Journal of Statistical Physics, Vol. 66, Nos. 1/2, 1992

Cluster Size Distribution at Criticality

D. C. Rapaport[1]

Received July 9, 1991

Monte Carlo studies of the cluster size distribution for the site percolation problem on the triangular lattice are extended to lattices with up to 4×10^{11} sites. Agreement with the predictions of scaling theory at p_c is excellent over a range of cluster sizes spanning five orders of magnitude.

KEY WORDS: Percolation; cluster properties; scaling theory.

The percolation problem[1] is one of the most concisely formulated lattice statistical models that continues to defy analytical solution; it has, however, proved to be an apparently willing candidate for numerical study. A variety of methods, including Monte Carlo simulation[2,3] and exact series analysis,[4] have been used to elucidate many of the most interesting properties of this model, and seemingly disparate results have been consolidated within the framework of scaling theory.[1]

In this brief communication we describe an extension of an earlier study[3] of the distribution of cluster sizes on the two-dimensional triangular lattice to systems whose size (i.e., area) is larger by a factor of 16. The aim of this work is to establish the degree to which the essentially phenomenological scaling theory is capable of describing the behavior of the model.

When the site occupation probability is at the critical value p_c ($p_c = 0.5$ for the triangular lattice), the prediction of scaling theory[1] is that the number of clusters of s sites, normalized per lattice site, is

$$n_s \propto s^{-\tau} \tag{1}$$

[1] Physics Department, Bar-Ilan University, Ramat-Gan 52900, Israel.

0022-4715/92/0100-0679$06.50/0

For two-dimensional lattices the result $\tau = 187/91$ is believed to be exact. The scaling expression (1) is actually the leading-order term in a large-s expansion for n_s, but the neglect of higher-order corrections should not be significant at sufficiently large s. On the other hand, there are limitations as to how large s can become on a finite lattice, because, as soon as the typical cluster is capable of spanning the lattice, finite-size effects will begin to affect the results. Thus there is a range of s over which (1) might be expected to provide a good fit to the data; the extent of this range must be determined by numerical means, with previous results[3] already suggesting a broad range of applicability. Extending the linear size of the lattice from the earlier $L = 1.6 \times 10^5$ to $L = 6.4 \times 10^5$ should allow an even more precise test of the scaling theory.

It is inefficient[3] to generate the cluster distribution of a lattice this size in a single calculation, especially in view of the potentially large memory requirements of the direct approach and the nature of the computations involved. A far more effective approach is to subdivide the lattice into a large number of sublattices which are generated independently and then joined. This approach has been described at length elsewhere[5] in a review of methods for lattice-statistical problems; the principal algorithms employed in this calculation will be found there. One particular advantage of the method is that the properties of all the sublattices are obtained at no extra cost.

A very brief summary of the approach is as follows. The computation involves independently generating 1024 sublattices, each with $L = 2 \times 10^4$. For each sublattice, the sites are randomly populated with probability $p = p_c$, and the complete clusters thus formed—those that do not extend to any of the sublattice edges—are enumerated; in addition, a list of those incomplete clusters that do actually touch the sublattice edges is compiled. The technique used for cluster accounting is based on the well-known multiple labeling method.[6] Sets of four distinct sublattices are then joined, resulting in 256 sublattices with edge $L = 4 \times 10^4$; clusters that are linked across mutual sublattice boundaries are suitably merged. This process is repeated several more times until only a single full-size $L = 6.4 \times 10^5$ lattice remains. At each stage of the procedure the cluster distributions of the combined sublattices are computed after periodic boundary conditions have been taken into consideration, and of course the final full-size lattice is treated in similar fashion. Thus, the quality of the results based on smaller L benefits from a comparatively large sample size, whereas the largest lattice is realized only once. The computations required approximately 200 hr on an IBM 6000/320 workstation.

In order to reduce both the amount of data that must be recorded and the inherent statistical scatter, the cluster results are stored in histogram

form with a bin width which doubles at each successive interval. The results of the computation are the partial sums

$$G_s = \sum_{s'=2^{k-1}}^{2^k-1} n_{s'}, \qquad s = 2^k \tag{2}$$

The expected behavior of G_s can be derived from the scaling form (1) by integration; to leading order the result is

$$G_s = \text{const} \times s_{\text{av}}^{1-\tau} \tag{3}$$

where

$$s_{\text{av}} = [2^{k-1}(2^k - 1)]^{1/2} \tag{4}$$

is the geometric mean of the limits of the kth histogram bin.

The results for G_s are shown in Fig. 1 plotted on a log-log scale. In addition to the mean G_s evaluated over all the lattice realizations for each size L, the high and low extremes are also shown (except for the largest L); clearly there exists a broad range of cluster sizes for each L over which the spread in G_s is extremely narrow.

Linear least-square fits based on the logarithm of the functional form of the scaling prediction (3) are shown in Fig. 2 for each of the lattice sizes considered. The contributions arising from extreme values of s, where deviations from linearity were apparent in Fig. 1, have been eliminated, both for small s, where scaling breaks down, and for large s, where poor cluster statistics and finite lattice size dominate; the remaining points are seen to lie extremely close to the optimal linear fit. For the largest lattice,

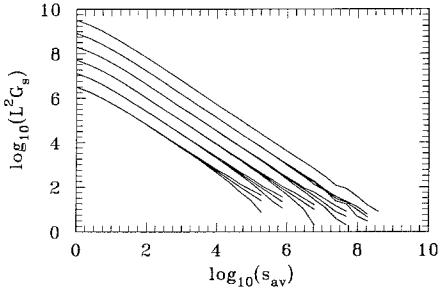

Fig. 1. Double logarithmic plot of the cluster size distribution. The lattice edge ranges from $L = 2 \times 10^4$ to $L = 6.4 \times 10^5$, with L doubling between lattices. The per-site normalization has been removed to separate the curves which would otherwise overlap (the higher curves are for larger L). In addition to the means over all lattice realizations, minimum and maximum cluster counts are also shown.

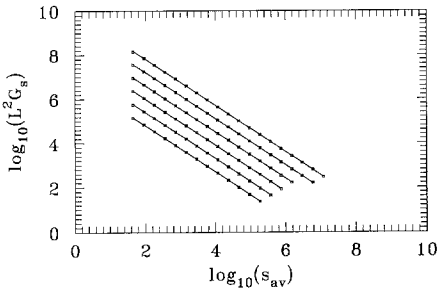

Fig. 2. Linear least-square fits to the data of Fig. 1 after removal of data points at the small- and large-s limits. The data points themselves are also shown.

the linear behavior (on a log-log scale) extends from approximately $L = 100$ to $L = 10^7$, a range of cluster sizes covering five orders of magnitude.

The gradients of the lines in Fig. 2 are estimates of $1 - \tau$; the values of τ deduced from the fits for the four largest lattice sizes lie between 2.0510 and 2.0516. If the contributions of the four smallest s values are removed from the fit calculations, the τ estimates are then found to lie between 2.0534 and 2.0558, with the value for the largest lattice being 2.0541. The second set of values differs from the expected exact result, $\tau = 187/91 = 2.0549...$, by less than 1 part per 1000, which is excellent by Monte Carlo standards. The succinct conclusion is that the evidence in favor of scaling remains strong.

REFERENCES

1. D. Stauffer, *Introduction to Percolation Theory* (Taylor and Francis, London, 1985); K. Binder and D. Stauffer, in *Applications of the Monte Carlo Method in Statistical Physics*, K. Binder, ed. (Springer, Berlin, 1984), p. 241.
2. A. Margolina, H. Nakanishi, D. Stauffer, and H. E. Stanley, *J. Phys.* **17**:1683 (1984).
3. D. C. Rapaport, *J. Phys. A* **19**:291 (1986).
4. M. F. Sykes and M. Glen, *J. Phys. A* **9**:87 (1976).
5. D. C. Rapaport, *Computer Phys. Rep.* **5**:265 (1987).
6. J. Hoshen and R. Kopelman, *Phys. Rev. B* **14**:3438 (1976).

Communicated by D. Stauffer