Impacts of preference and geography on epidemic spreading

Xin-Jian Xu,[1] Xun Zhang,[2] and J. F. F. Mendes[1],*

[1]*Departamento de Física da Universidade de Aveiro, 3810-193 Aveiro, Portugal*
[2]*Institute of Theoretical Physics, Lanzhou University, Lanzhou Gansu 730000, China*

We investigate the standard susceptible-infected-susceptible model on a random network to study the effects of preference and geography on diseases spreading. The network grows by introducing one random node with m links on a Euclidean space at unit time. The probability of a new node i linking to a node j with degree k_j at distance d_{ij} from node i is proportional to k_j^A/d_{ij}^B, where A and B are positive constants governing preferential attachment and the cost of the node-node distance. In the case of $A=0$, we recover the usual epidemic behavior with a critical threshold below which diseases eventually die out. Whereas for $B=0$, the critical behavior is absent only in the condition $A=1$. While both ingredients are proposed simultaneously, the network becomes robust to infection for larger A and smaller B.

PACS number(s): 89.75.Hc, 87.23.Ge, 05.70.Ln, 87.19.Xx

The classical mathematical approach to diseases spreading either ignores the population structure or treats populations as distributed in a regular medium [1,2]. However, it has been suggested recently that many social, biological, and communication systems possess two universal characters, the small-world effect [3] and the scale-free property [4], which can be described by complex networks whose nodes represent individuals and links represent interactions among them [5,6]. In view of the wide occurrence of complex networks in nature, it is important to study the effects of topological structures on the dynamics of epidemic spreading. Pioneering works [7–11] have given some valuable insights: for homogeneous networks (e.g., exponential networks), there are critical thresholds of the spreading rate below which infectious diseases will eventually die out; on the contrary, even infections with low spreading rates will prevail over the entire population in heterogeneous networks (e.g., scale-free networks). This radically changes many conclusions drawn from classic epidemic modelling. Furthermore, it has been observed that the heterogeneity of a population network in which the disease spreads may have noticeable effects on the evolution of the epidemic as well as the corresponding immunization strategies [12–16].

For many real networks, however, individuals are embedded in a Euclidean space and the interactions among them usually depend on their spatial distances and take place among their nearest neighbors [17–19]. For instance, the number of long-ranged links and the number of edges connected to a single node are limited by the spatial embedding, particularly in planar networks. Preferential attachment is weakened by geographical embedding [18]. Also, people have proved that the characteristic distance plays a crucial role in the dynamics taking place on those networks [20–23]. Thus, it is natural to study associated influences of preference and geography on epidemic spreading. But up to now only a few of works address this problem, e.g., modeling transmission as a function of geographical distance [24,25] availably capturing the dynamics of diseases in wild and domesticated animals [26,27].

In this paper, we study the standard susceptible-infected-susceptible (SIS) model on a growing network in Euclidean space. On a vertical plane, the growth of the network depends jointly on two mechanisms, preference and geography. The placement of links is driven by competition between preferential attachment and distance dependence. In the case that the network grows with geographical constraint, we recover the usual epidemic behavior with a critical threshold below which diseases will eventually die out. While the network is totally governed by preferential attachment, the epidemic behavior depends on the preferential exponent. When both factors are considered simultaneously, it becomes difficult for epidemic spreads as the preference has an overwhelming majority than the geography.

Specifically in the two-dimensional x-y plane, we consider a square of unit size and with periodic boundary conditions. To construct a network of N nodes, let $(x_1, x_2, \ldots, x_N)$ and $(y_1, y_2, \ldots, y_N)$ be the $2N$ independent random variables identically and uniformly distributed within the interval $\{0, 1\}$. A specific set of values of the random variables $\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$ is chosen to represent the coordinates of the N randomly distributed nodes [23]. The network starts with m_0 nodes and then the other nodes with m links are added one by one at each time step according to their serial numbers $i = m_0 + 1$ to N. Following ideas proposed by Yook *et al.* [18], the probability that a new node i links to a old node j with k_j links at distance d_{ij} from node i is

$$\prod (k_i, d_{ij}) \sim \frac{k_j^A}{d_{ij}^B}, \tag{1}$$

where A and B are positive constants, governing preferential attachment and the cost of the node-node distance. We note following interesting features. (i) In the case of $A=0$, the network is geographically grown with an exponential distribution of nodes' degree [see Fig. 1(a)]. In the limit of $B \rightarrow \infty$, only the smallest value of d corresponding to the nearest node will contribute with probability 1 [18]. (ii) In the case of $B=0$, the network reduces to the Barabási-Albert (BA) graph only for $A=1$. In the region $0 < A < 1$, the nodes' degree distribution is stretched exponential. For $A>1$, a fi-
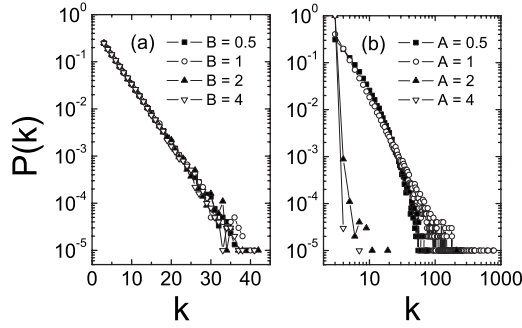
*jfmendes@fis.ua.pt

FIG. 1. Degree distributions of the generated networks with $m_0=m=3$ in two special cases $A=0$ (a) or $B=0$ (b). The size of the network is $N=10^5$.

nite number of nodes connect to nearly all other nodes [28]. That is illustrated by Fig. 1(b).

To estimate the effect of the network's topology on epidemic dynamics, we will investigate the standard SIS model [8]. This model relies on a coarse-grained description of individuals in the population. Namely, each node of the network represents an individual and each link is a connection along which the infection can spread to other individuals. The individuals can only exist in two states, susceptible and infected. At each time step, each susceptible node is infected with probability ν if it is connected to one or more infected nodes. At the same time, the infected nodes become susceptible again with probability δ, defining an effective spreading rate $\lambda = \nu / \delta$. We can set $\delta=1$ without lack of generality, since it only affects the definition of the time scale of the virus propagation. Individuals run stochastically through the cycle susceptible→infected→susceptible.

Let us focus on the case $A=0$ first, i.e., the network grows with geographical constraint. According to Eq. (1), the preferential attachment is excluded. Since all nodes are uniformly distributed in the square, the only effect of the factor B is determining the average path length of the network while the degree distribution of nodes has the same behavior. For small B, the role of node-node distance is weak and old nodes are linked with approximate randomness. When B becomes large, the geographical influence is strong and only nodes around the new one will be connected with large possibility, hence the local clustering. This feature is reflected in Table I, that is, the average path length L and the clustering coefficient C get larger with the increase of B.

Figure 2 shows the evolution of the infected nodes density

TABLE I. The average path length L and clustering coefficient C of the generated networks with size $N=10\,000$ in the case of $A=0$ for different values of B.

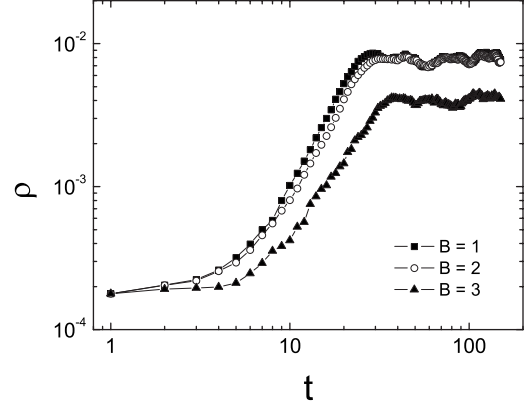| B	L	C
0.5	5.13(1)	0.00(1)
1	5.13(5)	0.00(2)
2	5.40(1)	0.03(9)
3	6.41(7)	0.19(2)
4	7.11(2)	0.31(0)

FIG. 2. Density of infected nodes ρ as a function of time t in the network with $A=0$ and $B=1,2,3$, respectively. The spreading rate is $\lambda=0.15$. Simulations were performed on networks with $N=10\,000$ and $m=m_0=3$.

as a function of time for epidemics with $\lambda=0.15$. We start from a single infected node of the network, and iterate the rules of the SIS model with parallel updating. Each curve represents the average over ten different starting configurations, performed on ten different realizations of the random networks. We clearly notice a great influence of the geography on the spreading velocity of diseases, namely, the smaller the parameter B is, the more fast the infection propagates. For finite B, since the geography does not change the network's connectivity distribution [see Fig. 1(a)], the evolution of node i's degree can be written as

$$\frac{dk_i(t)}{dt} \approx \frac{m}{t}, \qquad (2)$$

with the initial condition $k_i(i)=m$. One can easily write the solution

$$k_i(t) = m\left(\ln \frac{t}{i} + 1\right), \qquad (3)$$

and accordingly obtain the degree distribution

$$P(k) = \frac{1}{t}\sum_{i=1}^{t} \delta(k_i(t) - k) = \frac{1}{m}e^{-k/m+1}. \qquad (4)$$

In complex networks, the basic reproductive number takes the form $R_0 \sim \langle k^2 \rangle / \langle k \rangle$ [13]. In contrast to the classical result, it defines an epidemic threshold $\lambda_c = \langle k \rangle / \langle k^2 \rangle$. Combining Eq. (4), we have

$$\lambda_c = \frac{\int k \frac{1}{m} e^{-k/m+1} dk}{\int k^2 \frac{1}{m} e^{-k/m+1} dk} = \frac{2}{5m}. \qquad (5)$$

In Fig. 3, we plot the steady density of infected nodes ρ as a function of the spreading rate λ for the case of $A=0$, which is the time average of the fraction of infected individuals reached after an initial transient regime. Simulations were computed over 50 different starting configurations, per-
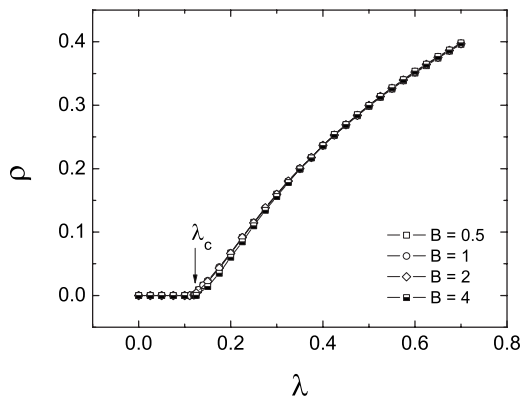
FIG. 3. Density of infected nodes ρ as a function of λ in the network with $A=0$ and $B=0.5, 1, 2, 4$, respectively.

formed on 50 different realizations of the networks. The size of networks is $N=10^5$. As shown in Fig. 3, all curves display the same behavior and the SIS model exhibits an epidemic threshold $\lambda_c=0.12(2)$ [29], despite the variety of B, which is in agreement with the analytical prediction $\lambda_c=2/(5 \times 3)$ $=0.133$ [Eq. (5)]. This is different than the results gained by Santos *et al.*, who also studied the SIS model on a homogeneous small-world network [30], the critical value of λ changes smoothly as one varies the rewiring probability without changing the degree distribution. In our model, the increase of B gives rise to the average path length and the clustering coefficient. Whereas in Ref. [30], those network features reduce with the increase of the rewiring probability.

Next, we will investigate the dynamics in the case of $B=0$, i.e., the network grows following preferential attachment. For the linear preference $A=1$, the generated network reduces to the exact BA graph. As shown in Fig. 4, the behavior of the density of infected nodes follows the property, $\rho \sim e^{-1/m\lambda}$ [11], which implies the surprising absence of any epidemic threshold in the model, i.e., $\lambda_c=0$. As to the nonlinear case, we first consider the region $A<1$ and go down a little from $A=1$. In this event, highly connected nodes become less attractive for attachment compared with the linear preference. The resulting degree distribution is of the form [28] $P(k) \sim k^{-A} \exp[-\mu k^{1-A}/(1-A)]$, where μ is a positive
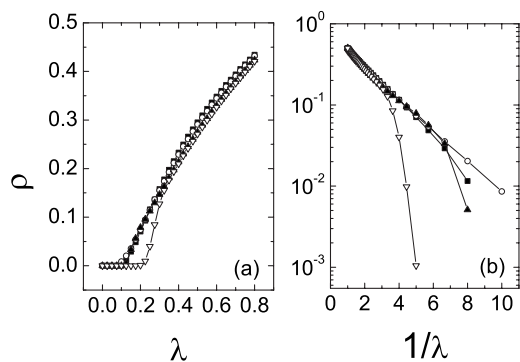

FIG. 4. Density of infected nodes ρ vs λ (a) and $1/\lambda$ (b) in the network with $B=0$. The preference exponents are $A=0.5$ (closed squares), 1 (open circles), 1.5 (closed diamonds), and 2 (open triangles), respectively.
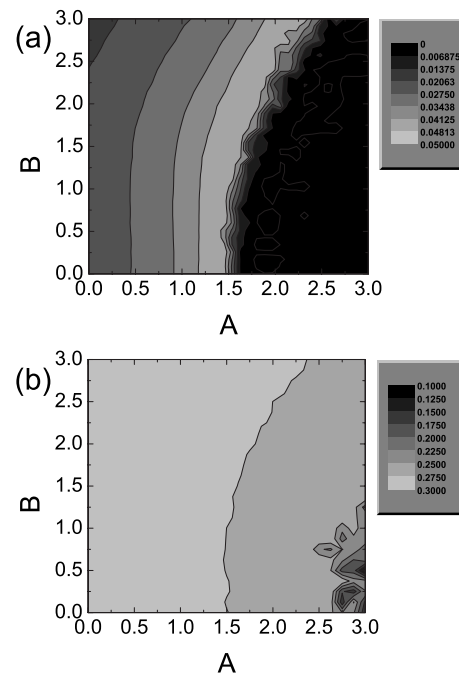

FIG. 5. Density of infected nodes ρ in A-B plane for the spreading rate $\lambda=0.15$ (a) and 0.5 (b), respectively.

constant depends on A, $\mu=\Sigma_{k=2}^{\infty}\Pi_{i=2}^{k}(1+\mu/i^A)^{-1}$. In the region $A>1$, the attractiveness of the old highly connected nodes increases which results in a small number of nodes that get all connections in the network. As shown in Fig. 4, there are epidemic thresholds of the SIS model for $A=0.5$, 1.5, and 2, respectively. Furthermore, the threshold increases as A becomes larger.

Finally, we plot the prevalence ρ in the A-B plane in Fig. 5 to shown the influence of the competition of two ingredients on infections. For $\lambda=0.15$, there exists a set of peaks in Fig. 5(a). Namely, given the value of B, as A is increased initially, the density of infected nodes ρ increases gradually and reaches a maximum for some value of A, and then decreases rapidly to 0 as A is increased further. Furthermore, the contour planes of ρ takes an excursion to right with the increase of B. We argue that the following factor should be taken into account to understand this performance. According to Eq. (1), as B increases, the new nodes are preferential to connect their nearest neighbors, which results in the decrease of the number of long-ranged links. The network becomes more local clustering and robust to epidemic spreading. To keep the same prevalence, the effect of node's degree should be strengthened, i.e., increases A. Thus the contour planes lean to right as B becomes larger. For $\lambda=0.5$, the prevalence ρ displays a different behavior. As shown in Fig. 5(b), ρ decreases monotonically as A gets larger, and if B decreases at the same time, the network becomes robust to diseases.

To summarize, we have studied the SIS model on a random network. On an Euclidean plane, the network grows depending jointly on the preference and the geography. The former indicates the attractiveness of highly connected nodes and the later denotes the geographical constraint. It is found that both factors have great influences on the infection. For

the network growing with the geographically constraint ($A=0$), we recover the usual epidemic behavior with a critical threshold below which diseases will eventually die out. While the network is purely governed by preferential attachment ($B=0$), the epidemic behavior depends on the preferential exponent and the critical phenomenon is absent only in the condition $A=1$. When both factors are present simultaneously, the network becomes robust to diseases as the preference has an overwhelming majority than the geography. In real world, agents located on different positions according to the competition between the preference and the geography. The above description of the spreading dynamics might contribute to understanding realistic epidemics.

[1] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases* (Hafner, New York, 1975).
[2] R. M. Anderson and R. M. May, Nature (London) **318**, 323 (1985).
[3] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[4] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[5] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
[6] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge, 2004).
[7] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, J. Math. Biol. **28**, 365 (1990).
[8] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, 1991).
[9] C. Moore and M. E. J. Newman, Phys. Rev. E **61**, 5678 (2000).
[10] A. L. Lloyd and R. M. May, Science **292**, 1316 (2001).
[11] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[12] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).
[13] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **26**, 521 (2002).
[14] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **65**, 036104 (2002).
[15] K. T. D. Eames and M. J. Keeling, Proc. R. Soc. London, Ser. B **270**, 2565 (2003).
[16] R. Cohen, S. Havlin, and D. ben-Avraham, Phys. Rev. Lett. **91**, 247901 (2003).
[17] R. Durrett, SIAM Rev. **41**, 677 (1999).
[18] S.-H. Yook, H. Jeong, and A.-L. Barabási, Proc. Natl. Acad. Sci. U.S.A. **99**, 13382 (2002).
[19] M. T. Gastner and M. E. J. Newman, Eur. Phys. J. B **49**, 247 (2006).
[20] A. F. Rozenfeld, R. Cohen, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **89**, 218701 (2002).
[21] C. P. Warren, L. M. Sander, and I. M. Sokolov, Phys. Rev. E **66**, 056105 (2002).
[22] S. S. Manna, G. Mukherjee, and P. Sen, Phys. Rev. E **69**, 017102 (2004).
[23] G. Mukherjee and S. S. Manna, Phys. Rev. E **74**, 036111 (2006).
[24] R. Durrett and S. Levin, Theor Popul. Biol. **46**, 363 (1994).
[25] H. W. Hethcote, SIAM Rev. **42**, 599 (2000).
[26] J. D. Murray, E. A. Stanley, and D. L. Brown, Proc. R. Soc. London, Ser. B **229**, 111 (1986).
[27] M. J. Keeling, M. E. J. Woolhouse, R. M. May, G. Davies, and B. T. Grenfell, Nature (London) **421**, 136 (2003).
[28] P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).
[29] Here, λ_c for different B only changes tinily in permillage and we deem the independence of the critical value on geographical constraint.
[30] F. C. Santos, J. F. Rodrigues, and J. M. Pacheco, Phys. Rev. E **72**, 056128 (2005).