

# SCIENTIFIC REPORTS



OPEN

## How Inhomogeneous Site Percolation Works on Bethe Lattices: Theory and Application

Jingli Ren<sup>1</sup>, Liying Zhang<sup>1</sup> & Stefan Siegmund<sup>2</sup>

Received: 29 October 2015

Accepted: 15 February 2016

Published: 01 March 2016

Inhomogeneous percolation, for its closer relationship with real-life, can be more useful and reasonable than homogeneous percolation to illustrate the critical phenomena and dynamical behaviour of complex networks. However, due to its intricacy, the theoretical framework of inhomogeneous percolation is far from being complete and many challenging problems are still open. In this paper, we first investigate inhomogeneous site percolation on Bethe Lattices with two occupation probabilities, and then extend the result to percolation with  $m$  occupation probabilities. The critical behaviour of this inhomogeneous percolation is shown clearly by formulating the percolation probability  $P_\infty(p)$  with given occupation probability  $p$ , the critical occupation probability  $p_c = \sup\{p | P_\infty(p) = 0\}$ , and the average cluster size  $\chi(p)$  where  $p$  is subject to  $P_\infty(p) = 0$ . Moreover, using the above theory, we discuss in detail the diffusion behaviour of an infectious disease (SARS) and present specific disease-control strategies in consideration of groups with different infection probabilities.

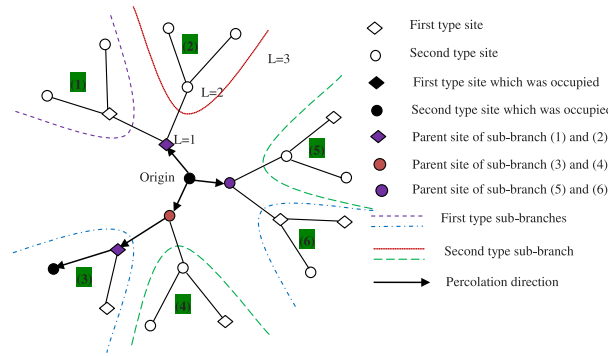
Percolation (for reviews see Stauffer 1979<sup>1</sup>, Essam 1980<sup>2</sup>) is the random occupation of sites or bonds on lattices or networks, named as site percolation or bond percolation, respectively. Site percolation is more general than bond percolation because every bond model may be reformulated as a site model (on a different graph) and the converse is in general not true<sup>3</sup>. The appeal of percolation is the occurrence of a critical phenomenon, which has attracted attention for a wide range of applications: liquid flows in porous media<sup>4,5</sup>, epidemic spread<sup>6-8</sup>, granular and composite materials<sup>9-12</sup>, forest fires<sup>13-15</sup> and fracture patterns and earthquakes in rocks<sup>16</sup>.

The research originated from homogeneous percolation, i.e., percolation with a single occupation probability. For example, Fisher and Essam (1961) solved homogeneous percolation problems on Bethe lattices<sup>17</sup>; Sykes and Essam (1964) studied the exact critical occupation probabilities in two dimensions<sup>18</sup>; Gerald (*et al.*, 2001) derived the value of critical occupation probabilities for a four-dimensional percolation problem on hyper-cubic lattices<sup>19</sup>.

Building on those results, researchers began to study inhomogeneous percolation<sup>20-26</sup>, in which sites (or bonds) may have different occupation probabilities. In 1982, Kesten<sup>20</sup> obtained a critical surface of occupation probabilities for inhomogeneous percolation on square lattices. In 2013, Grimmett<sup>21</sup> extended the result of inhomogeneous bond percolation to triangular and hexagonal lattices utilizing Russo-Seymour-Welsh (RSW) theory of box-crossings. Recently, Radicchi<sup>27</sup> studied percolation on not necessarily infinite graphs and used graph decomposition methods to identify abrupt and continuous changes in percolation.

In the current paper, we focus on inhomogeneous percolation on Bethe lattices. The interest in a good understanding of this inhomogeneous percolation process is twofold. Firstly, Bethe lattices might hold fractal structures, which allow more complex behaviour than square or triangular lattices. The results of this inhomogeneous percolation on Bethe lattices have more extensive applications, e.g., the spreading problem of social networks or infectious diseases (see section III). Secondly, as is well known, most results on percolation are obtained using the heuristic approximation or numerical approaches<sup>27-29</sup> and it is difficult to get the exact solutions for average cluster sizes and the percolation probability even for homogeneous percolation. In this paper, we investigate inhomogeneous site percolation on Bethe lattices from two occupation probabilities to  $m$  occupation probabilities: for the case of two occupation probabilities, we present the explicit formula of the critical occupation probability and the exact solution of average cluster size  $\chi(p)$ ; and for the case of  $m$  occupation probabilities, because of computational complexity, we formulate and obtain the numerical solutions of average cluster size and percolation probability, which might shed light on deriving the results of inhomogeneous percolation. Besides, we analyse in detail

<sup>1</sup>School of Mathematics and Statistics, Zhengzhou University, 450001 Zhengzhou, P.R. China. <sup>2</sup>Center for Dynamics & Institute for Analysis, Department of Mathematics, TU Dresden, 01062 Dresden, Germany. Correspondence and requests for materials should be addressed to J.R. (email: renjl@zzu.edu.cn)



**Figure 1.**  $PBL(3; (1, p_1), (2, p_2))$  with 3 sub-generations.

the spread of SARS (an infectious disease) using this inhomogeneous percolation on Bethe lattices with dynamically changing parameters. We present specific control strategies for SARS by comparing the critical infection probabilities, the average numbers of infected individuals for each day, and the probability of the large-scale outbreak of SARS.

### Basic theory

In ref. 30 a Bethe lattices is defined to be a tree where each site has  $Z$  neighbours,  $Z$  is also named coordination number. For the sake of convenience, we denote it as  $BL(Z)$ .

On a Bethe lattices, if all sites are occupied randomly with the same probability  $p$ , independent of its neighbours, we call the percolation process as homogeneous site percolation and write it as  $PBL(Z; p)$ . If the sites of a Bethe lattices  $BL(Z)$  are occupied with different probabilities, then the percolation is inhomogeneous. To be concrete, without loss of generality, assume that  $Z$  neighbours of each site are occupied randomly with  $m$  ( $m \in \mathbb{N}, m \leq Z$ ) occupation probabilities  $p_1, \dots, p_m$ , and then, we divide  $Z$  neighbours of each site into  $m$  different groups ( $m$  is the number of types of neighbours according to occupation probabilities), where,  $n_1$  of  $Z$  neighbours are sites with occupation probability  $p_1$ ,  $n_2$  are sites with occupation probability  $p_2, \dots$ , and  $n_m$  are sites with occupation probability  $p_m$ . The inhomogeneous percolation is indicated with  $PBL(Z; (n_1, p_1), (n_2, p_2), \dots, (n_m, p_m))$ . See Fig. 1 for an illustration of  $PBL(3; (1, p_1), (2, p_2))$ .

In this paper, we first consider inhomogeneous percolation on a Bethe lattices with  $m = 2$ , that is  $PBL(Z; (n_1, p_1), (n_2, p_2))$ . Clearly, in this case if  $p_1 = p_2$  or  $n_1 = 0$  or  $n_2 = 0$ , inhomogeneous percolation will specialize to homogeneous percolation on  $BL(Z)$ <sup>30</sup>. In a second step, we generalize the results of  $PBL(Z; (n_1, p_1), (n_2, p_2))$  to the case of  $PBL(Z; (n_1, p_1), (n_2, p_2), \dots, (n_m, p_m))$ .

**Critical surface of occupation probability.** Occupation probability is the probability with which the sites of a network are occupied. For  $PBL(Z; (n_1, p_1), (n_2, p_2))$ , assume that a grandparent site is a first type-site on an infinite lattice, then for the parent site, there are  $n_1 - 1$  sub-branches that begin with the first type-sites and  $n_2$  sub-branches beginning with the second type-sites (Fig. 1). According to the binomial distribution, only  $(n_1 - 1)p_1 + n_2p_2$  branches are accessible on average. On the other hand, if the grandparent site is a second type-site, then for the parent site, there are  $n_1$  sub-branches that begin with first type-sites and  $n_2 - 1$  sub-branches which begin with second type-sites. In this case, on average, only  $n_1p_1 + (n_2 - 1)p_2$  branches are accessible. Recalling that the ratio of these two types of sites is  $n_1/n_2$ , and according to expectation theory, on average, only

$$\frac{n_1[(n_1 - 1)p_1 + n_2p_2] + n_2[n_1p_1 + (n_2 - 1)p_2]}{n_1 + n_2} \tag{1}$$

branches are accessible at each step. In order to get an infinite cluster, it is necessary that the quantity (1) is equal or greater than one. Therefore, the critical condition such that an infinite cluster (percolating cluster) first occurs is

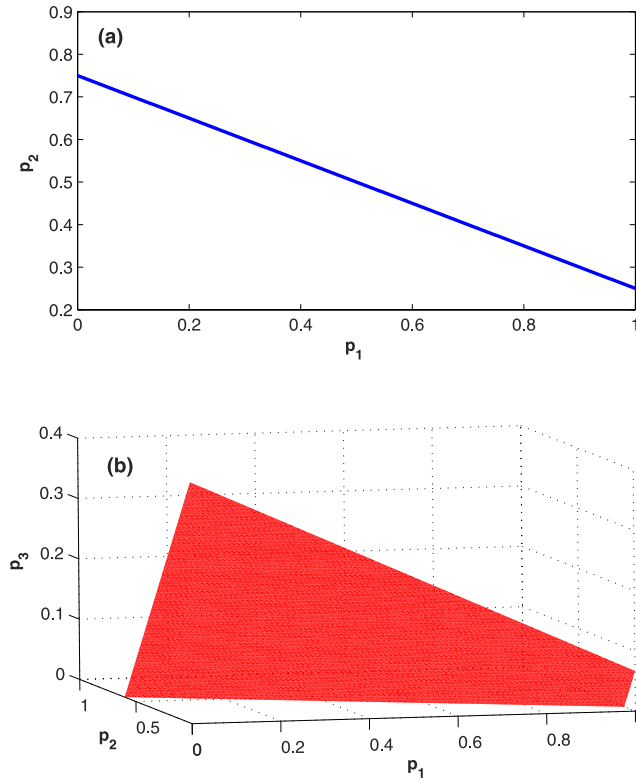
$$\frac{n_1(n_1 - 1)p_1 + n_2(n_2 - 1)p_2 + n_1n_2(p_1 + p_2)}{n_1 + n_2} = 1 \tag{2}$$

from equation (2), we derive the critical surface of occupation probability

$$n_1p_1 + n_2p_2 = \frac{n_1 + n_2}{n_1 + n_2 - 1} \tag{3}$$

this critical surface of  $PBL(Z; (n_1, p_1), (n_2, p_2))$  is a line with slope  $-n_1/n_2$  in the occupation probabilities set  $[0, 1]^2$ . As an example, see Fig. 2(a) for an illustration of the critical surface of  $PBL(3; (1, p_1), (2, p_2))$ .

In a similar way, we can derive the following result.



**Figure 2. Critical surface of occupation probability.** (a) Critical surface of occupation probability of  $PBL(3; (1, p_1), (2, p_2))$ . (b) Critical surface of occupation probability of  $PBL(6; (1, p_1), (2, p_2), (3, p_3))$ .

**Theorem 1.** The critical surface of  $PBL(Z; (n_1, p_1), (n_2, p_2), \dots, (n_m, p_m))$  is given by

$$n_1 p_1 + n_2 p_2 + \dots + n_m p_m = \frac{n_1 + n_2 + \dots + n_m}{n_1 + n_2 + \dots + n_m - 1} \tag{4}$$

**Proof of Theorem 1.** For  $PBL(Z; (n_1, p_1), (n_2, p_2), \dots, (n_m, p_m))$ , if the grandparent site is an  $i$ th ( $i = 1, \dots, m$ ) type-site, then for the parent site, there are  $n_i - 1$  sub-branches that begin with an  $i$ th type-site and  $n_k$  ( $k = 1, 2, \dots, m, k \neq i$ ) sub-branches beginning with  $k$ th type-sites (Fig. 1). According to the multi-binomial distribution, in this case only  $\sum_{k=1, k \neq i}^m n_k p_k + (n_i - 1)p_i$  branches are accessible on average.

Recalling the occupation probabilities for all types of sites and according to expectation theory, overall, only

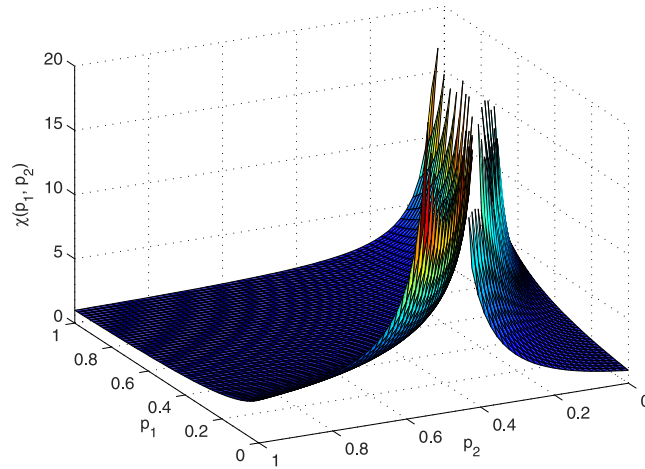
$$\frac{\sum_{i=1}^m n_i [\sum_{k=1, k \neq i}^m n_k p_k + (n_i - 1)p_i]}{\sum_{i=1}^m n_i} \triangleq \varphi(p) \tag{5}$$

branches are accessible on average. In view that  $\varphi(p)$  is equal to one on the critical surfaces, we have equation (4), the critical surface of  $PBL(Z; (n_1, p_1), (n_2, p_2), \dots, (n_m, p_m))$ .

It is clear that  $P_\infty(p) > 0$  if  $\varphi(p) > 1$ ;  $P_\infty(p) = 0$  if  $\varphi(p) \leq 1$ , where,  $p = (p_1, p_2, \dots, p_m) \in [0, 1]^m$ ;  $P_\infty(p)$  is the percolation probability (existing infinite clusters). The theory generalizes the concept of exact critical value  $p_c$  in homogeneous percolation. For consistency, we also indicated the critical surface of  $PBL(Z; (n_1, p_1), (n_2, p_2), \dots, (n_m, p_m))$  by  $p_c$ . The critical surface  $p_c = \{p | \varphi(p) = 1\}$  is the subset of the occupation-probability set  $\{p_1, p_2, \dots, p_m | 0 \leq p_1, p_2, \dots, p_m \leq 1\}$  with all  $p = (p_1, p_2, \dots, p_m)$  satisfying equation (4). Then, in the supercritical phase,  $p > p_c$  stands for  $\{p | \varphi(p) > 1\}$ , and there exists almost surely at least one infinite cluster of occupied sites. Contrarily, in the subcritical phase,  $p < p_c$  stands for  $\{p | \varphi(p) < 1\}$ , and all clusters of occupied sites are almost surely finite. For an illustration, Fig. 3(b) is the critical surface of  $PBL(6; (1, p_1), (2, p_2), (3, p_3))$ .

For an intuitive understanding, take  $p_1 = p$  as a reference, and  $p_i = k_i p$ ,  $i = 2, 3, \dots, m$ , we have from equation (4) that  $p_c = \frac{n_1 + n_2 + \dots + n_m}{(n_1 + n_2 + \dots + n_m - 1)(n_1 + k_2 n_2 + \dots + k_m n_m)}$ . In this case  $p_c$  is degenerated to a critical point, which is consistent with homogeneous percolation (in the supercritical phase,  $p > p_c$  and in the subcritical phase,  $p < p_c$ ).

The critical surface of occupation probabilities consists of those probabilities at which the percolation behaviour of the system changes essentially. We can control the percolation behaviour if we know the critical surface of occupation probabilities.



**Figure 3.** Average cluster size  $\chi(p)$  of PBL(3; (1,  $p_1$ ), (2,  $p_2$ )). The left-hand curve is a sketch of the average cluster size  $\chi(p)$  and the right-hand curve is a sketch of the mean size of finite clusters of occupied sites when  $p > p_c$ .

**Average cluster size of occupied sites.** The average cluster size  $\chi(p_1, p_2)$  of occupied sites is the mean size of the finite (non-percolating) clusters of occupied sites. It is closely related to the critical surface of occupation probabilities. We consider  $\chi(p_1, p_2)$  and the relationship between  $\chi(p_1, p_2)$  and critical surface.

**Theorem 2.** Assume that the Bethe lattice  $BL(Z)$  is infinite with occupation probabilities  $p = (p_1, p_2)$  such that  $\{p_1, p_2 | 0 \leq p_1, p_2 \leq 1\}$ , then the average cluster size of occupied sites of PBL ( $Z; (n_1, p_1), (n_2, p_2)$ ) satisfies

$$\chi(p_1, p_2) = \frac{(n_1 + n_2)(1 + p_1 + p_2 + p_1 p_2 B) + (n_1 p_1 + n_2 p_2) A}{(n_1 + n_2)[1 + (n_1 + n_2 - 1)p_1 p_2] |1 - B|} \tag{6}$$

where  $A = n_1 p_1 + n_2 p_2 + (n_1 + n_2)p_1 p_2, B = (n_1 - 1)p_1 + (n_2 - 1)p_2 + (n_1 + n_2 - 1)p_1 p_2$ .

**Proof of Theorem 2.** If the Bethe lattice is infinite, all sites are equivalent for evaluating the average cluster size of occupied sites. Let  $\chi_i(p_1, p_2)$  be the average cluster size of the centre sites which are of type  $i$  ( $i = 1, 2$ ), and  $B_{ij}$  ( $i, j = 1, 2$ ) be the contribution (to the average cluster size) from a sub-branch which begins with a  $j$ th type-site and whose parent site is an  $i$ th type-site (Fig. 1). Then

$$\chi_i(p_1, p_2) = 1 + n_i B_{i1} + n_i B_{i2}, \quad i = 1, 2 \tag{7}$$

where, the first term is the contribution from the centre site itself, the second term is the contribution from the first type branches, and the third term is the contribution of the second type branches.

According to expectation theory, on average, the average cluster size is

$$\chi(p_1, p_2) = \frac{n_1 \cdot \chi_1(p_1, p_2) + n_2 \cdot \chi_2(p_1, p_2)}{n_1 + n_2} \tag{8}$$

Based on definition of inhomogeneous percolation on a Bethe lattice, the following recurrence relations can be concluded

$$B_{ij} = p_j \left[ 1 + \sum_{k=1, k \neq i}^2 n_k B_{jk} + (n_i - 1) B_{ji} \right], \quad i, j = 1, 2 \tag{9}$$

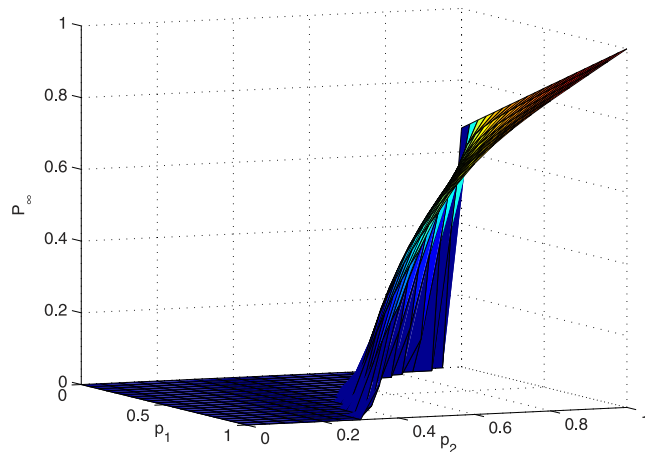
Solving  $B_{ij}$  from equation (9), we have from (7) and (8) that

$$\chi_1(p_1, p_2) = \frac{1 + p_1 + p_2 + p_1 A + p_1 p_2 B}{[1 + (n_1 + n_2 - 1)p_1 p_2] |1 - B|}, \quad \chi_2(p_1, p_2) = \frac{1 + p_1 + p_2 + p_2 A + p_1 p_2 B}{[1 + (n_1 + n_2 - 1)p_1 p_2] |1 - B|}$$

and

$$\chi(p_1, p_2) = \frac{(n_1 + n_2)(1 + p_1 + p_2 + p_1 p_2 B) + (n_1 p_1 + n_2 p_2) A}{(n_1 + n_2)[1 + (n_1 + n_2 - 1)p_1 p_2] |1 - B|}$$

where,  $A = n_1 p_1 + n_2 p_2 + (n_1 + n_2)p_1 p_2$  and  $B = (n_1 - 1)p_1 + (n_2 - 1)p_2 + (n_1 + n_2 - 1)p_1 p_2$ . For  $p_1 = p_2$  or  $n_1 = 0$  or  $n_2 = 0$ , equation (6) reduces to  $\chi(p) = \frac{1+p}{1-(Z-1)p}$ .



**Figure 4.** Percolation probability  $P_\infty(p)$  of PBL (3; (1,  $p_1$ ), (2,  $p_2$ )).

The average cluster size (of occupied sites)  $\chi(p_1, p_2)$ , which is a function of the occupation probability  $p = (p_1, p_2)$ , can reveal the intensity of percolation. For  $p < p_c$ ,  $\chi(p)$  increases rapidly with  $p$  (Fig. 3), and diverges in a power law of the distance between  $p$  and  $p_c$ , as  $p$  approaches  $p_c$  from below. For  $p > p_c$ , there exist infinite clusters of occupied sites and their number increases as  $p \rightarrow 1$ . On the other hand, the numbers of finite clusters (of occupied sites) and their sizes are reducing. Therefore, for  $p > p_c$ , the average sizes of finite clusters  $\chi(p_1, p_2)$  decrease with  $p$  increasing (Fig. 3).

Similarly, generalizing the result to PBL ( $Z; (n_1, p_1), (n_2, p_2), \dots, (n_m, p_m)$ ), we first solve  $B_{ij}$  ( $i, j = 1, 2, \dots, m$ ) from equation (10), where the  $B_{ij}$  have the same meaning as above

$$B_{ij} = p_j \left[ 1 + \sum_{k=1, k \neq i}^m n_k B_{jk} + (n_i - 1) B_{ji} \right], \quad i, j = 1, 2, \dots, m. \quad (10)$$

Then, substitute the solution of equation (10) into equation (11) to get the values of  $\chi_k(p)$ ,  $k = 1, 2, \dots, m$ , here  $p = (p_1, p_2, \dots, p_m)$ .

$$\chi_k(p) = 1 + \sum_{j=1}^m n_j B_{kj}, \quad k = 1, 2, \dots, m \quad (11)$$

According to expectation theory, the average cluster size is

$$\chi(p) = \frac{\sum_{k=1}^m n_k \cdot \chi_k(p)}{\sum_{k=1}^m n_k} \quad (12)$$

Because the explicit expression of  $\chi(p)$  is very complex, we provide only the derivation process.

**Percolation probability.** In this part, we mainly discuss the percolation probability, i.e., the probability that the origin site belongs to a percolating infinite cluster. Percolation probability is indicated as  $P_\infty(p)$  that can reveal the intensity of percolating, for  $p > p_c$  (Fig. 4).

For PBL ( $Z; (n_1, p_1), (n_2, p_2)$ ), in order to determine  $P_\infty(p)$ , let  $P_\infty^i(p)$  denote the probability that an  $i$ th type origin site belongs to a percolating infinite cluster, where  $p = (p_1, p_2)$ ,  $i = 1, 2$ .

A site belongs to a percolating infinite cluster, which means, not only the site itself is occupied, but also at least one of the  $Z = n_1 + n_2$  branches (originating from the site) connects to the percolating cluster. Both of these are independent of each other, so,  $P_\infty^i(p_1, p_2) = p_i (1 - Q_{i1}^{n_1} Q_{i2}^{n_2})$ . According to mean theory, it can be concluded that

$$\begin{aligned} P_\infty(p_1, p_2) &= \frac{n_1 \cdot P_\infty^1(p_1, p_2) + n_2 \cdot P_\infty^2(p_1, p_2)}{n_1 + n_2} \\ &= \frac{n_1 \cdot p_1 \cdot (1 - Q_{11}^{n_1} Q_{12}^{n_2}) + n_2 \cdot p_2 (1 - Q_{21}^{n_1} Q_{22}^{n_2})}{n_1 + n_2} \end{aligned} \quad (13)$$

where,  $Q_{ij}$ ,  $i, j = 1, 2$ , is the probability that a sub-branch does not connect to the percolating cluster and the sub-branch begins with a  $j$ th type-site, whose parent site is an  $i$ th type-site. Then, we have

$$\begin{cases} Q_{1j} = (1 - p_j) + p_j Q_{j1}^{n_1-1} Q_{j2}^{n_2} \\ Q_{2j} = (1 - p_j) + p_j Q_{j1}^{n_1} Q_{j2}^{n_2-1}, \end{cases} j = 1, 2 \quad (14)$$

for each equation of (14), the first term is the probability that the root site of a sub-branch is not occupied and the second term is the probability that the root site of the sub-branch is occupied but no child sub-branch connects to the percolating cluster.

If  $p \leq p_c$ , equation (14) has only the trivial solution  $Q_{ij} = 1, i, j = 1, 2$ , then from (13) we have  $P_\infty(p) = P_\infty^1(p) = P_\infty^2(p) = 0$  (Fig. 4). If  $p > p_c$ , there is a nontrivial solution for equation (14) and then equation (13) has a nonzero solution. In this case, it is not easy to get the exact nontrivial solution of equation (14) for it is a set of multivariable high-order equations. By employing fixed-point iteration, we get a numerical solution of equation (14) instead and then obtain the percolation probability from (13). From the numerical solution (Fig. 4), it can be seen that  $P_\infty(p)$  picks up abruptly at  $p_c$  then increases rapidly with  $p$  increasing.

We extend the result to  $PBL(Z; (n_1, p_1), (n_2, p_2), \dots, (n_m, p_m))$  by a similar analysis. In this general case,  $Q_{ij}, i, j = 1, 2, \dots, m$  satisfy ( $Q_{ij}$  meaning the same as above)

$$Q_{ij} = (1 - p_j) + p_j Q_{ji}^{n_i-1} \prod_{k=1, k \neq i}^m Q_{jk}^{n_k}, \quad i, j = 1, 2, \dots, m. \quad (15)$$

Firstly, derive the solution of (15) by fixed-point iteration, and then substitute  $Q_{ij}$  into (16),

$$P_\infty(p_1, p_2, \dots, p_m) = \frac{\sum_{k=1}^m n_k \cdot p_k \cdot \left(1 - \prod_{j=1}^m Q_{kj}^{n_j}\right)}{\sum_{k=1}^m n_k}. \quad (16)$$

This way, we get  $P_\infty(p)$ , where  $p = (p_1, p_2, \dots, p_m)$ .

**Percolation model of the disease spreading (SARS in Beijing, 2003).** Severe acute respiratory syndrome (SARS) is a viral respiratory illness caused by a corona-virus. In Beijing (China), about 2523 cases have been infected with SARS in 2003. At the beginning of emergence, because of the lack of understanding of SARS and the high mobility of the modern-social activities, SARS spread rapidly. Afterwards, when people found the high infectivity and death rate of SARS, they begun to limit social activity of the public and take strict isolated measure to prevent the spreading of disease then the disease was contained. Is this the proper infection control measure of SARS? Which kind of infectious diseases are suitable for this approach?

In fact, the spreading of SARS is a percolation process. Considering the differences of physical resistibility or intimate contact with the infected individual, we divided people into two groups: the people with higher infection probability (e.g., infants and elderly or healthcare workers), named as susceptible persons; and the people with lower infection probability, named as common persons. Then this disease is modelled as inhomogeneous percolation on Bethe lattices with two occupied probabilities, i.e.,  $PBL(Z; (S, p), (Z - S, kp))$ . Here,  $Z$  is the average contact number of each person,  $S$  (of  $Z$ ) denotes susceptible persons with infection probability  $p$  and  $Z - S$  (of  $Z$ ) is common persons with infection probability  $kp$  ( $0 \leq k \leq 1$ ).

The first case of SARS was confirmed on March 5 (in Beijing, 2003) and the government started reporting the cases from April 20. According to case-reporting data from April 20 to June 23 (in Beijing, 2003) (from the government bulletin) and the control activities of government, we find three other critical time points of SARS: May 1 (it is legal holidays from May 1 to May 7), May 14 (people generally panicked over SARS and limited their social activities), and May 30 (new cases of SARS considerably decreased). Correspondingly, the spreading process of SARS was divided into five stages. Then, by random simulation, we found that the spreading of SARS has a fifteen-day time delay. Therefore, we changed the five stages of the spreading process of SARS into: stage 1—from March 20 to May 4, stage 2—from May 5 to May 15, stage 3—from May 16 to May 28, stage 4—from May 29 to June 13, and stage 5—from June 14 to June 23.

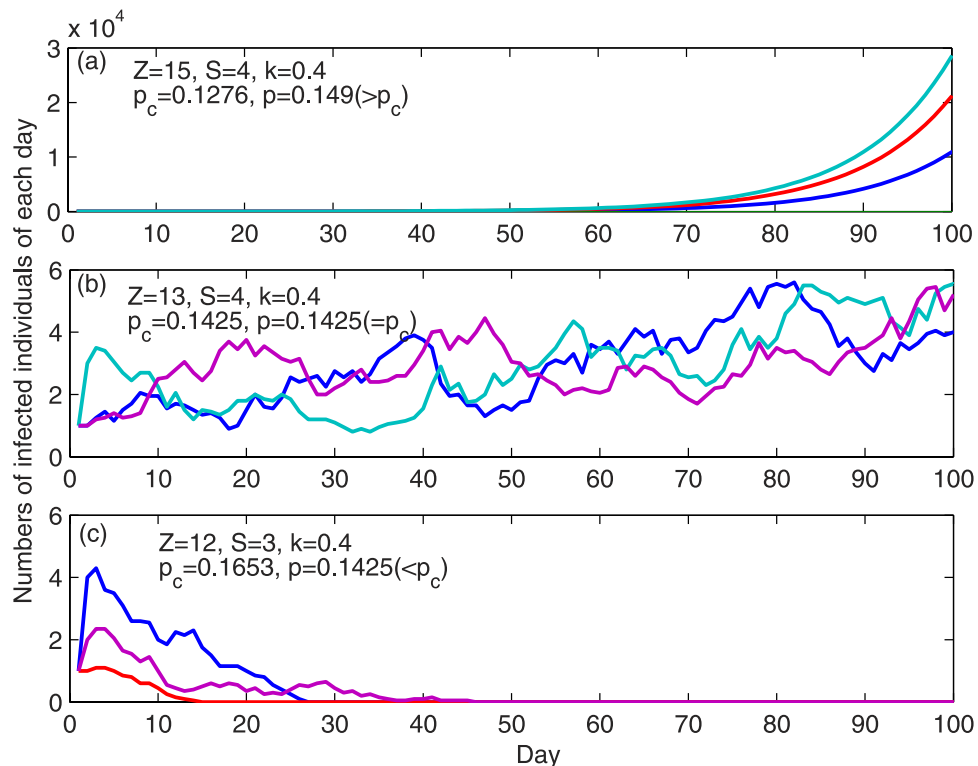
In the initial stage, the average contact number of one infected patient was around fifteen ( $Z = 15$ ), in which two or three person had infected SARS (Gong *et al.*<sup>31</sup>), so the infection probability was around 0.149 ( $p = 0.149$ ). We take  $S = 4$  according to the percentage of susceptible people in the population and we set  $k = 0.4$  by statistical investigation. It is worth noting that the infection probability  $p$  is the manifestation of the spreading intensity of the disease, which can only be slightly affected by the protective approach. Therefore, the infection probability was adjusted to 0.1425 in the second stage and third stage, and was adjusted to 0.141 in the fourth stage and fifth stage. The other parameters, i.e.,  $Z$  and  $S$  would change with prevention (isolation of infected patient and restriction of travel) and  $k$  remains unchanged in different stages. By statistical investigation, we set  $Z = 13$  and  $S = 4$  in second stage,  $Z = 12$  and  $S = 3$  in third stage,  $Z = 11$  and  $S = 3$  in fourth stage,  $Z = 8$  and  $S = 2$  in fifth stage.

Obviously, the spreading model of SARS (in Beijing) is inhomogeneous percolation on a Bethe lattice with dynamically changing parameters. See Table 1 for the model division and the parameters.

**Results of percolation model and control measures.** From equation (3), equation (6), and equation (13), we acquire the critical infection probabilities, average number of infected individuals, and the probability of large-scale outbreak for the SARS percolation model  $PBL(Z; (S, p), (Z - S, kp))$  in different spreading stages (Table 1).

	Mar. 20–May 4	May 5–May 15	May 16–May 28	May 29–June 13	June 14–June 23
Infection probability— $p$	0.149	0.1425	0.1425	0.141	0.141
Average contact number— $Z(S, Z-S)$	15(4, 11)	13(4, 9)	12(3, 9)	11(3, 8)	8(2, 6)
Critical infection probability— $p_c$	0.1276	0.1425	0.1653	0.1774	0.2597
Average infected persons by one patient	infinite	infinite	74.77	17.4	3.0277
Probability of large-scale outbreak— $p_{sc}$	0.0233	0.0045	0	0	0

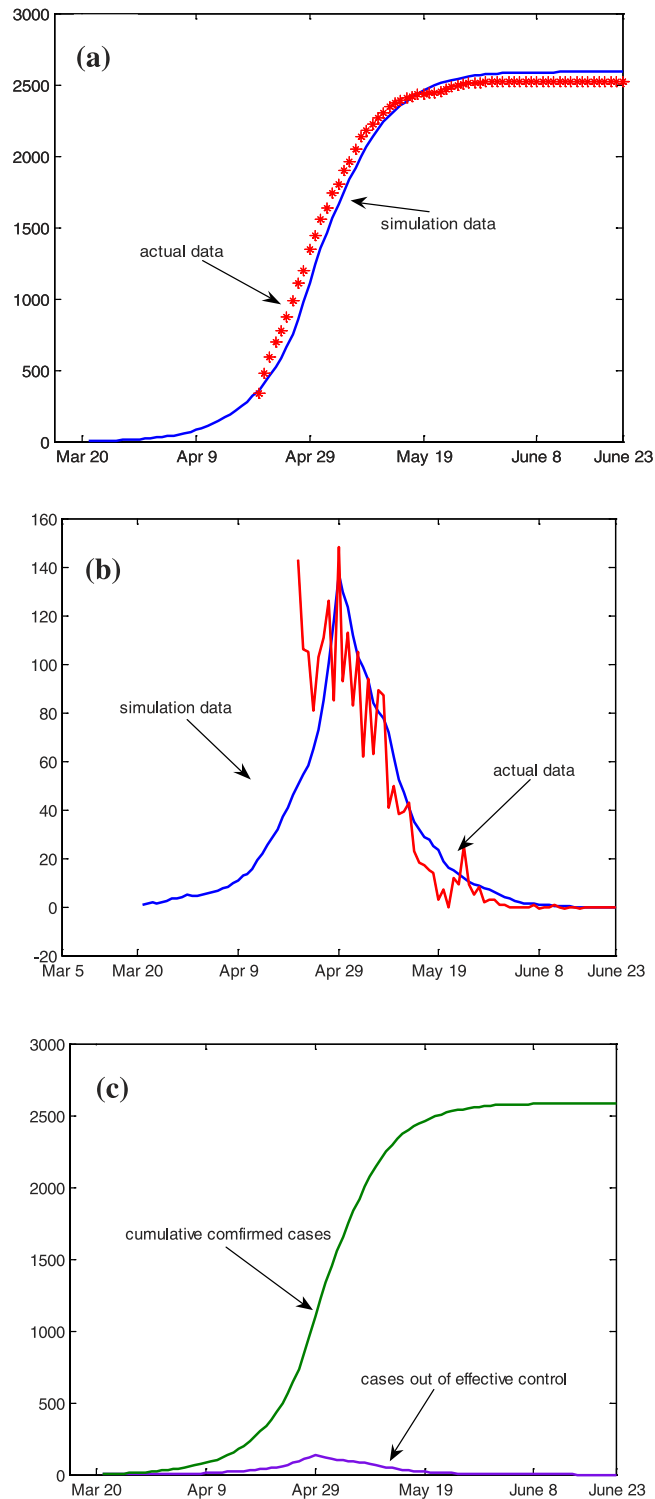
**Table 1.** The results of inhomogeneous percolation of SARS and some parameters by simulation and statistical analysis.



**Figure 5.** The average number of infected individuals versus every day with infection probabilities that are larger, equal, and smaller than critical infection probability.

It can be concluded from Table 1, that, if one is infected with SARS and would live as a normal person, then the disease would infect a massive crowd of healthy people except for  $Z \leq 13$  and  $S \leq 3$ . That warns us, at the initial stage of SARS, humankind should try their best to find SARS patients as early as possible and isolate them from healthy people. Nevertheless, during the incubation period, it is inevitable that some infected persons, who cannot be found and live as the normal person around us that is quite dangerous. At this time, the most effective way is to reduce outdoor activities of public then cut down the average contact number.

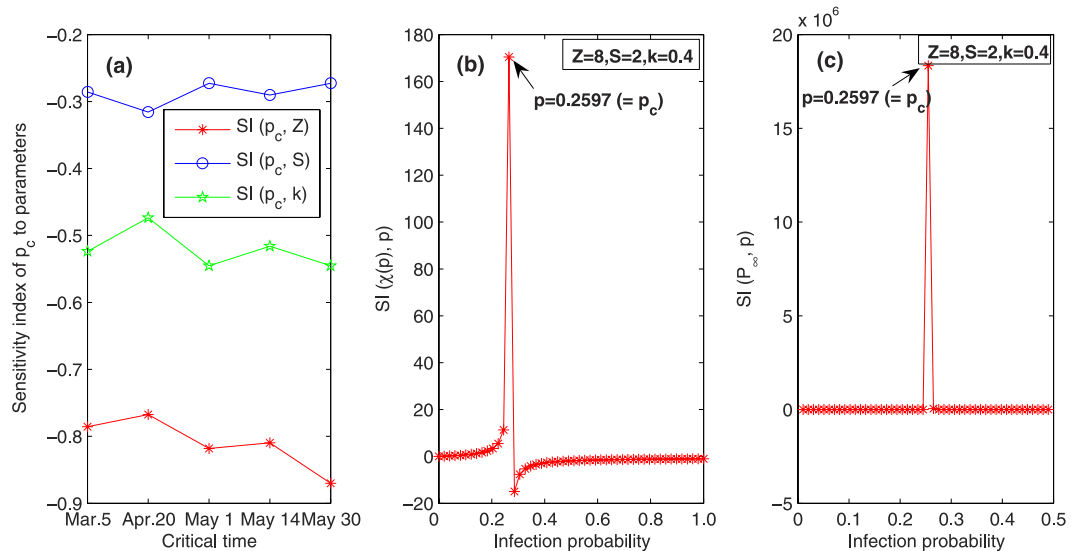
For more accurate and meticulous disease-control strategies, we scrutinized the dynamic-dependent relationship between  $p$ ,  $p_c$  and the average number of infected cases with subtle parameters by a Monte-Carlo simulation. In the initial stage of the SARS process, from March 20 to April 8, according to the above analysis,  $Z = 15$ ,  $S = 4$ ,  $k = 0.4$ , and  $p = 0.149$ , based on these, the average number of infected cases of each day was simulated. See Fig. 5(a) for an illustration. It is clear that the random variation of the infected number has an incremental trend and that implies that the disease would infect a large amount of persons. Then from April 9 to April 19, during the second stage of SARS, some protective measures were taken with the understanding of the disease, so parameters changed to  $Z = 13$ ,  $S = 4$ ,  $k = 0.4$ , and  $p = 0.1425$ . By simulation, we found that the infected number changes chaotically as in Fig. 5(b). With the time going by, from April 20 to May 4, the severity of the disease gradually being known, more protective measures were taken and parameters reduced further to  $Z = 12$ ,  $S = 3$ ,  $k = 0.4$ , and  $p = 0.1425$ . In this stage, the simulation revealed that the average infected number of each day fluctuates with a trend of decline and it would be zero after a period, which predicates the infectious disease can be controlled without any additional measures (Fig. 5(c)). In fact, by simulating the percolation, we acquire  $p_c = 0.1276 (< p)$  in the initial stage,  $p_c = 0.1425 (= p)$  in the second stage, and  $p_c = 0.1653 (> p)$  in the third stage, i.e., the initial stage is a supercritical phase of the spread of SARS, the second stage is a critical phase, and the third stage is a subcritical phase, which agrees with the simulation. We could conclude that, near the



**Figure 6.** Simulation results (a) Cumulative cases of SARS with simulation data and actual data. (b) Daily new cases of simulation and report. (c) The cumulative cases and daily new cases by simulation.

critical point, slightly adjusting of the system parameters would cause a fundamental change of the trend of infectious diseases. Therefore, in order control the large-scale outbreak of the disease, we must try our best to make  $p < p_c$ , even if the infection probability is only a little smaller than the critical infection probability. Actually, if  $p < p_c$ , the probability of a large outbreak of SARS is zero; and if  $p$  reaches and crosses  $p_c$ , the probability picks up as a power law with exponent one in terms of  $p - p_c$ , and the disease will outbreak rapidly.





**Figure 7.** Sensitivity index (a) Sensitivity index of critical infection probability to input parameters- $Z$ ,  $S$  and  $k$  on five critical days of SARS in Beijing (2003) (b) Sensitivity index of the average infected cases by one patient to input variable  $p$  (infection probability). (c) Sensitivity index of the probability of SARS large-scale outbreak to input variable  $p$ .

As we know, reducing outdoor activities of the public is a powerful strategy for infectious disease with latent period but it severely obstructs the people’s daily life and social economy. The methods in this paper will supply a quantitative measure for the risk of disease outbreaks and to guide our practice more appropriately.

**Predictions**

Based on the model of the spreading of SARS, we could supplement the data of cases from March 5 to April 20, during which the recording data was missing. See Fig. 6.

Figure 6(a) is the actual cumulative case-reporting data and simulative data. New cases of each day are displayed in Fig. 6(b). We can find that the inhomogeneous percolation is in good fitting with the dynamic process of SARS spreading and time delay is the considerable feature of SARS. The relationship between cumulative confirmed cases and cases out of effective control displays in Fig. 6(c). It can be seen that there will be many persons infected with SARS even if only a few cases of SARS are out of effective control.

**Sensitivity analysis.** The sensitivity index is the ratio of the change in output to the change in input of parameters or variables<sup>32</sup>. Taking into account the characteristics of the model, we employ a one factor at a time (OAT) approach, which is more agile and easy to interpret. The popular sensitivity index of OAT approach is  $SI(Y, \theta) = \frac{\partial Y}{\partial \theta} \cdot \frac{\theta}{Y}$ , where  $Y$  is the output,  $\theta$  is the input,  $SI(Y, \theta)$  is the sensitivity index of  $Y$  to  $\theta$ , and  $\frac{\partial Y}{\partial \theta}$  is the partial derivative of  $Y$  with respect to  $\theta$ . The quotient  $\frac{\theta}{Y}$  is introduced to normalize the index by removing the affects of units<sup>33</sup>.

First, we get  $p_c = \frac{Z}{(Z-1)[S+k(Z-S)]}$  from equation (3), and then derive that:

$$SI(p_c, Z) = - \frac{kZ^2 + (1-k)S}{(Z-1)[S+k(Z-S)]},$$

$$SI(p_c, S) = - \frac{(1-k)S}{S+k(Z-S)},$$

$$SI(p_c, k) = - \frac{k(Z-S)}{S+k(Z-S)}.$$

Where,  $p_c$  is the critical infection probability, one of the output of the SARS-percolation model;  $Z$ ,  $S$ ,  $k$  denote the average contact number of each patient, the number of susceptible persons of  $Z$ , the ratio of infection probabilities of susceptible person to common person (see page 12), respectively. They are all input parameters.

Based on the Table 1, the sensitivity indexes of  $p_c$  to three input parameters at five critical time points are obtained, which are all negative scalars. These suggest that the decreasing of each input parameters correspond to the increasing of  $p_c$ . Among them, the absolute value of  $SI(p_c, Z)$  is about 0.8, which is maximum,  $|SI(p_c, k)|$  is about 0.5, and  $|SI(p_c, S)|$  is about 0.3. It is clear that  $p_c$  has greater sensitivity to  $Z$ . See Fig. 7(a) for an illustration.

By a similar analysis with a numerical approach, we obtain  $SI(\chi(p), p)$  and  $SI(P_\infty(p), p)$ . The sensitivity index  $SI(\chi(p), p)$  is shown in Fig. 7(b). It indicates that  $\chi(p)$  is more sensitive near the  $p_c$ . Since  $\chi(p)$  is the average size

of all-finite clusters of infected cases, the  $SI(\chi(p), p)$  exhibits negative value when  $p$  is greater than  $p_c$ . By a numerical approach, the sensitivity of the probability of large-scalar outbreak of SARS  $SI(P_\infty(p), p)$  is displayed in Fig. 7(c).  $SI(P_\infty(p), p)$  has similar characteristics as  $SI(\chi(p), p)$ . They are both more sensitive near the  $p_c$ .

## Conclusion

In this paper, we present a theoretical framework for inhomogeneous site percolation on Bethe Lattices, and apply it to investigate a diffusion problem of an infectious disease. It is found that the inhomogeneous percolation on Bethe lattices serves as an appropriate model to describe the dynamic spreading behaviour of the infectious disease (SARS). The percolation model of SARS is not only in good agreement with the actual recorded data, but also can be used to predict the future trend of the disease and supply the missing data of the past. Moreover, it can provide quantitative results for government to make more proper disease-control strategies.

## References

1. Stauffer, D. Scaling theory of percolation clusters. *Phys. Rep.* **54**, 1–74 (1979).
2. Essam, J. W. Percolation theory. *Rep. Prog. Phys.* **43**, 833–912 (1980).
3. Grimmett, G. R. *Percolation 2nd edn*, 23–24 (Springer, Verlag Berlin Heidelberg, 1999).
4. Ghanbarzadeh, S., Prodanović, M. & Hesse, M. A. Percolation and grain boundary wetting in anisotropic texturally equilibrated pore networks. *Phys. Rev. Lett.* **113**, 048001 (2014).
5. Moon, K. & Girvin, S. M. Critical behaviour of superfluid<sup>4</sup>He in aerogel. *Phys. Rev. Lett.* **75**, 1328 (1995).
6. Moore, C. & Newman, M. E. J. Epidemics and percolation in small-world networks. *Phys. Rev. E* **61**, 5678 (2000).
7. Janssen, H. K., Müller, M. & Stenull, O. Generalized epidemic process and tricritical dynamic percolation. *Phys. Rev. E* **70**, 026114 (2004).
8. Newman, M. E. J., Jensen, I. & Ziff, M. R. Percolation and epidemics in a two-dimensional small world. *Phys. Rev. E* **65**, 021904 (2002).
9. Odagaki, T. & Toyofuku, S. Properties of percolation clusters in a model granular system in two dimensions. *J. Phys.-Condens. Mat.* **10**, 6447–6452 (1998).
10. Tobochnik, J. Granular collapse as a percolation transition. *Phys. Rev. E* **60**, 7137 (1999).
11. Bondt, S. D., Froyen, L. & Deruytere, A. Electrical conductivity of composites: a percolation approach. *J. Mater. Sci.* **27**, 1983–1988 (1992).
12. Bentz, D. P. & Garboczi, E. J. Modelling the leaching of calcium hydroxide from cement paste: effects on pore space percolation and diffusivity. *Mater. Struct.* **25**, 523–533 (1992).
13. Bramwell, S. T. *et al.* Universal fluctuations in correlated systems. *Phys. Rev. Lett.* **84**, 3744 (2000).
14. Squires, S. *et al.* Weakly explosive percolation in directed networks. *Phys. Rev. E* **87**, 052127 (2013).
15. Guisoni, N., Loscar, E. S. & Albano, E. V. Phase diagram and critical behavior of a forest-fire model in a gradient of immunity. *Phys. Rev. E* **83**, 011125 (2011).
16. Sahimi, M., Robertson, M. C. & Sammis, C. G. Fractal distribution of earthquake hypocenters and its relation to fault patterns and percolation. *Phys. Rev. Lett.* **70**, 2186 (1993).
17. Fisher, M. E. & Essam, J. W. Some cluster size and percolation problems. *J. Math. Phys.* **2**, 609–619 (1961).
18. Sykes, M. F. & Essam, J. W. Exact critical percolation probabilities for site and bond problems in two dimensions. *J. Math. Phys.* **5**, 1117–1127 (1964).
19. Gerald, P., Ziff, R. M. & Stanley, H. E. Percolation threshold, fisher exponent, and shortest path exponent for four and five dimensions. *Phys. Rev. E* **64**, 026115 (2001).
20. Harry, K. *Percolation Theory for Mathematicians*. 53–55 (Birkh01user Boston, 1982).
21. Grimmett, G. R. & Manolescu, I. Inhomogeneous bond percolation on square, triangular and hexagonal lattices. *Ann. Probab.* **41**, 2990–3025 (2013).
22. Iliev, G. K., Rensburg, E. J. J. V. & Madras, N. Phase diagram of inhomogeneous percolation with a defect plane. *J. Stat. Phys.* **158**, 255–299 (2014).
23. Gennadiy, B. & Yessica, C. S. Percolation and lasing in real 3D crystals with inhomogeneous distributed random pores. *Phys. B: Condens. Mat.* **453**, 12 (2014).
24. Ziff, R. M. *et al.* The critical manifolds of inhomogeneous bond percolation on bow-tie and checkerboard lattices. *J. Phys. A: Math. Theor.* **45**, 494005 (2012).
25. Scullard, C. R. & Ziff, R. M. Critical surfaces for general inhomogeneous bond percolation problems. *J. Stat. Mech.* **03**, p03021 (2010).
26. Bender, A. *et al.* Inhomogeneous markov approach to percolation theory based propagation in random media. *IEEE. T. Antenn. Propag.* **56**, 3271–3284 (2008).
27. Radicchi, F. Percolation in real interdependent networks. *Nature Phys.* **11**, 597–602 (2015).
28. Karrer, B., Newman, M. E. J. & Zdeborová, L. Percolation on sparse networks. *Phys. Rev. Lett.* **113**, 208702 (2014).
29. Hamilton, K. E. & Pryadko, L. P. Tight lower bound for percolation threshold on an infinite graph. *Phys. Rev. Lett.* **113**, 208701 (2014).
30. Christensen, K. & Moloney, N. R. *Complexity and Criticality* 15–22 (Imperial College Press, London, 2005).
31. Gong, J. H. *et al.* Simulation and analysis of control of severe acute respiratory syndrome. *J. Remote Sens.* **7**, 260–265 (2003).
32. Hamby, D. M. A review of techniques for parameter sensitivity analysis of environmental models. *Environ. Monit. Assess.* **32**, 135–154 (1994).
33. Saltelli, A., Tarantola, S. & Campolongo, F. Sensitivity analysis as an ingredient of modelling. *Stat. Sci.* **15**, 377–395 (2000).

## Acknowledgements

This work is supported by the NSFC (No. 11271339) project, the Plan for Scientific Innovation Talent of Henan Province, and the ZDGD 13001 program.

## Author Contributions

R.J.L. conceived the project and derived the theory. Z.L.Y. established and analysed the model. S.S. partly contributed to the theory and gave some advices on the model. All authors contributed to writing the manuscript.

## Additional Information

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Ren, J. *et al.* How Inhomogeneous Site Percolation Works on Bethe Lattices: Theory and Application. *Sci. Rep.* **6**, 22420; doi: 10.1038/srep22420 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>