18.821 PROJECT 1: PERCOLATION

MATTHEW BEVERIDGE, BARIŞ EKIM, JUSTIN LIM

Abstract. We examine the process of percolation on graphs G, where each edge is open with probability p. For d-regular graphs, we show that an infinite cluster cannot exist if $p < \frac{1}{d-1}$. For the square lattice $\mathbb{Z}^2$, we derive upper and lower bounds on the value of p such that an infinite cluster can exist. Finally, we examine a different process called invasion and conjecture a relationship between invasion and percolation.

1. Introduction

Percolation, in the natural context, is the process of a liquid passing through a filter, group of rocks, coffee grounds, etc. This same idea is used to model the transmission of disease between living organisms, and other similar “spreading” events. In the mathematical context, it describes the behavior of connected clusters in a random graph and is a crude model with tenuous connection to reality. To model percolation on a graph Γ, we choose an initial probability p, and say that an edge is *open* (transmission occurs) with probability p and *closed* (transmission does not occur) with probability $1-p$. The structural properties of the obtained random subgraph and the set of open edges could then be investigated question such as: “Is there an infinite component of the graph Γ?”, “Does it depend on the probability p?”, “For which critical value p_* such that for $p < p_*$ the probability that an infinite cluster exists is 0?”, and so on.

In this paper, we formally provide a definition of the percolation problem in terms of existence of an infinite cluster, prove several properties of the obtained components in d-regular graphs and the integer square lattice $\mathbb{Z}^2$, investigate the closely related process of *Invasion*, and provide supplementary empirical results.

For the d-regular graph, it is shown that the probability an infinite cluster exists is 0 if $p < \frac{1}{d-1}$ and conjectured 1 otherwise (Section 2). For the integer square lattice $\mathbb{Z}^2$, if $p < \frac{1}{2}$ then the probability an infinite cluster exists is 0 (Section 3). Additionally, we propose that the infinite sequence formed during invasion approaches the uniform distribution on $[0, p^*]$ where p^* is the "critical point" at which $\mathbb{P}[\textit{infinite cluster exists}] = 0$ when $p < p^*$ (Section 4).

2. Percolation on d-regular trees

As introduced in [BH57], the percolation problem is studied on a graph, where a subgraph is obtained a probability on the edges:

Definition 2.1. Given a graph $G = (V, E)$, an edge $e \in E$ is independently *open* with probability p and *closed* with probability $1-p$. Given two vertices $u, v \in V$, v is *reached* from u if there exists a path between u and v consisting only of open edges.

Key words and phrases. percolation, regular graphs.

Definition 2.2. The *cluster* $\Gamma(v)$ of a vertex $v \in V$ is a subgraph of G consisting of the vertices reached from v.

We first consider the d-regular tree G_d, which is the unique graph such that every vertex has exactly d neighbors and there are no cycles.

Theorem 2.3. *If $p < \frac{1}{d-1}$, then for any $v \in G_d$, $\mathbb{P}[|\Gamma(v)| = \infty] = 0$.*

Proof. Without loss of generality, let v be the root of G_d, and the set of vertices $V_\ell \subseteq V$ such that the path from some vertex $u \in V_\ell$ to v has length ℓ. Since any path in a d-regular tree is unique, any $u \in V_\ell$ is reached from v if and only if each edge in the path is open. Hence, $\mathbb{P}[u \in \Gamma(v)] = p^\ell$. Since every vertex in G_d has $d-1$ children with the exception of the root v (which has d children), $|V_\ell| = d(d-1)^{\ell-1}$. By the union bound,

$$\begin{aligned}
\mathbb{P}[\text{at least one of } V_\ell \text{ is reached from } v] &\leq \sum_{u \in V_\ell} \mathbb{P}[u \text{ is reached from } v] \\
&= d(d-1)^{\ell-1} p^\ell \\
&= \frac{d}{d-1}((d-1)p)^\ell
\end{aligned}$$

Since $(d-1)p < 1$, $((d-1)p)^\ell \to 0$ as $l \to \infty$, and therefore the left hand side goes to zero as well. However, if $|\Gamma(v)| = \infty$, at least one of V_ℓ must be reached for any ℓ. It follows that $\mathbb{P}[|\Gamma(v) = \infty|] = 0$. $\square$

Lemma 2.4. *The expected size $\mathbb{E}[|\Gamma(v)|] = \infty$ if $p \geq \frac{1}{d-1}$, and $\frac{1}{d-1}\left(\frac{d}{1-(d-1)p} - 1\right)$ otherwise.*

This is shown empirically in Figure 1. Here, it is visible where the average size of the clusters jumps significantly once $p \geq \frac{1}{d-1}$ until it reaches the maximum cluster size at for each respective d at $p = 1$.

Proof. For vertices $u, v \in G_d$, let X_u be the indicator variable such that $X_{uv} = 1$ if $u \in \Gamma(v)$ and 0 otherwise. By linearity of expectation,

$$\begin{aligned}
\mathbb{E}[|\Gamma(v)|] &= \mathbb{E} \sum_{u \in G_d} [X_{uv}] = \sum_{u \in G_d} \mathbb{E}[X_{uv}] = \sum_{u \in G_d} \mathbb{P}[u \in \Gamma(v)] \\
&= 1 + \sum_{\ell=1}^{\infty} \sum_{u \in V_\ell} \mathbb{P}[u \in \Gamma(v)]
\end{aligned}$$

Furthermore, from Theorem 2.3 we have that

$$\begin{aligned}
\mathbb{E}[|\Gamma(v)|] &= 1 + \sum_{\ell=1}^{\infty} d(d-1)^{\ell-1} p^\ell \\
&= 1 + \frac{d}{d-1} \sum_{\ell=1}^{\infty} ((d-1)p)^\ell
\end{aligned}$$

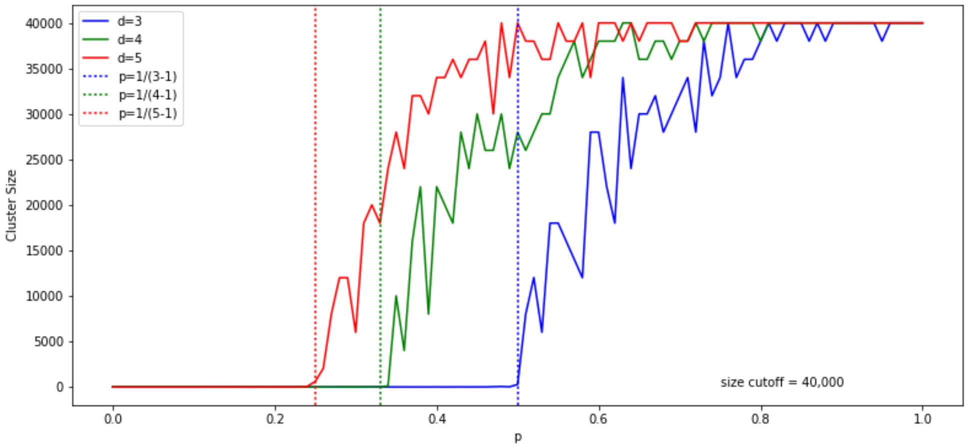

FIGURE 1. Average cluster size for 3, 4, and 5-regular graphs. The vertical lines correspond to the points where $p = \frac{1}{d-1}$. Note: the line for $p < \frac{1}{d-1}$ is not zero, it is simply small compared to the scale of the axis.

If $(d-1)p \geq 1$, clearly $\mathbb{E}[|\Gamma(v)|] = \infty$; otherwise if $(d-1)p < 1$, then

$$\begin{aligned}\mathbb{E}[|\Gamma(v)|] &= 1 + \frac{d}{d-1}\left(\frac{1}{1-(d-1)p} - 1\right) \\ &= \frac{1}{d-1}\left(\frac{d}{1-(d-1)p} - 1\right)\end{aligned}$$

□

Remark 2.5. Lemma 2.4 allows for a different proof of Theorem 2.3: if $p < \frac{1}{d-1}$, then since the expected size $\mathbb{E}[|\Gamma(v)|]$ is finite, $\mathbb{P}[|\Gamma(v)| = \infty] = 0$.

Theorem 2.6. *The probability that an infinite cluster exists is 0 if $p < \frac{1}{d-1}$ and 1 otherwise.*

Proof. Note that the union bound holds as long as the set of events in question is countable. Since G is countable and $\mathbb{P}[\Gamma(v) \text{ is infinite}] = 0$ by Theorem 2.3,

$$\mathbb{P}[\text{some cluster is infinite}] \leq \sum_{v \in G} \mathbb{P}[\Gamma(v) \text{ is infinite}] = 0$$

and therefore the probability on the left hand side must be 0.

In the case with $p \geq \frac{1}{d-1}$, the result that $\mathbb{P}[\text{infinite cluster exists}] = 1$ is conjectured from the empirical results in Figure 1.

□

3. PERCOLATION ON THE SQUARE LATTICE $\mathbb{Z}^2$

Consider the integer square lattice $\mathbb{Z}^2$, where each vertex (x, y) is connected to $(x \pm 1, y)$ and $(x, y \pm 1)$ by edges. Like in the previous section, we say that each edge is open with probability p and closed with probability $1 - p$, and denote by $\Gamma(x, y)$ the cluster containing (x, y).

Conjecture 3.1. If $p < \frac{1}{2}$, $\mathbb{P}[|\Gamma(0,0)| = \infty] = 0$. Consequently, the probability that an infinite cluster exists in the square lattice is also 0.

This is motivated by our simulations. Figure 2 shows the average size of $\Gamma(0,0)$ on a 200×200 lattice, as p varies between 0 and 1. The red line is $p = 0.5$. Note that the maximum size in the figure is 40000, which is the entire lattice. We conjecture that this is the the "switching point" at which $\Gamma(0,0)$ tends to have infinite size.

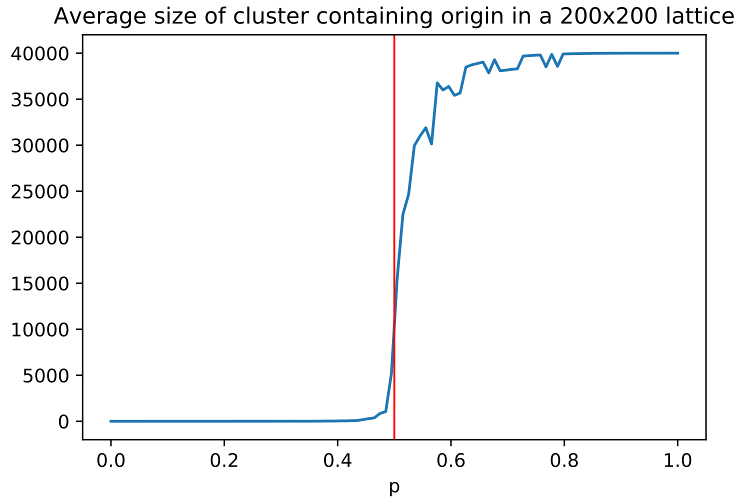

FIGURE 2. Average size of $\Gamma(0,0)$ on a square lattice of size 200×200, for 100 values of p equally spaced in $[0,1]$. Each value is taken as the average over 30 iterations.

Theorem 3.2. *If $p < 1/3$, the probability that an infinite cluster exists in the square lattice is 0.*

Proof. Let X_ℓ be the indicator variable for the event that a path of length ℓ starting from $(0,0)$ and consisting of only open edges exists. Such a path exists with probability p^ℓ. Starting from $(0,0)$, there are 4 vertices that can be picked in the first step, and at most 3 vertices for any other step; therefore, the number of such paths is at most $4(3^{\ell-1})p^\ell$. Since $p < 1/3$, $4(3^{\ell-1})p^\ell \to 0$ as $\ell \to \infty$, it follows that $\mathbb{P}[|\Gamma(0,0)| = \infty] = 0$. $\square$

Theorem 3.3. *Consider the "directed" square lattice $\mathbb{Z}^2_{dir}$, where each edge is directed "outwards" from the origin. Specifically, we can move from (x,y) to $(x \pm 1, y)$ and $(x, y \pm 1)$ only in the direction that increases $|x| + |y|$. For percolation on $\mathbb{Z}^2_{dir}$, if $p < \frac{1}{2}$, $\mathbb{P}[|\Gamma(0,0)| = \infty] = 0$.*

Proof. Consider $x, y \geq 0$. Since we can only move away from the origin, any path Q from the origin to (x,y) must have length exactly $x+y$. Furthermore, it is well-known that there are exactly $\binom{x+y}{x}$ such paths Q. By the union bound,

$$\begin{aligned}\mathbb{P}[(x,y) \text{ reached}] &\leq \sum_{\text{paths } Q} \mathbb{P}[Q \text{ is open}] \\ &= \binom{x+y}{x} p^{x+y}\end{aligned}$$

Let $L_n = \{(x, y) \in \mathbb{Z}^2 \mid x + y = n \text{ and } x, y \geq 0\}$, as shown in Figure 3. Using the union bound again,

$$\mathbb{P}[L_n \text{ reached}] \leq \sum_{x+y=n \text{ and } x,y>0} \binom{x+y}{x} p^{x+y}$$
$$= \sum_{x+y=n \text{ and } x,y>0} \binom{n}{x} p^n = (2p)^n$$

using the well-known identity $\sum_{x=0}^{n} \binom{n}{x} = 2^n$. Next, let L_n^1, L_n^2, L_n^3 be $L_n = L_n^0$ successively rotated 90 degrees clockwise around the origin. By symmetry, $P(L_n \text{ reached})$ is the same as $P(L_n^i \text{ reached})$, for each $i = 1, 2, 3$. Let $S_n = L_n \cup L_n^1 \cup L_n^2 \cup L_n^3$, as shown in Figure 4.

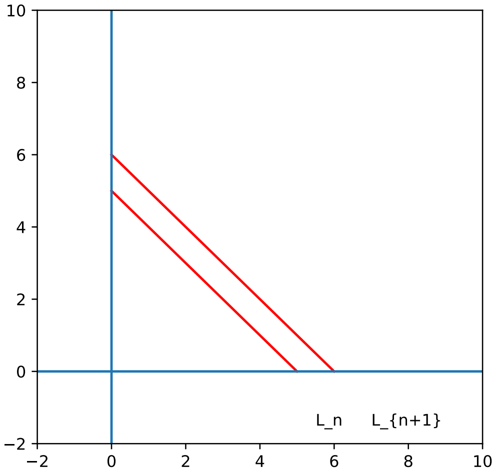

FIGURE 3. Illustration of L_n and L_{n+1}.

Therefore

$$P(S_n \text{ reached}) = \sum_{i=0}^{3} P(L_n^i \text{ reached}) \leq 4(2p)^n$$

Finally, since $2p < 1$, we find that $P(S_n \text{ reached})$ must go to zero as $n \to \infty$. However, if $\Gamma(0, 0)$ is infinite, then S_n must be reached for all n. It follows that the probability that $\Gamma(0, 0)$ is infinite is zero as well, as desired. □

We now show that the probability that an infinite cluster exists is nonzero for sufficiently large p. For this proof, we need an additional definition which will be useful:

Definition 3.4. The *dual graph* $\mathbb{Z}_*^2$ of $\mathbb{Z}^2$ consists of the midpoint of 4 adjacent lattice points, where a midpoint is *open* if any of the four surrounding edges are open and *closed* if all four surrounding edges are closed.

Note that a cycle with closed edges in $\mathbb{Z}_*^2$ limits the size of any cluster in $\mathbb{Z}_2$:

Remark 3.5. $|\Gamma| < \infty$ if there exists a cycle consisting of closed edges in $\mathbb{Z}_*^2$ surrounding (0, 0).

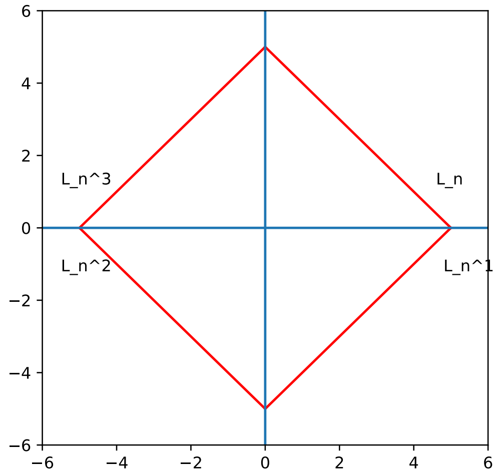

FIGURE 4. Illustration of S_n.

Using this remark, we can prove the following theorem:

Theorem 3.6. *For* $p > 2/3$, $\mathbb{P}[|\Gamma| = \infty] > 0$.

Proof. By Theorem 3.2, we know that there are $4(3^{\ell-1})$ paths of length ℓ on $\mathbb{Z}^2$. Similarly, there are $4(3^{\ell-1})$ paths of length ℓ on $\mathbb{Z}^2_*$ consisting of ℓ midpoints, and at most ℓ midpoints could result in a cycle. Therefore, there are $\ell \cdot 4(3^{\ell-1})$ cycles in $\mathbb{Z}^2_*$, and the probability of having a cycle consisting of closed edges in $\mathbb{Z}^2_*$ is $\ell \cdot 4(3^{\ell-1})(1-p)^\ell$. If $p > 2/3$, $\ell \cdot 4(3^{\ell-1})(1-p)^\ell \to 0$ as $\ell \to \infty$, and it follows that $\mathbb{P}[|\Gamma| = \infty] > 0$ for $p > 2/3$. □

4. INVASION

Consider the following process on a graph G, which we call an *invasion*. Assign independently to each edge a uniform random "probability" $c_e \in [0, 1]$. We form an infinite sequence of subgraphs $\Gamma_1, \Gamma_2, \ldots$ as follows. Γ_1 contains only a single vertex $\{v\}$, which is a fixed starting point. At step k, consider all the edges which do not lie in Γ_k, but are adjacent to it. Among these edges, take the one with the lowest probability, which we denote by c_k, and add that edge (with its other endpoint) to our graph, thereby forming Γ_{k+1}. Of interest to us will be the sequence of chosen probabilities $c_1, c_2, \ldots$ - what does it look like?

At first glance, it is not clear that the processes of percolation and invasion are related. In percolation, the edge "weights" are Bernoulli random variables, while in invasion, they are uniform random variables. In this section, we present some empirical results and conjecture a relationship between the processes of invasion and percolation on the same graph G.

Conjecture 4.1. Suppose p^* is the "critical point" for percolation on graph G. Specifically, if $p < p^*$, then P(infinite cluster exists) is zero, and 1 if $p > p^*$. For an invasion on G, the infinite sequence $c_1, c_2, \ldots$ approaches the uniform distribution on $[0, p^*]$.

Example 4.2. Figure 5 show the distribution of c_k for 1000 invasions on G_d, for $d = 4, 5, 6$, each of length 500. The red line in each case is at $p^* = \frac{1}{d-1}$, which is the conjectured switching points for d-regular graphs that we discussed earlier. Figure 6 shows the same plot for invasion on a square lattice, where we believe $p^* = \frac{1}{2}$. In each case, the c_k seem to be approximately uniformly distributed in $[0, p^*]$.

References

[BH57] Simon R Broadbent and John M Hammersley. Percolation processes: I. crystals and mazes. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 53, pages 629–641. Cambridge University Press, 1957.

Massachusetts Institute of Technology, Department of Mathematics, Cambridge MA 02139

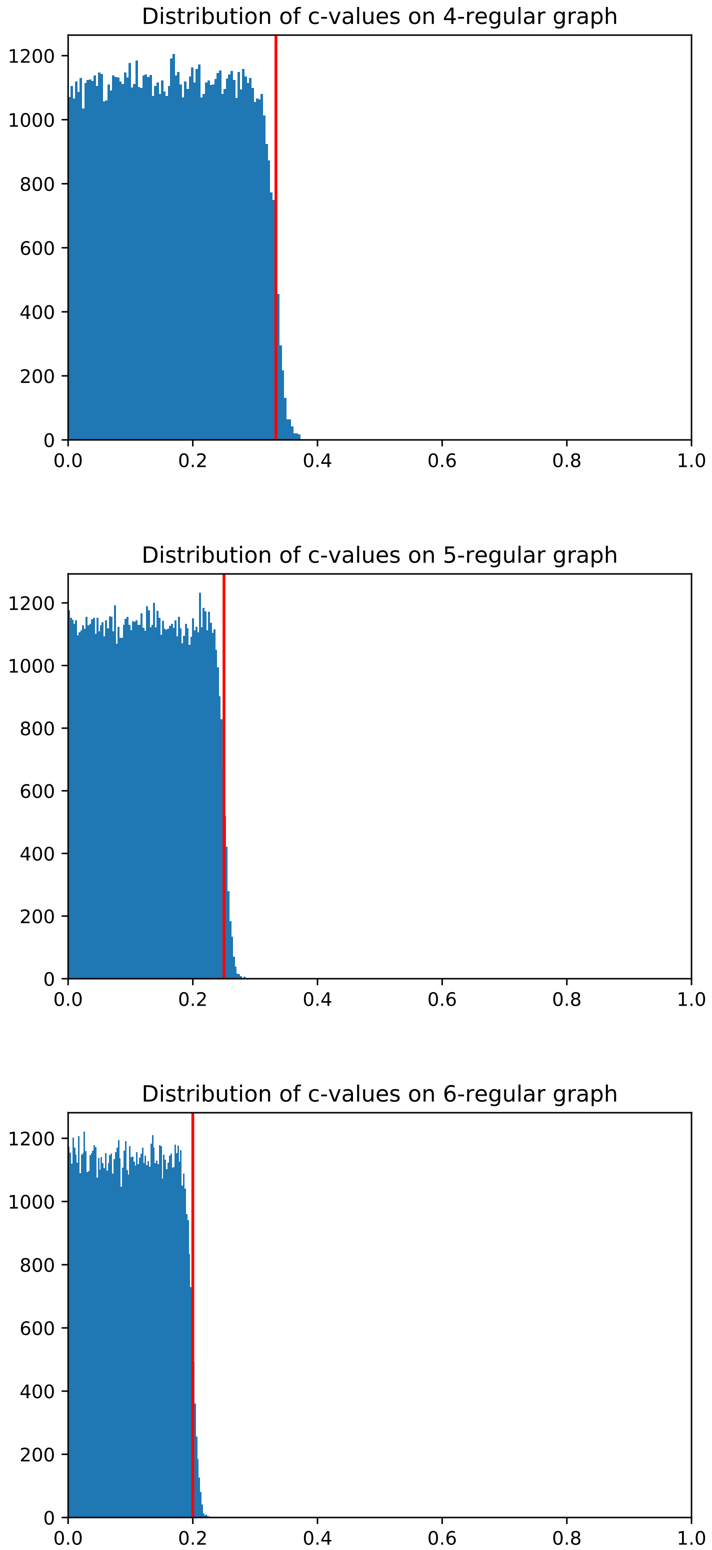

FIGURE 5. Distribution of c_k for invasion on 4, 5, and 6-regular graphs. Red lines are at $c = \frac{1}{3}, \frac{1}{4}, \frac{1}{5}$ respectively.

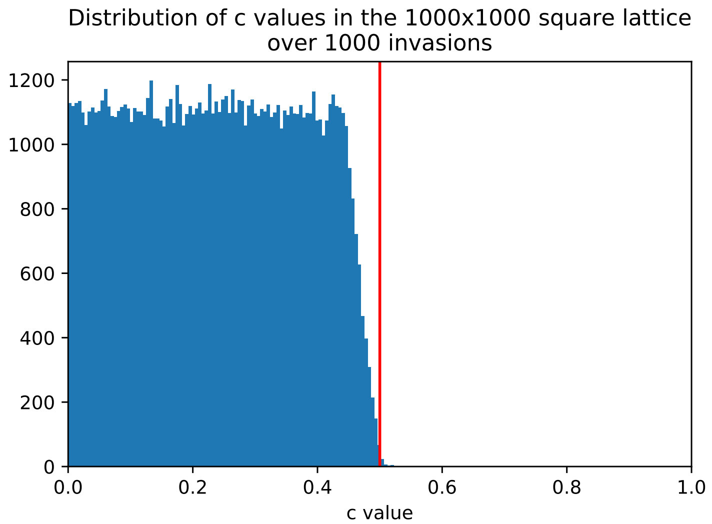

FIGURE 6. Distribution of c_k for invasion on the square lattice. The red line is at $c = \frac{1}{2}$.