An Introduction to Percolation Theory -II

Deepak Dhar

I.C.T.S., T.I.F.R., Bengaluru

Bengaluru School in Statistical Physics XV, Sep 3-20, 2024

Cluster enumerations

The key simplifying feature of the percolation problem is the fact that all local properties in this sytem are easy to determine, using a finite algorithm, even if the system size is large.

Consider the square lattice, site percolation problem. We pick a randomly generated configuration of the system at concentration p. Denote the probability that the origin belongs to a cluster of exactly s sites as $\mathrm{Prob}(s)$.

Then it is easy to see that
$\mathrm{Prob}(s = 0) = 1 - p \equiv q$,
$\mathrm{Prob}(s = 1) = pq^4$,
$\mathrm{Prob}(s = 2) = 4p^2q^6$,
and so on. Explicit expressions for $\mathrm{Prob}(s)$ for $s = 3, 4, 5..$ can be written down with a bit of effort. Each is a polynomial in p.

There is a finite number of connected clusters containg the origin having exactly s sites. Denote the number of clusters having s sites and t external perimeter sites as $A(s,t)$, then as the probability of a specified cluster having s sites and t perimeter sites is $p^s q^t$. Then, we have

$$\text{Prob}(s) = \sum_t A(s,t) p^s q^t.$$

We will write the total number of clusters having s sites as

$$A(s) = \sum_t A(s,t).$$

Pictures of such clusters were seen as reminiscent of a collection of cells on a petri-dish, and were called *animals*.

For larger values of s, it is convenient to get a computer program to calculate $A(s)$ for a given s. One well known efficient algorithm for this called the Martin's algorithm.

Martin's algorithm for enumeration of clusters

We will consider, for simplicity, clusters on the square lattice. Extension to other lattices is straight forward.

The algorithm is a particular application of the general 'depth first' search in computer science. It will generate all clusters of size $s \leq s_{max}$, efficiently, and without ommisions or repetitions. It is described next.

First we discuss how to generate a unique label for each allowed configuration of a given cluster.

Sketch of the computer program

1. Start at the origin, and give it label 1.
2. All unlabelled neighbors are given the next available labels (here $2, 3, 4, 5$) , with a given priority rule (say $E > N > W > S$), where the N, E, W, S refer to direction of the neighbor.
4. Go the the topmost unvisited occupied site in the list of labelled sites. Assign labels to its unlabelled neighbors sequencially, using the priority rule.
5. Go back to step 4.

The algorithm stops when there are no unlabelled occupied sites. The list of labelled sites will have more than N entries, as some are unoccupied. We will identify the cluster using the sequence 1001001000..., where the r-th binary digit gives the occupancy status of the site labelled r.

Clearly, the algorithm assigns a unique label to each cluster. Furthermore, given the label, the cluster structure can be uniquely inferred.

As each added site can add at most two to the number of perimeter sites to a cluster, the binary string corresponding to a cluster of s sites in the Martin's algorithm can have at most $3s + 2$ bits. Also, there are binary strings that do not correspond to any cluster, e.g. 100001.
We thus conclude that on the square lattice,

$$A(s) \leq 2^{3s+2}.$$

We can take the list of all allowed labels having length at most $3s + 2$, and arrange them sequentially them in a list using the dictionay order.

Generating the list of all allowed words upto a maximum length can be done efficiently, using the back-track algorithm.

A very elegant and short computer program in Fortran that implements this for a general translationally invariant lattice, using less than 40 lines of program, is given in
S. Redner, J. Stat. Phys., 29, 309 (1982).

It is easy to get a lower bound on $A(s)$. We can start at the origin, and add particles sequentially at the neighbor of last placed particle in the N or E direction. This generates a directed walk of $(s-1)$ steps, and we get

$$A(s) \geq 2^{s-1}.$$

Coupled with the upper bound derived earlier, it makes plausible that $A(s)$ grows exponentially with s. In fact, the expected asymptotic behavior for large s is

$$A(s) \approx C\lambda^s s^{-\theta}[1 + \text{ subleading correction terms}].$$

The argument is expected to be extended to other lattices. The exponent θ is expected to be the same, for different lattices of the same dimension.

We note that $Prob(s) = \sum_t p^s q^t A(s,t) \leq A(s) p^s$. Thus we conclude that
For all $p < (1/\lambda)$, $\mathrm{Prob}(s)$ decays exponentially with s for large s.

We can similarly show the two-point correlation function $G(\vec{R})$ that two point at a separaion $\vec{R}$ belong to the same cluster decays to exponentially fast with $|\vec{R}|$ as $|r| \to \infty$.

Proof: If two points belong to the same cluster, there exists a self-avoiding path on occupied sites from one to the other. There may be more than one path. But, clearly,

$$G(\vec{R}) \leq \sum_P \mathrm{Prob}(P),$$

where the sum is over all self avoiding path P between the two points. But the probability of a path of length L is p^L. If the number of self avoiding paths of length L is $N_{SAW}(L, \vec{R})$, we have $G(\vec{R}) \leq \sum_{L=|\vec{R}|}^{\infty} p^L N_{SAW}(L, \vec{R})$.

But clearly, $N_{SAW}(L, \vec{R}) \leq 4.3^{L-1}$.
Hence for $p < 1/3$, $G(\vec{R})$ decreases at least exponentially with $|\vec{R}|$.

Lower bounds on $P_\infty(p)$ at high densities

We will now sketch the proof that if p is large enough, then the probability of infinite cluster is bounded above zero. This argument is a variation of the Peierls' proof of the existence of spontaeous magnetization in the Ising model at low enough temperatures.

Consider site percolation on the infinite square lattice, with the density p of occupied sites. If the origin belongs to an infinite cluster, there should not exist any simple closed loop $\mathcal{L}$ of unoccupied sites that surrounds the origin.
Then, $1 - P_\infty(p) \leq \sum_{\mathcal{L}} \mathrm{Prob}(\mathcal{L})$.

(1) Let N_ℓ be the number of distinct shapes of closed loops of perimeter ℓ. The number of closed self-avoiding loops satisfies the bound $N_\ell < 4.3^{\ell-1}$.
(2) For a loop of perimeter ℓ, the probability that all sites are unoccupied is q^ℓ.
(3) A loop of perimeter ℓ has an area (number of sites enclosed) at most $\ell^2/16$. Then, there are at most $\ell^2/16$ ways of placing a particular loop $\mathcal{L}$ of length *ell* so that it encloses the origin.
Putting these together, we get
$1 - P_\infty(p) \leq \sum_{\ell=4,6,8..} 4.3^{\ell-1} q^\ell (\ell^2/16)$.

The RHS is a convergent sum for $q < 1/4$, and deceases as q decreases. So we can choose q small enough, so that the sum is less than 1. Hence proved.

We showed that in the low density phase, the cluster size distribution decays exponentially with the cluster size s for large s. In fact, while the result was proved for small p, one expects this result to be true for all $p < p_c$, and the probability of clusters of size greater than s is bounded from above by $\exp[-C(p)s]$, where $C(p)$ is a decreasing function of p that tends to zero, as p tends to p_c from below.

In the regime $p > p_c$, of course, the cluster size on an infinite lattice could be strictly infinite. But we can ask how does the probability of a finite large cluster of s sites decrease with s? Interestingly, it does not decay exponentially fast. There is a rigorous result due to Kunz and Suillard [Phys. Rev. Lett. 40(1978) p133] that it decays only as a stretched expnential, as $\exp[-Cs^{\frac{d-1}{d}}]$, in d-dimensions. We give only a heuristic proof here.

Consider a typical configuration of the system at $p > p_c$. We mark a hypercube of linear size ℓ. Then the probability that all the boundary sites of this hypercube are absent is $q^{2d\ell^{d-1}}$. But the number of sites in the largest cluster inside is with a large probability of $\mathcal{O}(\ell^d)$. Choose $\ell = \alpha s^{1/d}$, with α to be large enough so that is $> s$. Then, there exists a $\beta > 0$ such that

$$\mathrm{Prob}(s > s_0) \geq q^{\beta s^{\frac{d-1}{d}}}.$$

Relation of the Percolation model to the q-state Potts model

One can express the free energies of a class of theoretical models as the large deviation function of the number density of clusters in the percolation model.

Let us consider the moment generating function of the r-th moment of the total number of clusters on a lattice of V sites in the bond-percolation problem at concentration p . We define the large deviation function $f(x)$ by the relation that for large V,

$$\text{Prob[number density of clusters } = \text{x}] \sim e^{-Vf(x)}.$$

Then it turns out that this function $f(x)$ is actually related to the free energy function of other well-studied models, e.g. the Ising model.

We start by considering a model where the at each site i , there is a spin σ_i that can occur in of q discrete states (q is an integer). For example, $q = 2$ is the Ising model. The Hamilonian of the system is

$$H = -\sum_{<ij>} J_{ij}\delta(\sigma_i, \sigma_j).$$

Here the δ- function is the Kronecker δ- function, taking values 1 when $\sigma_i = \sigma_j$, and zero otherwise. Then the partition function is given by

$$\mathcal{Z}_V = \sum_{\{\sigma_i\}} \prod_{<ij>} (1 + v\delta(\sigma_i, \sigma_j)),$$

where we have written $v = e^{\beta J} - 1$, and the sum $\{\sigma_i\}$ is over all the q^V spin configurations.

We expand this product in powers of v, and in a graphical representation of each term, where we draw a bond between vertices i and j, iff we use term $v\delta(\sigma_i, \sigma_j)$ in the factor in the product. Then, the different terms correspond to all possible states of the bonds. The weight of a configuration with B bonds is v^B.

Now the summations over the spins can be done for each term, and a term with C clusters gives the sum as q^C. thus the partition function can be written as

$$\mathcal{Z}(v, q) = \sum q^C v^B.$$

If we put $q = 1$, one generates bond percolation configurations with correct relative weights, if we put $v = p/q$. For $q = 2$, we get the partition function of the Ising model, with appropriate identification of variables. In the limit $q = 1$, the partition function is trivial $\mathcal{Z}(q = 1, v) = (1 + v)^B$, but taking derivatives of $\frac{\partial^r}{\partial q^r} \log \mathcal{Z}$ we get the moments of cluster numbers $< C^r >$.

More interestingly, the expression for $\mathcal{Z}$ is a polynomila in q, and is well defined for all $q \geq 0$. One can take the large volume limits

$$\lim_{V \to \infty} \log \mathcal{Z}_V / V = f(v, q),$$

to define free energy per site $f(q, v)$. This allows us to define the Potts model free energy for all real q.
This generalization of the q-state Potts model for non-integer q was the deep insight of Fortuin and Kasteleyn [Physica. 57 (4) (1972) 536]. It is interesting because it turns out that the critical exponents of this model in 2-dimensions for general q can be determined exactly using the conformal field theory, and help us study what different critical exponents depend on and how.

Exercises

1. Write down the probabilities $\mathrm{Prob}(s)$ that the origin belongs to a cluster of size s for $s = 0$ to 5, as expicit polynomials in p alone. Verify that the sum differs from 1 only by terms of order p^6.

2. How well can you improve the bound in the Peierls' type proof given here?