Physics Letters A 381 (2017) 2578–2582

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

A relatively simple model for percolation properties of real networks

Tao Fu [a], Liling Zou [a], Chenguang Li [b,*], Junbo Zhao [c]

[a] Experimental School, Beijing University of Technology, Luyuannandajie 89, Beijing 101101, China
[b] Economics and Management School, North China University of Technology, Jinyuanzhuanglu 5, Beijing 100144, China [1]
[c] Economics and Management School, Beijing University of Technology, Pingleyuan 100, Beijing 100124, China

ARTICLE INFO

Article history:
Received 14 February 2017
Received in revised form 2 June 2017
Accepted 2 June 2017
Available online 7 June 2017
Communicated by F. Porcelli

Keywords:
Real network
Percolation
Generating function
Phase transition

ABSTRACT

Analyzing percolation rules of real networks has some great realistic significance. In this paper, we develop a relatively simple model based on generating function method to study percolation properties of real networks. We construct our model for both site and bond percolation, compare its estimates with those of the message passing algorithm and simulation results on computer-generated networks as well as practical networks, and discuss causes of the inaccuracy. The conclusions show that the accuracy of our model could be accepted though it is lower than that of the message passing algorithm and the discrepancies between the estimates of our model and the simulation values mainly come from the disagreement of those real networks with the model hypotheses.

1. Introduction

Percolation often used to model the spread of contagious diseases [1,2] and Internet resilience [3,4], is one of the most studied processes in statistical physics. It discusses the total connectivity of lattices or networks under the same site or bond occupation probability, so it can be divided into site percolation and bond percolation. The probability generating function [5] is a powerful tool in this field to estimate some percolation properties, and based on it, Newman et al. [6] derived exact expressions of the mean size of non-percolating components, the size of the giant component, and the position of the phase transition for random networks with arbitrary degree distributions while Leicht and D'Souza [7] calculate those properties for interacting networks composed of multiple subnetworks. Those models above took the degree distribution of networks (or interacting networks) as the input, and in the strict sense, their estimations correspond to a series of networks (or multiple networks) with a concrete degree (or multi-degree) distribution. Meanwhile, they didn't discuss the occupation probability too much. In recent years, the research interest in this field has turned to analyses of some concrete real networks. Karrer et al. [8] reformulated bond percolation on sparse, locally tree-like networks as a message passing process. Based on the generating function method, they set up self-consistent equations, which take bond occupation probability as the input, and by solving those equations, they predicted the giant component size as well as the average size of non-percolating components for those concrete networks. Meanwhile, they also estimated the critical percolation probability of such a network by taking the reciprocal of the leading eigenvalue of the non-backtracking matrix. Hamilton and Pryadko [9] derived similar result for the site percolation threshold on any arbitrary quasitransitive tree and quasi-regular graph. Radicchi [10] used the message passing algorithm proposed by Karrer et al. [8] to study the bond and site percolation in real interdependent networks. Then Radicchi and Castellano [11] tried to overcome the main limitation of that message passing algorithm by excluding redundant paths caused by triangles, and their modification improved the accuracy of that algorithm more effectively in site percolation than bond percolation.

Though the message passing algorithm contributes a lot to studying percolation rules of real networks, its calculation process is somewhat complex and demanding. In this paper, we develop a relatively simple percolation model for real networks by adding occupation probability on the model of Newman et al. [6]. We construct our model for both site and bond percolation, then compare estimates of the message passing algorithm and our model with simulation results on four concrete networks, and at last, we discuss causes of the inaccuracy of both the two models.

* Corresponding author.
E-mail address: lichenguang@emails.bjut.edu.cn (C. Li).
[1] Present address.

http://dx.doi.org/10.1016/j.physleta.2017.06.005
0375-9601/

2. A relatively simple model for percolation properties of real networks

2.1. The basic model based on generating function

Let us begin with reviewing the model of Newman et al. They defined the generating function for the degree of a node as

$$G_0(x) = \sum_{k=0}^{\infty} p_k \cdot x^k, \tag{1}$$

where p_k is the probability that a randomly chosen node has degree k. Then the generating function for the remaining degree of a node pointed by a randomly chosen edge was

$$G_1(x) = \sum_{k=1}^{\infty} \frac{p_k \cdot k \cdot x^{k-1}}{\sum_{k=1}^{\infty} p_k \cdot k}. \tag{2}$$

They used $H_0(x)$ and $H_1(x)$ to denote the generating functions for the size of a component which includes a randomly chosen node and which is reached by a randomly chosen edge, respectively. They had

$$H_0(x) = x \cdot G_0\big(H_1(x)\big), \tag{3}$$

and based on Fig. 1 (a), they got

$$H_1(x) = x \cdot G_1\big(H_1(x)\big). \tag{4}$$

By demanding both sides of Eqs. (4) and (3) on x, and substituting $x = 1$ into them, they had

$$H_1'(1) = \frac{1}{1 - G_1'(1)}, \tag{5}$$

and

$$H_0'(1) = 1 + \frac{G_0'(1)}{1 - G_1'(1)}. \tag{6}$$

As $H_1'(1) \geq 1$, and according to Eq. (5), they derived the expression for the position of the phase transition:

$$G_1'(1) = 1. \tag{7}$$

When $G_1'(1) \geq 1$, there exists a giant component, and at this moment, they predicted u the probability that a randomly chosen edge is not part of the giant component by solving the equation

$$u = G_1(u). \tag{8}$$

Note that $u = 1$ is always a solution of Eq. (8), but this solution should be abandoned except when $G_1'(1) = 1$, and some solution between 0 and 1 should be reserved. Then they calculated S the fraction of the graph occupied by the giant component as

$$S = 1 - G_0(u). \tag{9}$$

Though this algorithm is mainly applied to a series of networks with the same degree distribution, it could also be used to analyze the single network, as we can easily attain $G_0(x)$ and $G_1(x)$ for that single network, and calculate $G_0'(1)$ and $G_1'(1)$ based on them. Papers like Fu et al. [12] and Li and Zhang [13] used similar algorithm to study some real interacting networks, but they did not incorporate the occupation probability into their theory model.

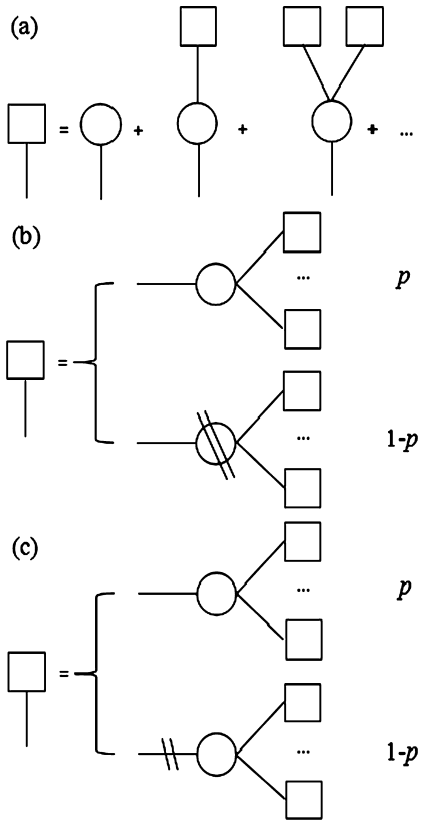

Fig. 1. A diagramatical representation of the sum rule for the connected component reachable by following a randomly chosen edge, where squares denote components and circles denote nodes. (a) excerpted from [6] discusses networks with the same degree distribution never under the node or edge removal, while (b) and (c) are for site and bond percolation, respectively, for instance, squares in (b) denote the connected components after the node removal with the probability $(1 - p)$.

2.2. The extended model for site percolation on real networks and the comparison between its estimates and simulation values

Now we start our model construction from the case of site percolation, and here we use p to denote site occupation probability. Based on Fig. 1(b) and in analogy with Eqs. (3) and (4), we have

$$H_1(x) = 1 - p + p \cdot x \cdot G_1\big(H_1(x)\big), \tag{10}$$

and

$$H_0(x) = 1 - p + p \cdot x \cdot G_0\big(H_1(x)\big). \tag{11}$$

By demanding both sides of Eqs. (10) and (11) on x, and substituting $x = 1$ into them, we have

$$H_1'(1) = \frac{p}{1 - p \cdot G_1'(1)}, \tag{12}$$

and

$$H_0'(1) = p + \frac{p^2 \cdot G_0'(1)}{1 - p \cdot G_1'(1)}. \tag{13}$$

From Eq. (13), we derive the phase transition expression:

$$p_c \cdot G_1'(1) = 1. \tag{14}$$

So the critical value of the node occupation probability is $p_c = \frac{1}{G_1'(1)}$, when p is above or equal to this value, the phase transition will take place. At this moment, we have

$$u = 1 - p + p \cdot G_1(u), \tag{15}$$

and