E)}{2} > 0$ as claimed.

We mentioned that the critical probability for the 2-dimensional lattice is $\frac{1}{2}$. If we have time and there is interest we will prove this later.

The reference for this last result and a general introduction to Percolation is Grimmett's excellent text pages 1-19.

The branching process approach also told us that $G_{n,p}$ almost surely had only components of size $O(n^{\frac{2}{3}})$ when $p < \frac{1-\epsilon}{n}$ for some $\epsilon = 0$. but had a component with a constant proportion of the vertices with positive probability when $p > \frac{1+\epsilon}{n}$ for some $\epsilon > 0$.

To prove the first result we noted that the expected size of the component of

$G_{n,p}$ containing a vertex is strictly less than the expected size of the tree built by a branching process where the number of children of a node is $Bin(n,p)$ (since in building the component of $G_{n,p}$ the number of new nodes added when we explore the i^th node of the component is $n-k$ where k is the number of nodes already in the tree). Since this is a constant provided $p < \frac{1-\epsilon}{n}$ for some $\epsilon > 0$, it follows that the probability v is in a component of size exceeding $n^{2/3}$ is $O(n^{\frac{-2}{3}})$. Hence the expected number of vertices in such components is $O(n^{1/3})$. Since if there is such a large component there are $n^{2/3}$ vertices in it, Markov's inequality tells us that the probability $G_{n,p}$ has a component of size exceeding $n^{2/3}$ is $O(n^{1/3})$.

To prove the second result we noted that the probability that a vertex v was in a component of size δn is at least the probability that a branching process where the number of children is $Bin((1-\delta)n,p)$ has size at least δn(since until we grow a component of size δn this process underestimates the number of new vertices added to this component each time we explore one of its nodes). This probability is at least the probability the branching process does not become extinct. So, if $p > \frac{1+\epsilon}{n}$ for some $1 > \epsilon > 0$ then the expected number of nodes of $G_{n,p}$ in components of size at least $\frac{\epsilon n}{2}$ is νn for some $\nu > 0$ and hence the probability there is such a component is at least ν.