Localization and Spreading of Diseases in Complex Networks

A. V. Goltsev, 1,2 S. N. Dorogovtsev, 1,2 J. G. Oliveira, 1,3 and J. F. F. Mendes 1

1 Department of Physics & ISN, University of Aveiro, 3810-193 Aveiro, Portugal
2 Ioffe Physico-Technical Institute, 194021 St. Petersburg, Russia
3 Departamento de Engenharia Física, Faculdade de Engenharia, Universidade do Porto, rua Dr. Roberto Frias, 4200-465 Porto, Portugal

Using the SIS model on unweighted and weighted networks, we consider the disease localization phenomenon. In contrast to the well-recognized point of view that diseases infect a finite fraction of vertices right above the epidemic threshold, we show that diseases can be localized on a finite number of vertices, where hubs and edges with large weights are centers of localization. Our results follow from the analysis of standard models of networks and empirical data for real-world networks.

PACS numbers: 05.10.-a, 05.40.-a, 05.50.+q, 87.18.Sn

Survey of infectious diseases reveals that before an outbreak, often, if not typically, a disease is localized within a small group of individuals. Changes in environmental conditions or increase in the frequency of external contacts result in an epidemic outbreak. In the present paper we propose an approach that enables us to describe quantitatively this important localization-delocalization phenomenon. Our approach is based on the SIS model of spreading of diseases in weighted and unweighted networks, where the weights of edges encode frequency of contacts between vertices. It is widely accepted that in uncorrelated networks the epidemic threshold Λc of the infection rate λ is λc = q / q2, where (q) and (q2) are the first and second moments of the degree distribution. So in networks with a finite (q2), the threshold should be non-zero, while it is zero if (q2) diverges. One should stress however that all these well-known results were obtained only within a mean-field theory, actually within an annealed network approximation in which a random network is substituted for its fully connected weighted counterpart. Contrastingly, one can show exactly for an arbitrary graph that λc is actually determined by the largest eigenvalue Λ1 of the adjacency matrix Aij of the graph, and λc = 1 / Λ1. For uncorrelated networks, it was found that Λ1 is determined by the maximum degree qmax, Λ1 ∝ qmax. Then, if in the infinite size limit, qmax tends to infinity, as, e.g., in the Erdős-Rényi graphs, this leads to an amazing conclusion that the epidemic threshold is absent even in (infinite) networks with a finite (q2) in contrast to the mean-field result.

In the present paper we develop a spectral approach to the SIS model on complex networks. We show that the contradiction between the mean-field approximation and the exact result can be resolved if we take into account localization of diseases. It turns out that, in contrast to the mean field theory, in which a finite fraction of vertices are infected at λ > λc, there are actually two scenarios of the spreading of diseases. If Λ1 corresponds to a localized eigenstate, then, at λ right above λc = 1 / Λ1, disease is mainly localized on a finite number of vertices, i.e., the fraction of infected vertices is negligibly small in large networks. With further increase of λ, disease gradually infects more and more vertices until it will infect a finite fraction of vertices. In the second scenario, Λ1 corresponds to a delocalized state. Then already at λΛ1−1≪1, disease infects a finite fraction of vertices. Analysing network models and real-world networks, we show that hubs, edges with large weights, and other dense subgraphs can be centers of localization.

We consider the standard SIS model of disease spreading in a complex network of size N having the adjacency matrix with arbitrary entries Aij ≥ 0. Infected vertices become susceptible with unit rate, and each susceptible vertex becomes infected by its infective neighbor with the infection rate λ. The probability ρi(t) that vertex i is infected at time t is described by the evolution equation

\[ \frac{d\rho_i(t)}{dt} = -\rho_i(t) + \lambda[1 - \rho_i(t)] \sum_{j=1}^{N} A_{ij} \rho_j(t). \]

In the steady state, at t → ∞, the probability ρi ≡ ρi(∞) is determined by a non-linear equation,

\[ \rho_i = \frac{\lambda \sum_j A_{ij} \rho_j}{1 + \lambda \sum_j A_{ij} \rho_j}, \]

which has a non-zero solution ρi > 0 if λ is larger than a so-called epidemic threshold λc. In this case, the prevalence ρ ≡ \(\sum_i \rho_i/N\) is also non-zero. This transition contrasts to a “thermodynamic phase transition”, since the critical feature at λc is present even for finite nets.

Spectral approach.—To solve the SIS model, we use the spectral properties of the adjacency matrix \( \hat{A} \). The eigenvalues \( \hat{A} \) and the corresponding eigenvectors \( f \) with components \( f_i \) are solutions of the equation \( \Lambda f = \hat{A} f \). Since the matrix \( \hat{A} \) is real and symmetric, its N eigenvectors \( f(\Lambda) (\Lambda_{\text{max}} ≡ \Lambda_1 ≥ \Lambda_2 ≥ \ldots Λ_N) \) form a complete orthonormal basis. According to the Perron-Frobenius theorem, the largest eigenvalue \( \Lambda_1 \) and the corresponding principal eigenvector \( f(\Lambda_1) \) of a real nonnegative symmetric matrix are nonnegative. The probabilities \( \rho_i \)
can be written as a linear superposition,
\[ \rho_i = \sum_{\lambda} c(\lambda) f_i(\lambda). \]  
(3)

The coefficients \( c(\lambda) \) are projections of the vector \( \rho \) on \( f(\lambda) \). Substituting Eq. (3) into Eq. (2), we obtain
\[ c(\lambda) = \lambda \sum_{\lambda'} \Lambda' c(\lambda') \sum_{i=1}^{N} \frac{f_i(\lambda') f_i(\lambda)}{1 + \lambda \sum_{\lambda} \Lambda c(\lambda)f_i(\lambda)}. \]  
(4)

In order to find the epidemic threshold \( \lambda_c \) and \( \rho(\lambda) \) near \( \lambda_c \), it is enough to take into account only the principal eigenvector \( f(\lambda_1) \) in Eqs. (3) and (4), i.e., \( \rho_i \approx c(\lambda_1) f_i(\lambda_1) \). Solving Eq. (4) with respect to \( c(\lambda_1) \) in the leading order in \( \tau \equiv \lambda \Lambda_1 - 1 \ll 1 \), we find \( \lambda_c = 1/\Lambda_1 \) and \( \rho \approx \alpha_1 \tau \), where the coefficient \( \alpha_1 \) is
\[ \alpha_1 = \sum_{i=1}^{N} f_i(\lambda_1)/[N \sum_{i=1}^{N} f_i^2(\lambda_1)]. \]  
(5)

This expression is asymptotically exact if there is a gap between \( \lambda_1 \) and \( \lambda_2 \) (see also Ref. [10]). Thus, at \( \tau \ll 1 \), \( \rho \) is determined by the principal eigenvector. Contributions of other eigenvectors are of the order of \( \tau^2 \). Considering two largest eigenvalues in Eq. (4), \( \Lambda_1 \) and \( \Lambda_2 \), and their eigenvectors, we obtain \( \rho(\lambda) \approx \alpha_1 \tau + \alpha_2 \tau^2 \) and so on.

The usual point of view is that \( \alpha_1 \) is of the order of \( O(1) \), and so a finite fraction of vertices is infected right above \( \lambda_c \). To learn if another behavior is possible, we study whether \( \Lambda_1 \) corresponds to a localized or delocalized state. We introduce the inverse participation ratio
\[ IPR(\Lambda) \equiv \sum_{i=1}^{N} f_i^4(\Lambda). \]  
(6)

If, in the limit \( N \to \infty \), \( IPR(\Lambda) \) is of the order of \( O(1) \), then the eigenvector \( f(\Lambda) \) is localized. If \( IPR(\Lambda) \to 0 \) then this state is delocalized. For a localized \( f(\Lambda) \) the components \( f_i(\Lambda) \) are of the order of \( O(1) \) only at few vertices. For a delocalized \( f(\Lambda) \) we usually have \( f_i(\Lambda) \sim O(1/\sqrt{N}) \ll 1 \). From Eq. (6) it follows that if the principal eigenvector \( f(\Lambda_1) \) is localized, then \( \alpha_1 \sim O(1/N) \) and so \( \rho \approx \alpha_1 \tau \sim O(1/N) \). In this case, above \( \lambda_c \) disease is localized on a finite number \( N_{\rho} \) of vertices. If \( f(\Lambda_1) \) is delocalized, then \( \rho \) is of the order of \( O(1) \), and disease infects a finite fraction of vertices right above \( \lambda_c \). These two contrasting scenarios are shown in Fig. 1 for the SIS model on the karate-club network [11] and the weighted collaboration networks of scientists posting preprints on the astrophysics archive at arXiv.org, 1995–1999, and the condensed matter archive at January 1, 1995 – March 31, 2005 [12]. The astro-ph and karate-club nets have delocalized principal eigenstates while the cond-mat-2005 net has a localized principal eigenstate. Numerical solution of Eq. (2) gives \( \alpha_1 \approx 1.8 \times 10^{-3} \) for the astro-ph net and \( \alpha_1 \approx 1.5 \times 10^{-4} \) for the cond-mat-2005 net.

There is an iterative method to find \( \Lambda_1 \) and \( IPR(\Lambda_1) \),
\[ \Lambda_1 = \lim_{n \to \infty} \Lambda_1(n) = \lim_{n \to \infty} (g(n) \tilde{A}g(n))/|g(n)|^2, \]  
(7)
\[ IPR(\Lambda_1) = \lim_{n \to \infty} \sum_{i=1}^{N} (g_i(n))^4/|g(n)|^4, \]  
(8)
where \( g(n+1) = \tilde{A}g(n) \) and \( g(0) \) is a positive vector. \( \Lambda_1(n) \) is the lower bound of \( \Lambda_1 \). For unweighted networks, i.e., \( A_{ij} = 0, 1 \), choosing \( g(0) = 1 \), we obtain that the first iteration \( n = 1 \) gives \( g_i(1) = g_i \) and
\[ \Lambda_1(1) = \frac{1}{\langle q^2 \rangle N} \sum_{i,j} q_i A_{ij} q_j = \Lambda_{MF} + \frac{\langle q \rangle \sigma^2 r}{\langle q^2 \rangle}, \]  
(9)
where \( \Lambda_{MF} \equiv \langle q^2 \rangle / \langle q \rangle \) and \( r \) is the Pearson coefficient (see also Ref. [14]). Eq. (9) shows the influence of degree-degree correlations on \( \Lambda_1 \). The first iteration also gives the mean-field result \( IPR = \langle q^4 \rangle / [N \langle q^2 \rangle^2] \sim O(1/N) \). Already a few iterations give good approximations for \( \Lambda_1 \) and \( IPR \) if the principal eigenstate is delocalized but more iterations are needed if this eigenstate is localized.

**Bethe lattice.**—What structural elements of a network are responsible for the localization of the principal eigenstate? Let us consider a Bethe lattice (see Fig. 2). The adjacency matrix of an unweighted regular Bethe lattice with vertices of degree \( k \) has the largest eigenvalue \( \Lambda_1 = k \). The principal eigenvector \( f(\Lambda_1) = N^{-1/2} \) is delocalized. Other eigenstates are not important for us. Let us introduce a hub of degree \( q > k \) connected to the neighbors by edges with a weight \( w \geq 1 \) [see Fig. 2(b)]. Other edges have weight 1. Solving the equation \( \Lambda f = \tilde{A} f \), we find
\[ \Lambda_1 = qw^2/\sqrt{qkw^2 - B}, \]  
(10)
\[ IPR = f_0(\Lambda_1)(1 + qw^2)/(a^2 - B), \]  
(11)
\[ f_0(\Lambda_1) = [(qw^2/2 - B)/(qkw^2 - B)]^{1/2}. \]  
(12)
Here $B \equiv k - 1$ is the branching coefficient of the graph, $a \equiv |qw^2 - B|^{1/2}$, and $f_0(\Lambda_1)$ is the component of the principal eigenvector $f(\Lambda_1)$ at the hub. $f_i(\Lambda_1)$ depends only on distance $n$ from the hub to vertex $i$, i.e., $f_i(\Lambda_1) = f_n(\Lambda_1)$, and exponentially decays with $n \geq 1$,\[ f_n(\Lambda_1) = w f_0(\Lambda_1)/a^n. \] (13)

This exponential decay leads to a finite IPR, so this eigenstate is localized. If $qw^2 \gg B$, then $IPR \to (1 + 1/q)/4$. The second eigenstate with $\Lambda_2 = k$ is delocalized.

The criterion $\Lambda_1 > k$, leads to the condition that $q$ must be larger than $q_{loc} \equiv (B^2 + B)/w^2$. For the SIS model on this Bethe lattice, Eqs. (5) and (13) give $a_1 \propto 1/N$.

Now we consider the Bethe lattice with two hubs of degrees $q_1$ and $q_2$ connected by an edge with a weight $w \geq 1$ [see Fig. 2(b)]. Other edges have weight 1. We find that two eigenstates with eigenvalues $\Lambda_1$ and $\Lambda_2$ can appear above the delocalized eigenstate with $\Lambda_3 = k$,

\[ \Lambda_{1(2)} = a \pm B/a, \]

\[ a_\pm^2 = \frac{1}{2} (Q_1 + Q_2 + w^2) \pm \frac{1}{2} [(Q_1 + Q_2 + w^2)^2 - 4 Q_1 Q_2]^{1/2}, \]

\[ \Psi_1^1 = a_\pm + Q_1, \quad \Psi_1^2 = a_\pm + Q_2 \]

\[ IPR(\Lambda_{1(2)}) = \left| \Psi_1^1(a_\pm + Q_1) + \Psi_1^2(a_\pm + Q_2) \right|/\left| a_\pm - B \right|. \] (14)

The signs $\pm$ correspond to $\Lambda_1$ and $\Lambda_2$, respectively, and $Q_{1(2)} \equiv q_{1(2)} - B - 1$. The components $f_i(\Lambda_{1(2)})$ decrease with increasing distance $n$ from vertex $i$ to the hubs as $\Psi_{1(2)}/a^n$ if vertex $i$ belongs to branches connected to hub 1 or hub 2, respectively. $\Psi_{1(2)}$ are the components of $f(\Lambda_1)$ at hubs 1 and 2, respectively. Their ratio is $\Psi_2/\Psi_1 = (a_\pm^2 - Q_i)/(wa_\pm)$.

The criterion for localization is given by the condition $\Lambda_1, \Lambda_2 > k$. If $q_1 = q_2$ and $w \gg 1$, $IPR(\Lambda_1)$ reaches the maximum value $0.5$ that means localization on two hubs, $\Psi_1 \to \Psi_2 \to 1/\sqrt{2}$.

Scale-free networks.—To study appearance and properties of localized eigenstates in uncorrelated complex networks, we use the static model [15] that generates unweighted scale-free networks with degree distribution $P(q) \propto C q^{-\gamma}$ at $q \gg 1$. Using software OCTAVE, for each realization of a random network of size $N$ with the mean degree $\langle q \rangle$ and $\gamma = 4$, we calculated eigenvalues, eigenvectors, and IPR($\Lambda$) of the adjacency matrix. In networks of size $N = 10^5$, we found that several (typically, from one to three for different realizations) eigenstates appear above the upper delocalized eigenstate. These states are localized at hubs and their properties are described well by Eqs. (10)–(13) with $w = 1$ if the branching coefficient $B$ in these equations is replaced by the averaged branching coefficient $B = \langle q^2 \rangle/\langle q \rangle - 1$. In these scale-free graphs, the upper delocalized eigenstate $\Lambda_d$ is slightly above the mean-field value $\Lambda_{MF} = \langle q^2 \rangle/\langle q \rangle$. The maximum degree $q_{max}$ fluctuates from realization to realization. In realizations with a small $q_{max}$, the principal eigenvector is delocalized and $\Lambda_1 = \Lambda_d$. Therefore, the criterion for localization of the principal eigenstate at a vertex with degree $q_{max}$ can be written as

\[ \Lambda_1 = q_{max}/\sqrt{q_{max} - B} \geq \Lambda_d. \] (15)

The equality here gives the threshold degree $q_{loc}$. For $N = 10^5$, $\langle q \rangle = 10$, and $\gamma = 4$, our numerical calculations give $\langle q^2 \rangle/\langle q \rangle \approx 14.1$ and $\Lambda_d \approx 15.1$. According to Eq. (15), a localized state appears above $\Lambda_d$ if $q_{max}$ is larger than $q_{loc} \approx 214$. Since the averaged value of $q_{max}$ depends on $N$, at small $N$ the probability to generate a graph with $q_{max} > q_{loc}$ is small [16]. Only sufficiently large graphs can have a localized principal eigenstate.

Fig. 3 represents results of our numerical solution of Eq. (4) for the SIS model on a scale-free network of size $N = 10^5$ with the principal eigenvector localized at hub with $q = 323$. Equations (10)–(13) and (5) give $\Lambda_1 = 18.35$, $IPR = 0.23$, and $\alpha_1 \approx 1.4 \times 10^{-3}$. These values agree well with the measured values $\Lambda_1 = 18.47$, $IPR = 0.21$, and $\alpha_1 \approx 1.7 \times 10^{-3}$. The eigenvector with $\Lambda_2$ is also localized at another hub with $q = 254$. The third eigenvalue $\Lambda_3$ corresponds to a delocalized eigenvector. The first two eigenstates allow to describe $\rho(\lambda)$ close to $\Lambda_3 = 1/\Lambda_1$. Accounting for the delocalized eigenstate $\Lambda_3$ gives better results in a broader range of $\lambda$ (see Fig. 3).
from the mean-field value $\Lambda_{MF}$.

In real-world networks the inverse participation ratio

delocalized, the epidemic occurs in the whole region above $\lambda_c = 1/\Lambda_1$. We suggest that further investigations of real-world networks will give many new examples of the disease localization-delocalization phenomena.

This work was partially supported by the FCT projects PTDC: FIS/71551/2006, FIS/108476/2008, SAU-NEU/103904/2008, MAT/114515/2009, and PEst-C/CTM/LA0025/2011.

Table I. Characteristics of real-world networks. $N$ is size, $\gamma$ is the degree distribution exponent, $q_{max}$ is the maximum degree, $q_{loc}$ is the localization threshold found from Eq. (15), $\Lambda_1$ is the largest eigenvalue and $\Lambda_1(1)$ is its lower bound, Eq. (9), respectively. $D$ and $A$ stand for assortative and disassortative mixing. Two last columns represent weighted networks.

| Network               | $N$   | $\gamma$ | $q_{max}$ | $q_{loc}$ | $\langle q^2 \rangle / \langle q \rangle$ | $\Lambda_1$ | $\Lambda_1(1)$ | $IPR(\Lambda_1)$ weighted |
|-----------------------|-------|----------|-----------|----------|-----------------------------------------|-------------|----------------|---------------------------|
| cond-mat 2005 [12]    | 40421 | 3.0      | 278       | 2604     | 27.35                                   | A           | 51.29          | 35.205                    | 0.0081 | 47.63 | 0.3415 |
| hep-th [12]           | 8361  | –        | 50        | 521      | 8.687                                    | A           | 23             | 10.632                    | 0.0417 | 40.52 | 0.3531 |
| astro-ph [12]         | 16706 | –        | 360       | 5415     | 44.92                                    | A           | 73.89          | 56.287                    | 0.005  | 33.7575 | 0.0525 |
| power grid [17]       | 4941  | exponential | 19   | 53       | 3.87                                    | –           | 7.483          | 3.9                    | 0.041  |
| fp5 [18]              | 27985 | 2.2      | 2942      | 38610    | 211.0                                    | –           | 197.03         | 176.3                    | 0.0035 |
| CAIDA (router-internet) [19] | 192244 | 2.7 | 1071     | 11947    | 37.89                                   | –           | 109.5          | 42.9                    | 0.010  |
| karate club [11]      | 34    | –        | 17        | 37       | 7.77                                    | D           | 6.72           | 6.01                     | 0.073  |

Real networks.—The largest eigenvalue $\Lambda_1$, $IPR(\Lambda_1)$, and other parameters of a few weighted and unweighted real-world networks are given in Table I. Note first that in all these unweighted real networks the inverse participation ratio $IPR(\Lambda_1)$ is small that evidences a delocalized $\Lambda_1$. Second, in unweighted networks, $\Lambda_1$ differs strongly from the mean-field value $\Lambda_{MF} = \langle q^2 \rangle / \langle q \rangle$. $\Lambda_1 < \Lambda_{MF}$ in networks with disassortative mixing and $\Lambda_1 > \Lambda_{MF}$ in assortative networks. Thus, degree-degree correlations strongly influence $\Lambda_1$. A similar observation has been made in [7]. Qualitatively, this result agrees with the lower bound of $\Lambda_1$ determined by Eq. (9). So we suggest that assortative mixing decreases the epidemic threshold $\lambda_c = 1/\Lambda_1$ compared to the mean-field value $1/\Lambda_{MF}$ while disassortative mixing increases $\lambda_c$.

Table I shows that in contrast to the unweighted hep-th and cond-mat-2005 networks, their weighted versions have a localized principal eigenvector with a large $IPR$. Localization occurs at vertices linked by edges with large weights. In the cond-mat-2005 network, localization occurs at vertices of degrees 37 and 28 connected by an edge with weight 34.3 that is much larger than the averaged weight $\bar{w} = 0.51$. Components of the principal eigenvectors decrease exponentially from the center of localization in agreement with the solution Eq. (14). In the hep-th network, this strong edge has weight 34 larger than $\bar{w} = 0.97$ and connects vertices of degrees 34 and 33. In unweighted networks localization does not occur because the localization threshold $q_{loc}$ obtained from the criterion Eq. (15) exceeds $q_{max}$. The prevalence $\rho(\lambda)$ in two weighted networks is represented in Fig. II(a).

In conclusion, based on a spectral approach to the SIS model, we showed that if the principal eigenvector of the adjacency matrix of a network is localized, then at the infection rate $\lambda$ right above the threshold $1/\Lambda_1$, disease is mainly localized on a finite number of vertices. Importantly, a strict epidemic threshold in this case is actually absent, and a real epidemic affecting a finite fraction of vertices occurs after a smooth crossover, at higher values of $\lambda$. On the other hand, if the principal eigenvector is delocalized, the epidemic occurs in the whole region above $\lambda_c = 1/\Lambda_1$. We suggest that further investigations of real-world networks will give many new examples of the disease localization-delocalization phenomena.

1. R. Albert and A.-L. Barabási, Rev. Mod. Phys. 74, 47 (2002); S. N. Dorogovtsev and J. F. F. Mendes, Adv. Phys. 51, 1079 (2002); M. E. J. Newman, SIAM Review 45, 167 (2003); S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. 80, 1275 (2008).
2. R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. 86, 3200 (2001); Phys. Rev. E 63 066117 (2001); M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. 90, 028701 (2003).
3. Y. Wang et al, 22nd International Symposium on Reliable Distribute Systems (SRDS03) (IEEE, 2003), p. 25; F. Chung, L. Lu, and V. Vu, PNAS 100, 6313 (2003); M. Krivelevich and B. Sudakov, Combinatorics, Probab. Comput. 12, 61 (2003); S. N. Dorogovtsev et al, Phys. Rev. E 68, 046109 (2003).
4. S. Gómez et al, Europhys. Lett. 89, 38009 (2010).
5. B. A. Prakash et al, arXiv:1004.0904.
6. C. Castellano and R. Pastor-Satorras, Phys. Rev. Lett. 105, 218701 (2010).
7. C. Castellano and R. Pastor-Satorras, arXiv:1105.5545.
8. P. Moretti et al, arXiv:1106.4215.
9. H. Minc, Nonnegative matrices (A Wiley-Interscience Publication, New York, 1988).
10. P. Van Mieghem, EPL (2012), to be published.
11. W. W. Zachary, J. Anthropol. Res. 33, 452 (1977).
12. M. E. J. Newman, PNAS 98, 404 (2001).
13. M. E. J. Newman, Phys. Rev. Lett. 89, 208701 (2002).
14. S. N. Dorogovtsev et al, Phys. Rev. E 81, 031135 (2010).
15. K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. 87, 278701 (2001).
16. D.-H. Kim and A. E. Motter, Phys. Rev. Lett. 98, 248701 (2007).
17. D. J. Watts and S. H. Strogatz, Nature 393, 440 (1998).
[18] A. Almendral et al, New J. Phys. 9, 183 (2007).

[19] CAIDA’s router-level topology measurements, http://www.caida.org/tools/measurements/skitter/router_topology/.