Discrete Applied Mathematics 222 (2017) 124–142

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Algebraic bounds for heterogeneous site percolation on directed and undirected graphs

Kathleen E. Hamilton [a,∗,1], Leonid P. Pryadko [b]

[a] Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, TN, 37821, USA

[b] Department of Physics & Astronomy, University of California, Riverside, CA, 92521, USA

ARTICLE INFO

Article history:
Received 19 April 2016
Received in revised form 22 December 2016
Accepted 29 December 2016
Available online 12 February 2017

Keywords:
Percolation
Non-backtracking spectra
Vertex connectivity
Infinite graph
Digraph

ABSTRACT

We analyze site percolation on directed and undirected graphs with site-dependent open-site probabilities. We construct upper bounds on cluster susceptibilities, vertex connectivity functions, and the expected number of simple open cycles through a chosen arc; separate bounds are given on finite and infinite (di)graphs. These produce lower bounds for percolation and uniqueness transitions in infinite (di)graphs, and for the formation of a giant component in finite (di)graphs. The bounds are formulated in terms of appropriately weighted adjacency and non-backtracking (Hashimoto) matrices. It turns out to be the uniqueness criterion that is most closely associated with an asymptotically vanishing probability of forming a giant strongly-connected component on a large finite (di)graph.

1. Introduction

We are currently living in an age where many scientific and industrial applications rapidly generate large datasets. The connectivity and underlying structure of this data is of great interest. As a result, graph theory has enjoyed a resurgence, becoming a prominent tool for describing complex connections in various kinds of networks: social, biological, technological [2,1,13,20,19,61,64,65], and many others. Percolation on graphs has been used to describe internet stability [16,14], spread of contagious diseases [30,57,68] and computer viruses [62]; related models describe market crashes [28] and viral spread in social networks [72,44,40]. General percolation theory methods are increasingly used in quantum information theory [23,48,36,45,18]. Percolation is also an important phase transition in its own right [43,26,25,73] and is well established in physics as an approach for dealing with strong disorder: quantum or classical transport [4,46,39], bulk properties of composite materials [9,58], diluted magnetic transitions [70], or spin glass transitions [21,60,22,17,69].

Recently, we suggested [31] a lower bound on the site percolation transition on an infinite graph,

$$p_c \geq 1/\rho(H). \quad (1)$$

Here, $\rho(H)$ is the spectral radius of the non-backtracking (Hashimoto) matrix [35] H associated with the graph. This expression has been proved [31] for infinite *quasi-transitive* graphs, a graph-theoretic analog of translationally-invariant systems with a finite number of inequivalent sites. The bound (1) is achieved on any infinite quasi-transitive tree [31], and

∗ Corresponding author.
E-mail addresses: hamiltonke@ornl.gov (K.E. Hamilton), leonid.pryadko@ucr.edu (L.P. Pryadko).

1 This submission was written by the author acting in her own independent capacity and not on behalf of UT-Battelle, LLC, or its affiliates or successors.

http://dx.doi.org/10.1016/j.dam.2016.12.027

it also gives numerically exact percolation thresholds for several families of large random graphs, as well as for some large empirical graphs [41].

In most applications of percolation theory, one encounters large, but finite, graphs. The expectation is that the corresponding crossover retains some properties of the transition in the infinite graphs, e.g., the formation of large open clusters be unlikely if the open site probability p is well below p_c. However, the bound (1) tells nothing about the structure of the percolating cluster on finite graphs, and neither it gives an algorithm for computing the location of the crossover in the case of a finite graph [64]. In particular, Eq. (1) misses the mark entirely for any finite tree where $\rho(H) = 0$.

In this work, we construct several spectral and algebraic bounds for transitions associated with heterogeneous site percolation on directed and undirected graphs, both finite and infinite, and analyze the continuity of these bounds for a sequence of finite digraphs weakly convergent to an infinite graph. Namely, for finite digraphs, we construct explicit upper bounds for the local in-/out-/strong-cluster susceptibilities (average size of a cluster connected to a given site), the strong connectivity function (probability that a given pair of sites belongs to the same strongly-connected cluster), and the expected number of simple cycles passing through a given arc. We also construct some analogous bounds for infinite digraphs, which result in non-trivial lower bounds for the transitions associated with divergent in-/out-cluster susceptibilities, emergence of infinite in-/out-clusters, and the strong-cluster uniqueness transition.

Our results imply that Eq. (1) and its analogue for heterogeneous site percolation on a general digraph give a universal bound for the strong-cluster *uniqueness* transition, below which a strongly connected infinite cluster cannot be unique. Such a bound is continuous for an increasing sequence of subgraphs if the percolation problem on the limiting digraph has a finite *minimum return probability*, the probability that any arc and its inverse are connected by an open non-backtracking path. Finite minimum return probability also guarantees that below this bound, the strong connectivity decays exponentially with the distance, and the expected size of a strongly connected cluster scales sublinearly with the number of vertices in a digraph. In comparison, the bound (1) applies only conditionally to the percolation transition proper, e.g., for a weakly-convergent sequence of quasi-transitive digraphs of increasing size, where the number of inequivalent vertex classes remains uniformly bounded.

The remainder of this paper is organized in four sections. In Section 2 we define several matrices associated with heterogeneous site percolation and introduce other notations. Our main results are given in Section 3 which contains bounds for finite digraphs, and in Section 4 where infinite digraphs are discussed. Finally, in Section 5 we compare effectiveness of different criteria in limiting the emergence of a *giant component*, an open cluster which contains a finite fraction of all vertices in the digraph.

2. Definitions and notations

We consider only simple directed and undirected graphs with no loops or multiple edges. A general digraph $\mathcal{D} = (\mathcal{V}, \mathcal{E})$ is specified by its sets of vertices (also called sites) $\mathcal{V} \equiv \mathcal{V}(\mathcal{D})$ and edges $\mathcal{E} \equiv \mathcal{E}(\mathcal{D})$. Each edge (bond) is a pair of vertices, $(u, v) \subseteq \mathcal{E}$ which can be directed, $u \to v$, or undirected, $u \leftrightarrow v$. A directed edge $u \to v$ is also called an arc from u to v; an undirected (symmetric) edge can be represented as a pair of mutually inverted arcs, $u \leftrightarrow v \equiv \{u \to v, v \to u\}$. A digraph with no undirected edges is an oriented graph. We will denote the set of arcs in a (di)graph $\mathcal{D}$ as $\mathcal{A} \equiv \mathcal{A}(\mathcal{D})$. Each vertex $v \in \mathcal{V}$ in a digraph $\mathcal{D}$ is characterized by its in-degree $\mathrm{id}(v)$ and out-degree $\mathrm{od}(v)$, the number of arcs in $\mathcal{A}(\mathcal{D})$ to and from v, respectively. A digraph with no directed edges is an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. For every vertex in an undirected graph, the degree is the number of bonds that include v, $\deg(v) = \mathrm{id}(v) = \mathrm{od}(v)$.

We say that vertex u is connected to vertex v on a digraph $\mathcal{D}$, if there is a path from $u = u_0$ to $v \equiv u_\ell$,

$$\mathcal{P} \equiv \{u_0 \to u_1, u_1 \to u_2, \ldots, u_{\ell-1} \to u_\ell\} \subseteq \mathcal{A}(\mathcal{D}). \tag{2}$$

The path is called non-backtracking if $u_{i-1} \neq u_{i+1}$, $0 < i < \ell$, and self-avoiding (simple) if $u_i \neq u_j$ for $0 \leq i, j \leq \ell$. The length of the path is the number of arcs in the set, $\ell = |\mathcal{P}|$. The distance from u to v on $\mathcal{D}$, $d(u, v)$, is the minimum length of a path from u to v. We call path (2) open if $u_0 \neq u_\ell$, and closed otherwise. A closed path is a cycle; it can be non-backtracking or self-avoiding (simple). Connectivity on an undirected graph is a symmetric relation: we just say that vertices u and v are connected (or not). On a digraph, we say that vertices u and v are strongly connected iff u is connected to v and v is connected to u; u and v are weakly connected on $\mathcal{D}$ if they are connected on the undirected graph underlying $\mathcal{D}$. A *ray* is a semi-infinite simple path, characterized as in- or out-going according to the directionality of the constituent arcs. A *strong ray* is a strongly connected union of in- and out-going rays; it has the property that the intersection between the vertex sets is an infinite set.

A digraph $\mathcal{D}$ is called *transitive* iff for any two vertices u, v in $\mathcal{V} \equiv \mathcal{V}(\mathcal{D})$ there is an automorphism (symmetry) of $\mathcal{D}$ mapping u onto v. Digraph $\mathcal{D}$ is called *quasi-transitive* if there is a finite set of vertices $\mathcal{V}_0 \subset \mathcal{V}$ such that any $u \in \mathcal{V}$ is taken into $\mathcal{V}_0$ by some automorphism of $\mathcal{D}$. We say that any vertex which can be mapped onto a vertex $u_0 \in \mathcal{V}_0$ is in the equivalence class of u_0. The square lattice is an example of a transitive graph; a two-dimensional lattice with r inequivalent vertex classes defines a (planar) quasi-transitive graph.

A graph $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ is called a covering graph of $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ if is there is a function $f : \mathcal{V}' \to \mathcal{V}$, such that an edge $(u', v') \in \mathcal{E}'$ is mapped to the edge $(f(u'), f(v')) \in \mathcal{E}$, with an additional property that f be invertible in the vicinity of each vertex, i.e., for a given vertex $u' \in \mathcal{V}'$ and an edge $(f(u'), v) \in \mathcal{E}$, there must be a unique edge $(u', v') \in \mathcal{E}'$ such that $f(v') = v$. The *universal cover* $\widetilde{\mathcal{G}}$ of a connected graph $\mathcal{G}$ is a connected covering graph which has no cycles (a tree); it is