A Brief Introduction to Percolation Theory

Joshua Mankelow
1902187

April 18, 2021

Abstract

Consider a cube of water-permeable material. What is the probability that that if water is poured on top of the cube it may drain all the way through the cube and out the opposite face? Initially developed by Paul Flory and Walter Stockmayer in 1944, percolation theory attempts to answer such questions by rephrasing them in terms of vertices (sites) and edges (bonds) of graphs and examining the connectedness of such graphs. The connectedness of these graphs—in the infinite case—is determined by a threshold probability, p_c, describing whether the water may pass through each site or bond. This essay will introduce the ideas of site and bond percolation as well as the notion of clusters and critical (threshold) probabilities. We will also analyse the one dimensional case to garner a basic understanding before exploring higher dimensional cases. After discussing the concepts of percolation theory, we will move on and look at the many applications of the theory discussed in the earlier parts of the essay.

Contents

1 Introduction

Let us consider the example from the abstract of water filtering through a porus medium, but this time in two dimensions. How do we model this? One might imagine that the medium consists of many particles arranged (for simplicity) in an $n \times n$ square lattice and linked to each of their nearest neighbours. Clearly, this is the lattice on $\mathbb{Z}^2$. To set up the problem, each of the particles will be expressed as a vertex in a graph and each of the links will be an edge connecting two directly adjacent vertices. In the context of percolation theory, a vertex is called a site and an edge is called a bond; these sites and bonds form a graph which we refer to as a network.

So what does percolation actually mean? In order to talk about percolation, we need to think about what makes a percolation problem. If we think about the context we're in, some of the sites in the network will alow the water through and some of them won't. The sites that allow water to pass through are labeled **open** and the sites that don't allow water to pass through are labeled **closed**. This gives us **site** percolation. If we were to consider open and closed bonds instead of sites, then we would have **bond** percolation. We also let these sites or bonds be open with probability p and closed with probability $1 - p$. I will refer to this p as the **percolation probability**.

Now that we've defined site and bond percolation, what's the problem that we're trying to solve? In the case of water being poured on a porus medium, we would like to know whether there is a route that the water could take from the top of the medium to the bottom. This is called an **open path**. We shall model this using site percolation (in fact, all examples in this essay will be using site percolation unless explicitly stated otherwise).

Definition 1.1. Let $N = (V, E)$ be a network, we say that a path in N is **open** if every site in the path is open. [1]

Definition 1.2. Let $N = (V, E)$ be a network and let $A, B \in V$. The sites A, B are **openly connected** if there exists an open path connecting A and B. Throughout this essay, I will interchange the term **openly connected** with **connected**.

Definition 1.3. Let $N = (V, E)$ be a network and let $A, B \in V$. The sites A, B are **openly disconnected** if there does not exist an open path connecting A and B. Throughout this essay, I will intercahnge the term **openly disconnected** with **disconnected**.

Returning to the example, the probability that an open path from the top of the network to the bottom exists depends on both our choices of both p and n. As a result of our context, our value for n should be large—this

[1] This definition is trivially different for bond percolation.

is the case with most percolation models—but we shall use small n for the sake of example and simplicity. Let us now fix n and see what happens as we vary p. Obviously we have two trivial cases, $p = 0$ and $p = 1$, where the network is completely disconnected and completely connected respectively. What about when $p \in (0, 1)$? Let's inspect three different values of p on our network: $p = 0.25$, $p = 0.5$ and $p = 0.75$ as shown in figures 1a, 1b and 1c on page 5.

As one might expect, as p increases, so does the average size of a cluster of open sites. Also observe that the overall connectedness of the network increases as p increases. I.e. the probability of having an open path from the top of the netwoork to the bottom increases with p. This is the result that we expect given the context — as we reduce the number of things blocking the way for the water, it's easier for it to pass through the block of our chosen medium. Now let us consider two sites A and B in our example network. As we increase p, it's obvious that the probability of these two sites A and B being connected will increse. The question is: What's the relationship between p and the probability of A and B being connected? It turns out that there's a probability, $p*$, such that

- for $p < p*$, the network is mostly disconnected so the probability of A and B being connected is close to 0.
- for $p > p*$, the network is mostly connected so the probability of A and B being connected is close to 1.

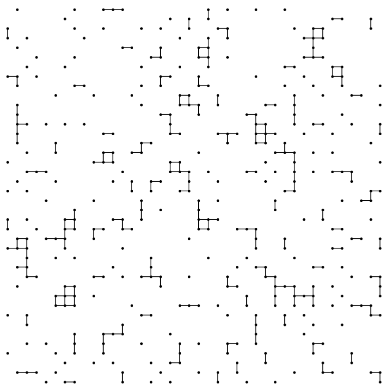
The idea of percolation theory is to understand where this transition from mostly disconnected to mostly connected occurs and what value of p gives rise to such a change in structure. This value of p is usually labelled p_c and is referred to as the **critical probability** or **critical threshold**. More formally, this **critical probability** is the probability at which, when considering an infinite network, a cluster of infinite size is guaranteed to exist.

1.1 Other network configurations

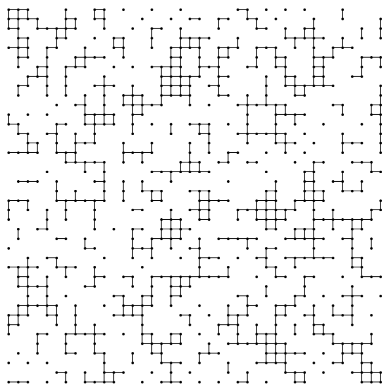
We have already seen the lattice on $\mathbb{Z}^2$ as an example, but there are many more. To remain within the scope of this essay, we shall only breifly mention some two and three dimensional examples and print their site and bond critical probabilities and a diagram.

Definition 1.4. A network is considered **regular** if every site in that network has the same number of bonds attached to it.

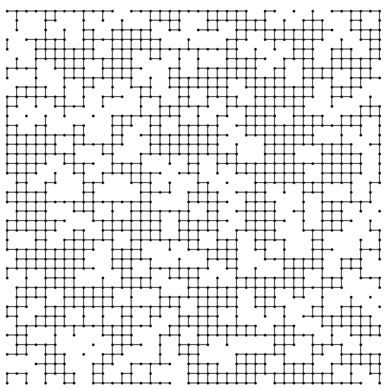
Definition 1.5. The **coordination number** of a regular network is the number of bonds attached at every site. This quantity is denoted using the letter Z. I.e. the lattice on $\mathbb{Z}^2$ has a coordination number $Z = 4$.



(a) $p = 0.25$



(b) $p = 0.5$



(c) $p = 0.75$

Figure 1: Examples of bond percolation for $p \in (0, 1)$ on a 40×40 network where if a site is open in the model it appears present in the diagram and bonds are only present if they connect two open sites.

Two dimensional network configurations

Clearly, one two dimensional network configuration is the lattice on $\mathbb{Z}^2$. In context this is referred to as the square lattice. Other regular two dimensional network configurations include, but are not limited to, the Bethe Lattice (Figure 4a), Honeycomb Lattice (Figure 4b), Kagome Lattice (Figure 4c) and the Triangular Lattice (Figure 4d). As one might imagine, each of these configurations has a different (but not necessarily distinct) critical probability. Below is a table showing the critical probabilities for each of the aforementioned network configurations. It should be noted that probabilities marked with a * (star) are exact results.

Configuration	Z	p_c for bond percolation	p_c for site percolation
Bethe	3	TO FIND	TO FIND
Honeycomb	3	$1 - 2\sin(\pi/18)$*	0.6962
$\mathbb{Z}^2$ (Square)	4	1/2*	0.5927
Kagome	4	0.522	0.652
Triangular	6	$2\sin(\pi/18)$*	1/2*

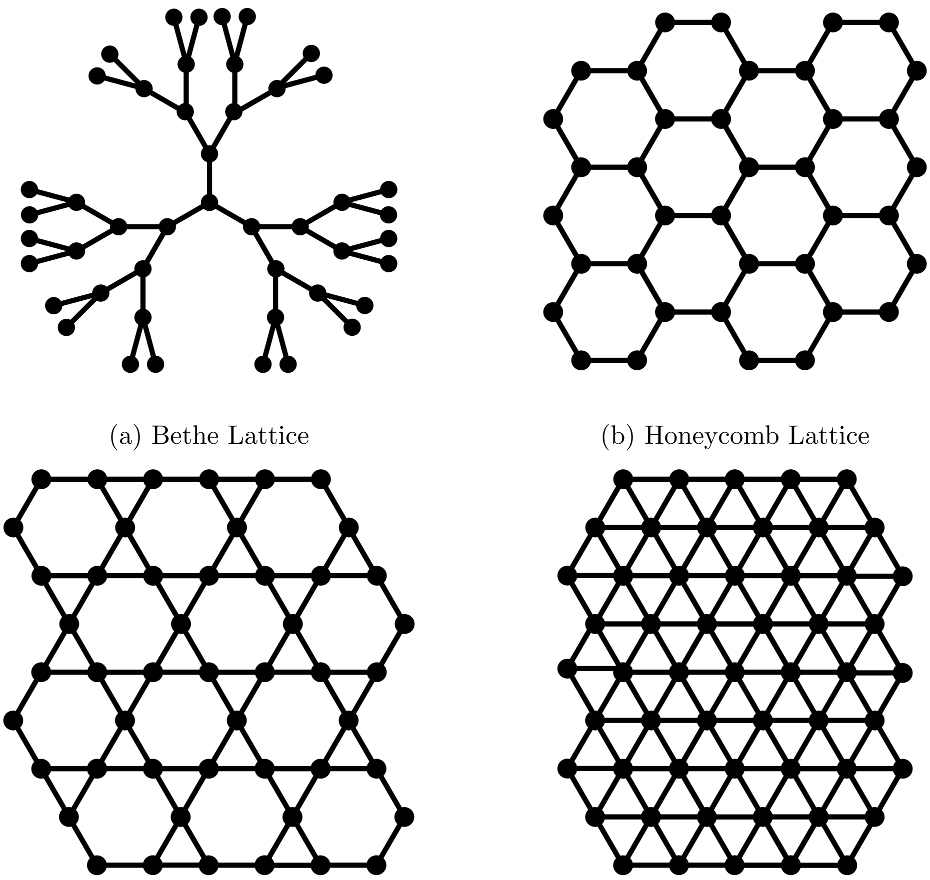
Figure 2: Critical probabilities for various configurations of two dimensional networks[Sahimi, 1994, p. 11]

Configuration	Z	p_c for site percolation	p_c for bond percolation
Diamond	4	0.3886	0.4299
Simple Cubic	6	0.2488	0.3116
BCC	8	0.1795	0.2464
FCC	12	0.198	0.119

Figure 3: Critical probabilities for various configurations of three dimensional networks[Sahimi, 1994, p. 11]

Three dimensional network configurations

It shouldn't be hard to guess that the lattice on $\mathbb{Z}^3$ is a potential configuration for three dimensional networks. We call this configuration the Simple Cubic Lattice (Figure 5b). Similarly to the two dimensional case, there many other regular three dimensional network configurations. These include, but again are not limited to, the Diamond Lattice (Figure 5a), the Body Centered Cubic (BCC) Lattice (Figure 5c) and the Face Centered Cubic (FCC) Lattice (Figure 5d). Notice how none of these results are precise.



(a) Bethe Lattice (b) Honeycomb Lattice

(c) Kagome Lattice (d) Triangular Lattice

Figure 4: Two dimensional regular network configurations

(a) Diamond Lattice

(b) Simple Cubic Lattice

(c) Body Centered Cubic Lattice

(d) Face Centered Cubic Lattice

Figure 5: Three dimensional regular network configurations

2 The one dimensional case

To develop an understanding of how to analyse these networks, we shall consider the one dimensional case where our network is the lattice on $\mathbb{Z}$, or a "chain".

Theorem 2.1. *The critical probability, p_c, for the lattice on $\mathbb{Z}$ when considering site percolation is $p_c = 1$.*

If we recall the definition of the critical probability, the value of p_c that we're looking for is intuitively 1. In order to prove this theorem, we shall introduce some more machinery. This machinery isn't necessary for the proof, but it helps with understanding and will allow us to analyse more interesting cases later on.

Definition 2.1. If the percolation probability on the chain is p, we define the **number of s-clusters per site** by the following quantity:

$$n_s = p^s(1-p)^2$$

This quantity represents the probability that any given site in the network is the left end of an s-cluster. So, if our network is of length $L >> s$, then we will have $Ln_s = Lp^s(1-p)^2$ clusters of size s on average. This quantity also allows us to explore the probability that any given site is part of an s-cluster. Such a probability is given by the quantity $n_s s$.

The idea of percolation is to understand whether a path exists from one side of the network to the other, so we have to use a slightly different definition for the one dimensional case. This new definition fixes an issue that I'll address after the proof.

Definition 2.2. Consider the lattice on $\mathbb{Z}$ of the form $N = (V, E)$. The critical probability, denoted p_c, is the percolation probability such that the probability that there **uniquely** exists a cluster, $C \subseteq V$, with $|C| = \infty$ is 1.

This proof and the following corollory are heavily inspired by the proof and subsequent corollory from Dietrich Stauffer's *Introduction to Percolation Theory.* [Dietrich Stauffer, 1991]

Proof. (Theorem 2.1) We shall prove by contradiction. Let us assume that $p_c \in [0, 1)$ and is fixed. This implies that a chain of length L will have, on average, $L(1-p)$ closed sites. As $L \to \infty$, $L(1-p) \to \infty$ showing us that there is at least one closed site in the chain and that means there is no continuous row of occupied sites. Thus $p_c = 1$ in order to have only one infinite cluster. $\square$

So why doesn't this proof work if we hadn't made that ammendment to the definition. Notice that $\mathbb{Z}$ is a countably infinite set and all of the $L(1-p)$

closed sites form a subset of $\mathbb{Z}$. This means that the set of all closed sites is also countably infinite. This situation may be rephrased in a way such that we must partition a countably infinite set into countably infinite subsets. prove this?

The above results allow us to get some more interesting information about the behaviour of our system. For example, we can deduce the following:

Corollary. *When considering the lattice on $\mathbb{Z}$ with percolation probability $p \in [0, 1)$, the following equality holds:*

$$\sum_s n_s s = p$$

This result comes from the fact that every open site must belong to a cluster of some size s. So summing $n_s s$ over all s must give us p. This equality may also be derived using the definition of n_s and the formula for a geometric series.

Proof.

$$\begin{aligned}
\sum_s n_s s &= \sum_s p^s (1-p)^2 s \\
&= (1-p)^2 \sum_s p \frac{d(p^s)}{dp} \\
&= (1-p)^2 p \frac{d(\sum_s p^s)}{dp} \\
&= (1-p)^2 p \frac{d(p/(1-p))}{dp} \\
&= p
\end{aligned}$$

$\square$

It's worth noting that this equality doesn't hold for $p = 1$, because $n_s = 1^s(1-1)^2 = 0$ so $\sum_s n_s s = 0$. The technique of considering the sizes of clusters and the number of empty sites surrounding them is also used when analysing more compelx structures. Say, for example, you have a cluster, C, of size $|C| = 9$ on the lattice on $\mathbb{Z}^2$. This cluster could take many different shapes with different numbers of empty sites surrounding it. For example, this cluster could be a straight line of open sites which has a total of 22 closed sites surrounding it. Therefore the probability that a cluster like this exists at any site given a percolation probability, p, is $p^9(1-p)^{20}$. This cluster could also be a square of open sites with side length 3. This would have 12 closed sites surrounding it and thus the probability of a cluster of this shape existing at any given site is $p^9(1-p)^{12}$. The ability for clusters to be the same size but have different "perimeters" is what makes analysing these problems in higher dimensions so difficult.

3 Applications

References

[Dietrich Stauffer, 1991] Dietrich Stauffer, A. A. (1991). *Introduction to Percolation Theory*. Taylor & Francis, 11 New Fetter Lane, London EC4P 4EE.

[Sahimi, 1994] Sahimi, M. (1994). *Applications of Percolation Theory*. Taylor & Francis, University of Southern California.