Ecological Modelling 220 (2009) 1034–1042

Contents lists available at ScienceDirect

Ecological Modelling

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

On the basic reproduction number and the topological properties of the contact network: An epidemiological study in mainly locally connected cellular automata

P.H.T. Schimit[a], L.H.A. Monteiro[b,a,*]

[a] Universidade de São Paulo, Escola Politécnica, Departamento de Engenharia de Telecomunicações e Controle, Av. Prof. Luciano Gualberto, travessa 3, n.380, CEP 05508-900, São Paulo, SP, Brazil

[b] Universidade Presbiteriana Mackenzie, Escola de Engenharia, Pós-graduação em Engenharia Elétrica, Rua da Consolação, n.896, CEP 01302-907, São Paulo, SP, Brazil

ARTICLE INFO

Article history:
Received 21 October 2008
Received in revised form
29 December 2008
Accepted 16 January 2009
Available online 23 February 2009

Keywords:
Basic reproduction number
Epidemiology
Ordinary differential equations
Probabilistic cellular automata
Random networks
SIR model

ABSTRACT

We study the spreading of contagious diseases in a population of constant size using susceptible-infective-recovered (SIR) models described in terms of ordinary differential equations (ODEs) and probabilistic cellular automata (PCA). In the PCA model, each individual (represented by a cell in the lattice) is mainly locally connected to others. We investigate how the topological properties of the random network representing contacts among individuals influence the transient behavior and the permanent regime of the epidemiological system described by ODE and PCA. Our main conclusions are: (1) the basic reproduction number (commonly called R_0) related to a disease propagation in a population cannot be uniquely determined from some features of transient behavior of the infective group; (2) R_0 cannot be associated to a unique combination of clustering coefficient and average shortest path length characterizing the contact network. We discuss how these results can embarrass the specification of control strategies for combating disease propagations.

1. Introduction

Models based on probabilistic cellular automata (PCA) have been employed for investigating the spreading of contagious diseases (e.g. Ahmed et al., 1998; Doran and Laffan, 2005; Fuentes and Kuperman, 1999; Monteiro et al., 2006a, 2007; Sirakoulis et al., 2000; Yakowitz et al., 1990). In these epidemiological studies, the population is divided in three groups, as in the classical SIR model proposed by Kermack and McKendrick in 1927 (e.g. Anderson and May, 1992; Murray, 2003): the group S represents the individuals that are susceptible and thus subjected to the infection; the group I is formed by the individuals that are infected and can transmit the disease for susceptible ones; and the group R is composed by the individuals that are recovered. Equivalences with models based on ordinary differential equations (ODEs) are explored in some works (e.g. Ahmed et al., 1998; Fuentes and Kuperman, 1999; Monteiro et al., 2006a, 2007), since ODE are a mean-field approximation for PCA if the three groups (S, I and R) are homogeneously distributed in space.

In PCA models, in order to represent an infection process with a short-range character, each cell in the lattice (usually corresponding to an individual in the population) interacts with two (e.g. Ahmed et al., 1998), four (e.g. Yakowitz et al., 1990) or eight surrounding neighbors (e.g. Doran and Laffan, 2005; Fuentes and Kuperman, 1999; Monteiro et al., 2006a, 2007; Sirakoulis et al., 2000). In fact, disease-causing contacts form a connection network among individuals and the assumption of a topology in which the couplings are local and regular is just a first approximation. When long-range interactions are also taken into account, random coupling topologies are frequently employed (e.g. Boccaletti et al., 2006; Kleczkowski and Grenfell, 1999; Newmann, 2002) and relations among topological properties and features of the epidemiological dynamics (for instance, the number of infective individuals in the permanent regime) are investigated (e.g. Hartvigsen et al., 2007; Pautasso and Jeger, 2008; Xu et al., 2006). Besides differences in the coupling topology, variations usually found in epidemiological studies based on PCA include: the use of non-isotropic lattices (e.g. Kansal et al., 2000); the consideration of more than one individual living in each lattice cell (e.g. Doran and Laffan, 2005); the absence of the recovered group for modelling diseases where the cure does not confer immunity (e.g. Fuentes and Kuperman, 1999); the inclusion of a latent group composed by infected individuals, but not yet infective ones (e.g. Ward et al., 2007), etc.

In this work, the random contact network is formed by C connections starting from each cell (individual) in the PCA lattice to

* Corresponding author at: Universidade Presbiteriana Mackenzie, Pós-graduação em Engenharia Elétrica, Rua da Consolação, n.896, CEP 01302-907, São Paulo, SP, Brazil. Tel.: +55 11 2114 8711; fax: +55 11 2114 8600.
E-mail addresses: pedro.schimit@poli.usp.br (P.H.T. Schimit), luizm@mackenzie.br, luizm@usp.br (L.H.A. Monteiro).

0304-3800/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2009.01.014

other cells pertaining to the square matrix of size $2r+1$ centered in such a cell; i.e., C connections start from the cell occupying the central position of its neighborhood matrix (two or more connections between the same two cells are allowed) and r is the maximum radius where a connection can be made. Of course, such a cell can receive connections starting from other cells, hence its actual number of neighbors is commonly greater than C. The case $r=1$ including all eight surrounding cells is called the Moore neighborhood of unitary radius (e.g. Wolfram, 1994). We assume that the probability q_i of settling a connection between the central cell and any cell pertaining to the layer i is given by

$$q_i = \frac{r+1-i}{r^2+r-\sum_{j=1}^{r} j} \quad (1)$$

where $i=1,2,\ldots,r$ and the layer i is formed by the cells with Moore radius equal to i. For instance, for $r=3$, then $q_1=1/2$, $q_2=1/3$, $q_3=1/6$. Thus, the probability of connecting the central cell to any of the 8 cells forming the layer $i=1$ is $1/2$, to any of the 16 cells forming the layer $i=2$ is $1/3$, and to any of the 24 cells forming the layer $i=3$ is $1/6$. In this random network, the cells are mainly locally connected. Such a topology (already used as coupling model of neural networks by Monteiro et al., 2006b) can present "high" clustering coefficient cc (here "high" means $cc \gg C/N$, where N is the number of cells constituting the lattice) and "small" average shortest path length l (here "small" means $l \sim \ln(N)/\ln(C)$) as the graphs called small-worlds (Watts and Strogatz, 1998). The main difference is that, in a small-world network, any cell has the same probability of receiving a rewired connection; in our network, this probability is given by q_i, which depends on the maximum radius r and on the distance i between the cells that may be connected.

A key parameter in epidemiological investigations, already found in the 19th-century scientific literature (Nishiura and Inaba, 2007), is the basic reproduction (reproductive) number R_0 defined as the expected number of secondary cases generated by one primary case of infection in a totally susceptible and sufficiently large population (e.g. Anderson and May, 1992; Murray, 2003). This parameter characterizes the transmission intensity of a particular disease in a particular population and it is interpreted as a threshold criterium: if $R_0>1$, the number of infected individuals increases and there occurs outbreak and/or persistence of the disease; if $R_0<1$, the number of infected individuals declines and the disease eventually disappears. Hence, determining the value of such a parameter is important to predict the temporal evolution of a disease in a population. Here we study the relations among R_0, the parameters r and C and the topological properties of the contact network generated by Eq. (1) and discuss how these results can affect the specification of strategies for controlling disease spreadings.

This manuscript is organized as follows. In Section 2, topological properties of the random connection network created by Eq. (1) are numerically investigated. In Section 3, the PCA model and the corresponding ODE model are described and analytical relations between them are derived. In Section 4, the influence of the contact network on the epidemiological dynamics is numerically examined. In Section 5, the relevance of this study is discussed.

2. Features of the PCA contact network

In the PCA model, individuals live in a toroidal surface: a square matrix formed by $n \times n = N$ cells with periodic boundary conditions (that is, the top edge of the matrix contacts the bottom edge and the left edge contacts the right edge; thus, a torus embedded in a three-dimensional space is formed from a two-dimensional plan in order to eliminate edge effects). In this section, the figures present average results obtained in ten contact networks formed by Eq. (1) for $n=200$ (the results are not qualitatively different for other values of n). Random networks are usually characterized by the degree distribution and two topological parameters (e.g. Barabasi, 2003; Boccaletti et al., 2006; Bollobàs, 2001): the clustering coefficient and the average shortest path length.

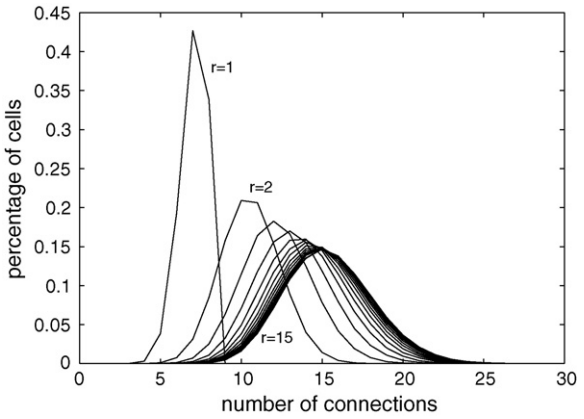

Fig. 1. Degree distribution for $1 \le r \le 15$ and $C=8$. The vertical axis indicates the percentage of the 40,000 (200×200) cells presenting the number of connections shown in the horizontal axis. For instance, for $r=2$, then about 20% of cells have 10 connections. By increasing the value of r, the distribution peak decreases, and the distribution width and the average value increase.

The degree distribution expresses how the number of connections per cell varies on the lattice. Figs. 1 and 2, which exhibit the percentage of cells with a given number of connections, show that by increasing the values of r and/or C, the distribution peak decreases and the distribution width and the average value increase.

The clustering coefficient is the typical number of connections existing between cells pertaining to a neighborhood of a cell divided by the total number of connections that could possibly exist. The higher this coefficient, the quicker the disease can spread into the neighborhood. Fig. 3 shows that this coefficient increases with C and tends to decrease with r.

The average shortest path length is the average shortest distance between any two cells composing the lattice. If this path length is small, then two cells are usually connected by just a few connections, which facilitates the disease propagation. Fig. 4 evidences that this length decreases with increasing values of r and/or C.

The influence of these topological parameters in the disease propagation are investigated in Section 4. In the next section,

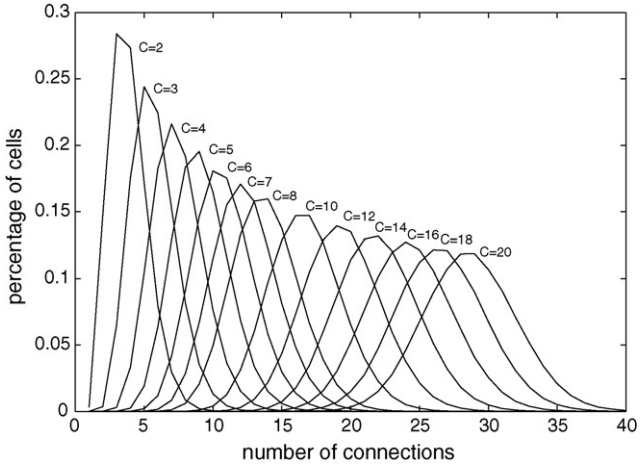

Fig. 2. Degree distribution for $r=5$ and $2 \le C \le 20$. By increasing the value of C, the distribution peak decreases, and the distribution width and the average value increase.

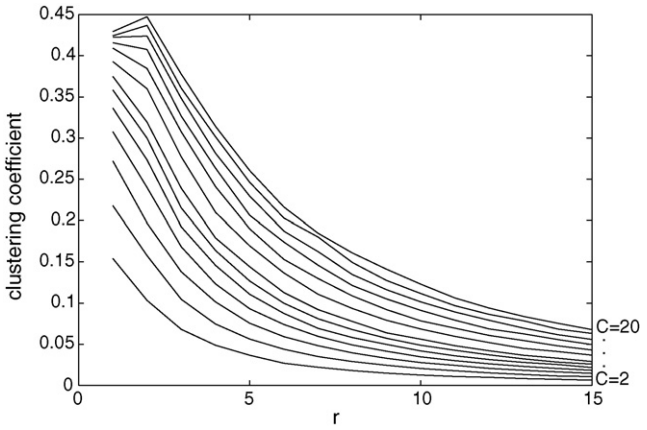

Fig. 3. Clustering coefficient for $1 \leq r \leq 15$ and $2 \leq C \leq 20$. The value of r appears in the horizontal axis and the value of C in the right edge. This topological parameter tends to decrease with r and it increases with C.

the epidemiological models based on PCA and ODE are presented.

3. PCA and ODE approaches

Each cell in the PCA lattice represents an individual that can be in one of three states: S, I or R. The time evolution of this population is ruled by the following set of probabilities of state transitions. At each time step, there is a probability P_i of a S-cell being infected according to $P_i(\nu) = (1 - e^{-k\nu})$, where ν is the number of connections with infective neighbors and k is a parameter related to the disease infectivity (observe that P_i is a monotone increasing function of ν, and that $P_i(0) = 0$, but $P_i(\nu \neq 0) < 1$). Each I-cell has probability P_c per time step of becoming cured and probability P_d per time step of dying because of the disease. At each iteration, recovered cells may die due to other causes with probability P_n. When infective and recovered individuals die, susceptible ones replace them. Therefore, the total number of individuals $N = n^2$ remains constant; an appropriate assumption for modelling diseases spreading quickly and/or populations where the deaths are balanced by the births. The states of all cells are simultaneously updated throughout a simulation.

This PCA model is a variation of the model proposed by Monteiro et al. (2006a). There an I-cell could also die by other causes and the neighborhood of a cell was restricted to its eight surrounding cells. Here an I-cell can die only because of the disease and the connections with other cells are determined by applying Eq. (1) at each time step. Thus, the neighbors may not be local and may not be the same from one time step to another.

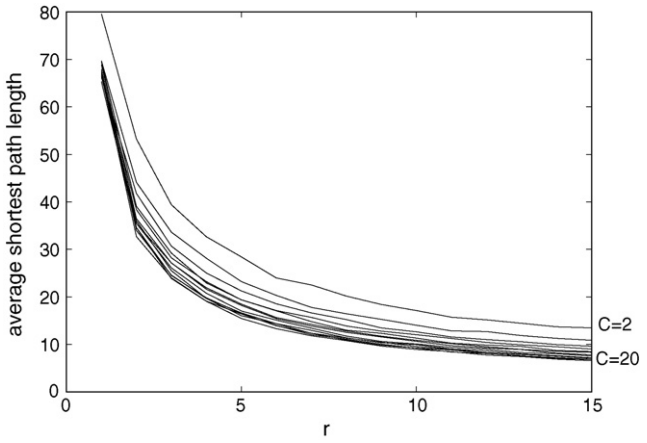

Fig. 4. Average shortest path length for $1 \leq r \leq 15$ and $2 \leq C \leq 20$. The value of r is given in the horizontal axis and the value of C in the right edge. This topological parameter decreases with r and/or C.

If S, I and R-individuals are homogeneously distributed over the space, then this PCA can be represented by the following ODE model:

$$\begin{aligned} \frac{dS(t)}{dt} &= -aS(t)I(t) + cI(t) + eR(t) \\ \frac{dI(t)}{dt} &= aS(t)I(t) - bI(t) - cI(t) \\ \frac{dR(t)}{dt} &= bI(t) - eR(t) \end{aligned} \tag{2}$$

where a is the infection rate constant; b is the recovering rate constant; c is the death rate constant related to the disease; e is the death rate constant related to other causes. Because $dS(t)/dt + dI(t)/dt + dR(t)/dt = 0$, the total number of individuals remains constant, thus: $S(t) + I(t) + R(t) = N$.

The stationary solutions (S^*, I^*, R^*) (where S^*, I^* and R^* are constants satisfying $dS(t)/dt = 0, dI(t)/dt = 0, dR(t)/dt = 0$ for any instant t) of Eq. (2) are

$$\frac{S^*}{N} = 1; \quad I^* = 0; \quad R^* = 0 \tag{3}$$

and

$$\frac{S^*}{N} = \frac{1}{R_0}; \quad \frac{I^*}{N} = \frac{e}{e+b}\left(1 - \frac{1}{R_0}\right); \quad \frac{R^*}{N} = \frac{b}{e+b}\left(1 - \frac{1}{R_0}\right) \tag{4}$$

where

$$R_0 \equiv \frac{aN}{b+c} \tag{5}$$

Stability analysis of Eq. (2)(Monteiro et al., 2006a, 2007) reveals that the disease-free stationary state given by Eq. (3) is asymptotically stable if $R_0 < 1$ and unstable if $R_0 > 1$; the endemic stationary state given by Eq. (4) is unstable if $R_0 < 1$ and asymptotically stable if $R_0 > 1$. For this ODE model, the bifurcation parameter R_0 defined by Eq. (5) corresponds to the basic reproduction number. In fact, each different epidemiological model yields a different expression to R_0 (e.g. Piqueira et al., 2004).

Because the ODE model is a mean-field approximation for the PCA model, the values of the epidemiological parameters a, b, c and e can be estimated from PCA simulations by the following expressions obtained from Eq. (2):

$$a \simeq \frac{\Delta I(t)_{S\to I}}{S(t)I(t)\Delta t} \quad b \simeq \frac{\Delta R(t)_{I\to R}}{I(t)\Delta t} \quad c \simeq \left(1 - \frac{\Delta R(t)_{I\to R}}{I(t)\Delta t}\right)\frac{\Delta S(t)_{I\to S}}{I(t)\Delta t}$$

$$e \simeq \frac{\Delta S(t)_{R\to S}}{R(t)\Delta t} \tag{6}$$

where $\Delta I(t)_{S\to I}/\Delta t$ is the increase per time step of infective individuals due to the contamination process; $\Delta R(t)_{I\to R}/\Delta t$ is the increase per time step of recovered individuals due to the healing process; $\Delta S(t)_{I\to S}/\Delta t$ is the increase per time step of susceptible individuals due to the death caused by the disease; and $\Delta S(t)_{R\to S}/\Delta t$ is the increase per time step of susceptible individuals due to the death for other causes. By assuming that the probability of a state transition at each iteration can be estimated from the relative frequency of its occurrence, then $\bar{P}_i \simeq \Delta I(t)_{S\to I}/[\Delta t S(t)]$; $P_c \simeq \Delta R(t)_{I\to R}/[\Delta t I(t)]$; $P_d \simeq \Delta S(t)_{I\to S}/[\Delta t I(t)]$; $P_n \simeq \Delta S(t)_{R\to S}/[\Delta t R(t)]$. Therefore:

$$a \simeq \frac{\bar{P}_i}{I(t)} \equiv \frac{\sum_\nu P_i(\nu) S_\nu}{I(t)\sum_\nu S_\nu} \quad b \simeq P_c \quad c \simeq (1 - P_c)P_d \quad e \simeq P_n \tag{7}$$

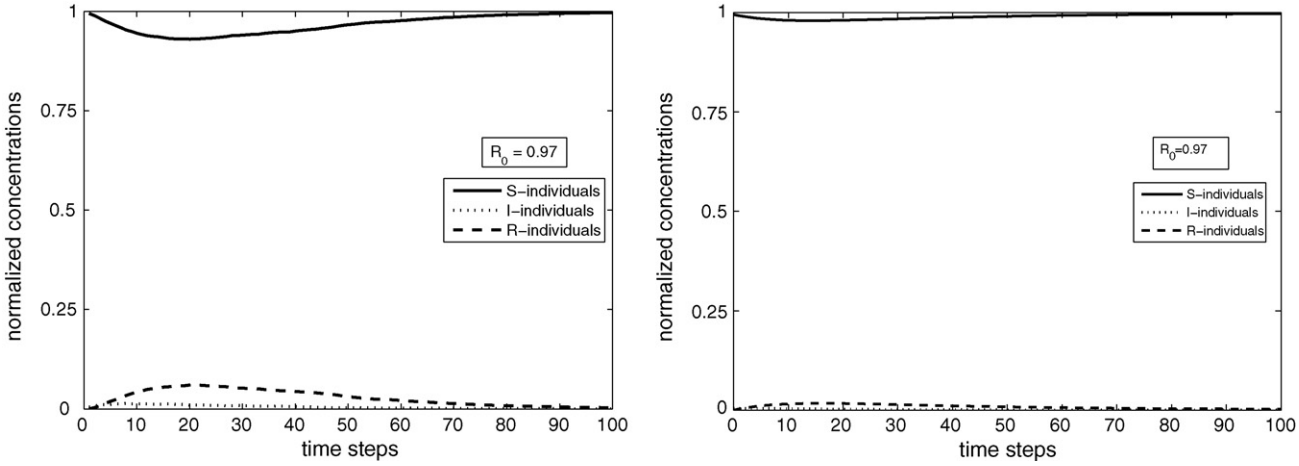

Fig. 5. PCA (left) and ODE (right) simulations for $C = 2$ and $r = 1$. As $R_0 < 1$, the disease-free stationary state is reached in the permanent regime.

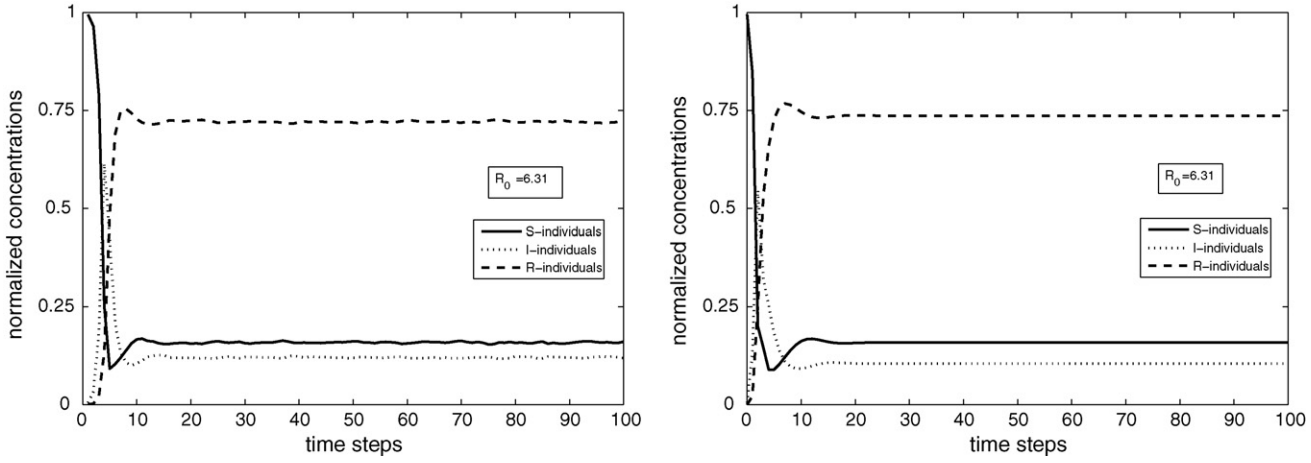

Fig. 6. PCA (left) and ODE (right) simulations for $C = 10$ and $r = 50$. As $R_0 > 1$, the endemic stationary state is reached in the permanent regime.

where S_ν is the number of susceptible cells with ν connections to infective neighbors. The expression for the parameter c is due to the fact that, at the time step t, an infective cell is firstly tested if it will be cured (with probability P_c) at $t + 1$; if not cured, then it is tested if will be dead (with probability P_d) at $t + 1$.

Thus, the epidemiological parameters b, c and e appearing in the linear terms of Eq. (2) are related to the probabilities of recovering and dying. The value of a is related to the average infection probability $\bar{P}_i$ and it is the unique epidemiological parameter influenced by r and C.

By using Eq. (6), simulations with the PCA model can be performed for estimating the values of the parameters a, b, c and e appearing in the ODE model, in order to obtain similar evolutions of $S(t)$, $I(t)$ and $R(t)$ in both approaches. Average values of these four epidemiological parameters are calculated by taking into account the last 20 time steps of a PCA simulation (when the system already reached its permanent regime). Then, these average values are substituted in Eq. (2), which are numerically solved.

Figs. 5(left) and 6(left) show the dynamical behaviors of the PCA for $n = 200$, $k = 1$, $P_c = 60\%$, $P_d = 30\%$, $P_n = 10\%$; Figs. 5(right) and 6(right) exhibit the temporal evolutions of the corresponding ODE with the parameter values determined by Eq. (6). At $t = 0$, S, I and R-cells are randomly distributed according to the proportions $S(0)/N = 99.5\%$, $I(0)/N = 0.5\%$, $R(0)/N = 0\%$. In Fig. 5(left), for $C = 2$ and $r = 1$ the corresponding numerical value of the basic reproduction number obtained from Eqs. (5) and (6) is $R_0 \simeq 0.97$; in Fig. 6(left) for $C = 10$ and $r = 50$, $R_0 \simeq 6.31$. Observe the good agreement between the two approaches.

Fig. 7 shows the asymptotically stable stationary solutions in function of R_0. As predicted from the ODE analyses, the disease only persists in the PCA simulations if $R_0 > 1$.

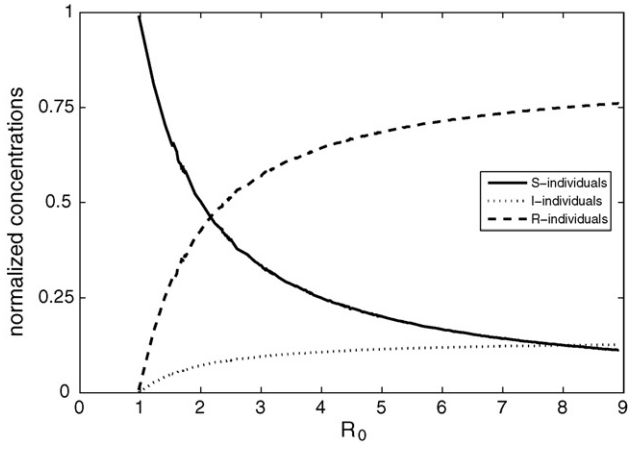

Fig. 7. Asymptotically stable values of S^*/N, I^*/N and R^*/N in function of R_0 obtained from Eqs. (3) and (4) for the same epidemiological parameter values used in Figs. 5 and 6 (i.e., $b = 0.6$, $c = 0.12$, $e = 0.1$).

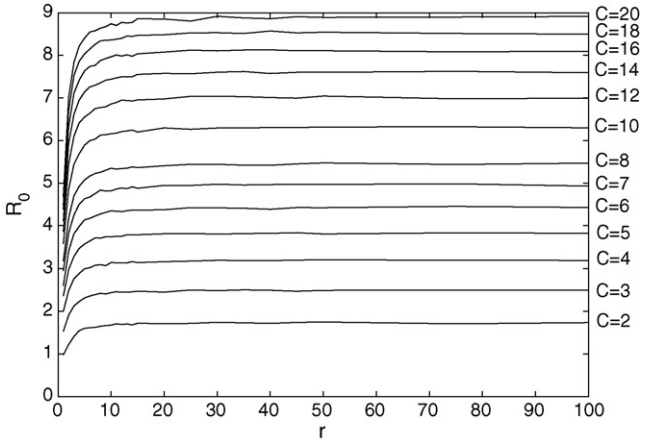

Fig. 8. R_0 for $1 \leq r \leq 100$ and $2 \leq C \leq 20$. The higher the value of r and/or C, the higher R_0.

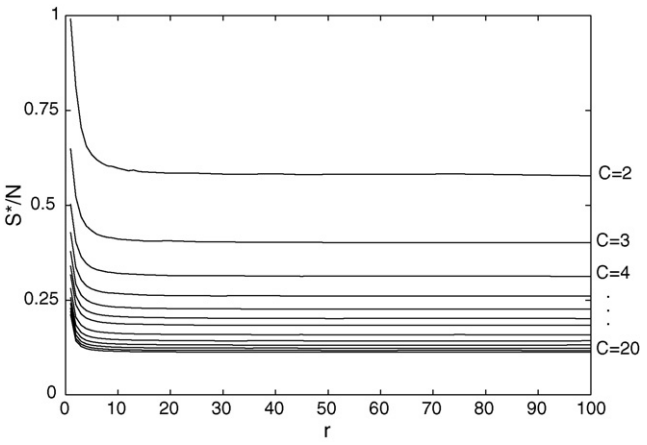

Fig. 9. S^*/N for $1 \leq r \leq 100$ and $2 \leq C \leq 20$. In the permanent regime, the susceptible group decreases with r and/or C.

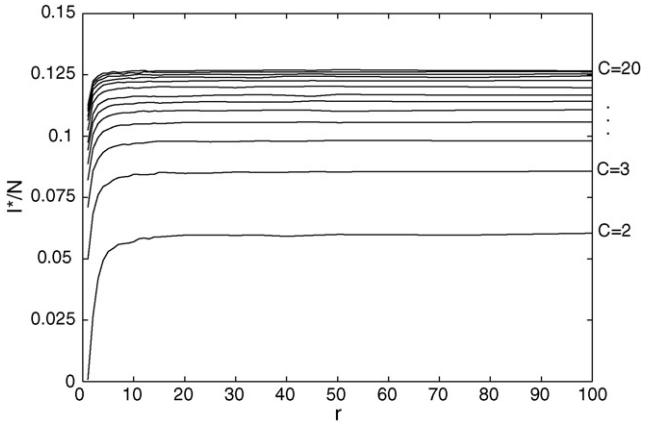

Fig. 10. I^*/N for $1 \leq r \leq 100$ and $2 \leq C \leq 20$. In the permanent regime, the infective group increases with r and/or C.

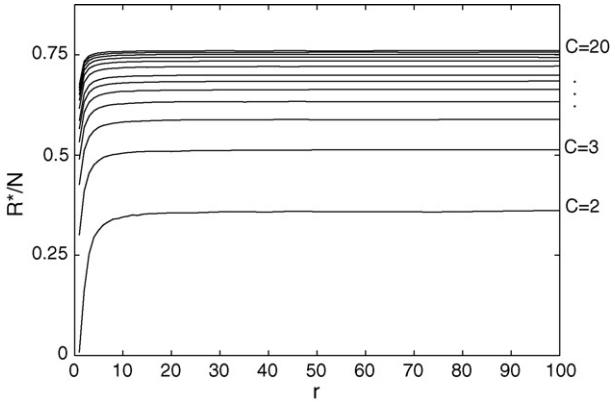

Fig. 11. R^*/N for $1 \leq r \leq 100$ and $2 \leq C \leq 20$. In the permanent regime, the recovered group increases with r and/or C.

4. Influence of the contact network on the disease dynamics

The aim here is to understand how the values of r, C and the two topological parameters described in Section 2 affect R_0, the permanent regime and the transient behavior of the PCA model. In this section, all figures show average results obtained in ten simulations with the same epidemiological parameter values used in Figs. 5–7. Thus, R_0 changes because r and/or C (and consequently a) vary. In each simulation, S, I and R-cells in the PCA lattice are randomly distributed at $t = 0$ respecting the proportions used in Figs. 5–7. Hence, the numerical values of the initial conditions are the same, but the corresponding geographical distributions may be altered from one simulation to another. And, from a time step to another, the neighborhood of each cell is recreated by following Eq. (1). Other sets of epidemiological parameter values (for instance,

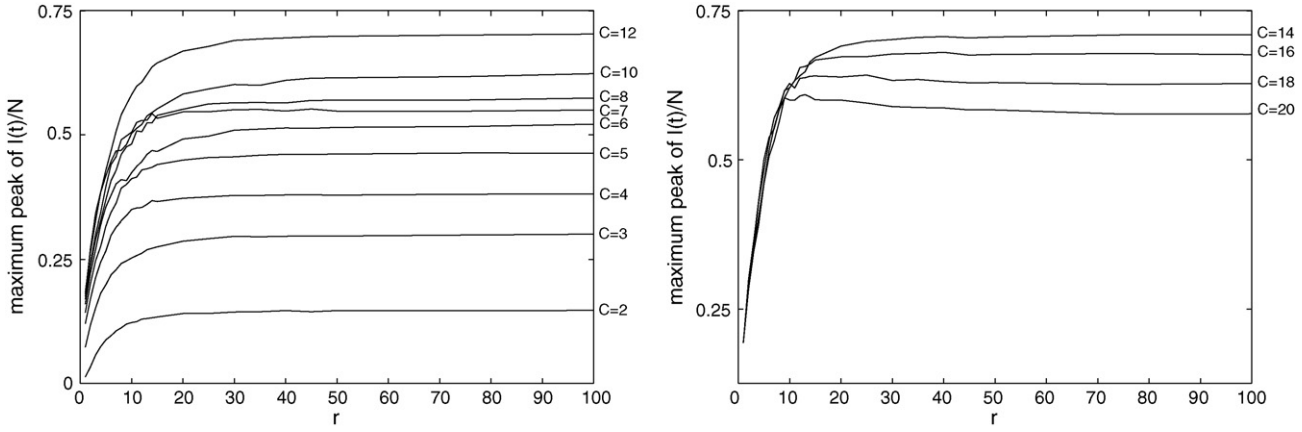

Fig. 12. Maximum peak of $I(t)/N$ in the PCA model for $1 \leq r \leq 100$ and $2 \leq C \leq 12$ (left) $14 \leq C \leq 20$ (right). The highest concentration of $I(t)/N$ initially increases with r and then reaches a saturation value, and there is a value of C for which this transient feature attains its maximum.

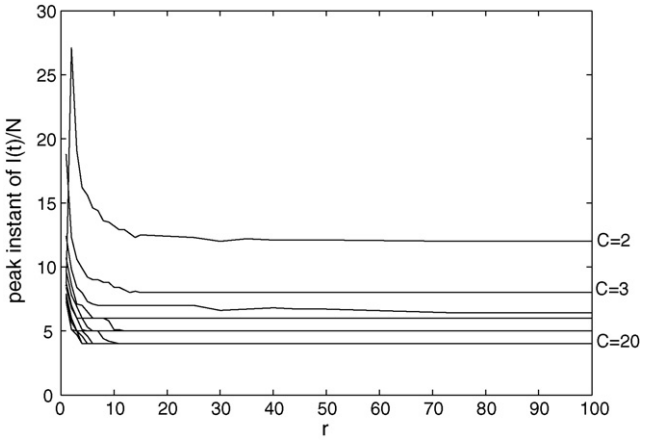

Fig. 13. Peak instant of $I(t)/N$ in the PCA model for $1 \le r \le 100$ and $2 \le C \le 12$. The instant corresponding to the highest concentration of $I(t)/N$ decreases with C, and there is a value of r (usually a low value) for which this transient feature reaches its maximum.

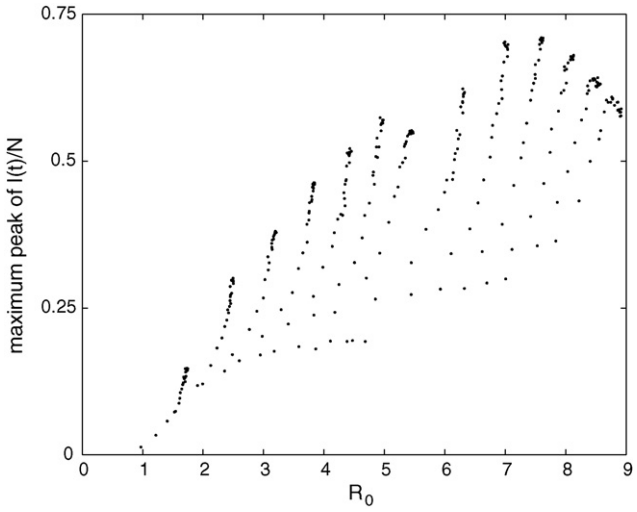

Fig. 15. Maximum peak of $I(t)/N$ in the PCA model in terms of R_0. Notice that a particular value of this transient feature can correspond to different values of R_0.

$k = 0.5$, $P_c = 80\%$, $P_d = 10\%$, $P_n = 10\%$) and/or initial conditions (for instance, $S(0)/N = 99\%$, $I(0)/N = 1\%$, $R(0)/N = 0\%$) were tested, and the results were qualitatively the same as reported below.

Fig. 8 presents R_0 in function of r and C for $1 \le r \le 100$ and $2 \le C \le 20$ (the disease frequently disappears for $C = 1$ and any r, and $C = 2$ and $r = 1$; for the other values of C and r it usually persists). Notice that R_0 increases with r and/or C (because a increases with r and/or C) and reaches a limit. Thus, the increase of the number of contacts (obtained from the parameter C) and the increase of the area where these contacts are made (related to the parameter r) help to support the persistence of a disease in a population.

Figs. 9–11 illustrate the asymptotically stable stationary concentrations of S, I and R in function of r and C for $1 \le r \le 100$ and $2 \le C \le 20$. The values of I^*/N and R^*/N grow with increases in r and/or C; the value of S^*/N diminishes. Observe that all these stationary concentrations attain limit values when the interactions among individuals are amplified. Observe also that densely connected populations present higher infective and recovered individuals than sparsely coupled populations.

Figs. 12–14 are related to some features of the transient behavior of $I(t)$ in the PCA model. These features are: the maximum peak, the instant when this peak occurs, and the settling time (here defined as the first time instant T where the average value of $I(t)$ in $T - 5 \le t \le T + 5$ does not vary more than 3% of its stationary value). Fig. 12 shows that the maximum peak increases with r and it increases with C until $C = 12$; from $C = 14$ to $C = 20$, this peak decreases as C grows. This is an interesting result because a lower maximum peak cannot be due to a less infectious disease or a more resistant population; it can be caused by increasing the number of connections beyond a critical value. Figs. 13 and 14 reveal that, after an initial growing, the peak instant and the settling time tend to decrease as r grows. The values of both these transitory features diminish when C increases.

The relation between maximum peak of $I(t)$ and R_0 is far from being trivial, as shown in Fig. 15; however, there is a tendency of increasing. Figs. 16 and 17 reveal that the peak instant and the settling time of $I(t)$ depend on R_0 in a similar way that they depend on r (Figs. 13 and 14). Thus, these both features initially increase with R_0 and then, after a critical number, they decrease and saturate. Figs. 16 and 17 show that (for the epidemiological parameter values used in these simulations) a disease with $R_0 \gtrsim 1$ and another with $R_0 \cong 3$ can present similar peak instants and similar settling times; therefore, these transitory features alone cannot be adequate for distinguishing them in terms of R_0.

Fig. 8 shows that the number R_0 increases with r and/or C (according to Eq. (5) R_0 is proportional to the epidemiological

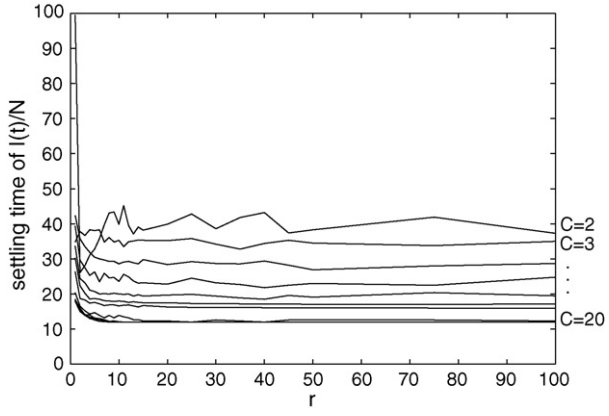

Fig. 14. Settling time of $I(t)/N$ in the PCA model for $1 \le r \le 100$ and $2 \le C \le 12$. The interval time needed for the system reaching its permanent regime tends to decrease with C, and there is a value of r (usually a low value) for which this transient feature attains its maximum.

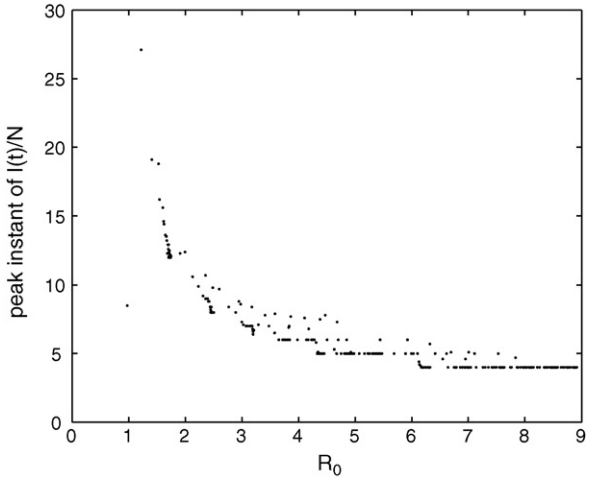

Fig. 16. Peak instant of $I(t)/N$ in the PCA model in terms of R_0. Observe that a particular value of this transient feature can correspond to distinct values of R_0.

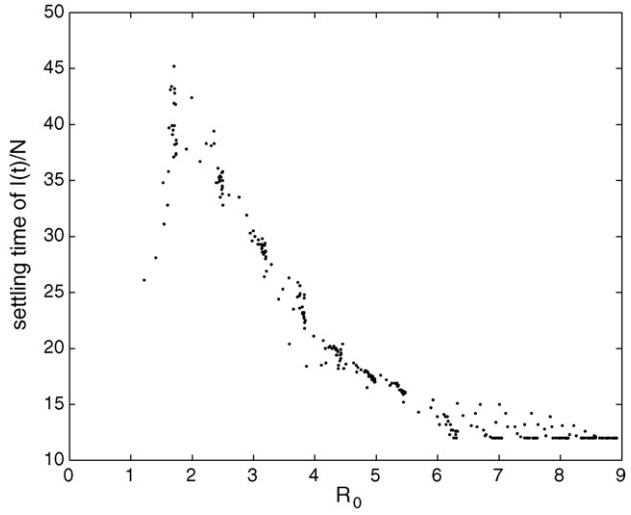

Fig. 17. Settling time of $I(t)/N$ in the PCA model in terms of R_0. Notice that a particular value of this transient feature can correspond to unalike values of R_0.

parameter a, which increases with r and/or C). Because the clustering coefficient and the average shortest path length vary with r and C as exhibited by Figs. 3 and 4, it is interesting to understand how these two topological parameters are related to R_0.

The relation between clustering coefficient and R_0 is illustrated in Fig. 18. For a fixed value of C, this coefficient decreases with R_0 because higher values of R_0 are due to higher values of r. Thus, by increasing the area where the contacts are made and keeping fixed the number of connections, the clustering coefficient must diminish, because connections between cells composing the neighborhood of a particular cell become more rare. The higher this fixed value of C, the higher the value of this coefficient. For a fixed value of r, the clustering coefficient increases with R_0 because higher values of R_0 are due to higher values of C. Therefore, by increasing the number of connections and keeping fixed the area where they are made, the clustering coefficient must increase, because more links between cells composing the neighborhood of a particular cell can be established. The higher this fixed value of r, the smaller the value of this coefficient.

Fig. 19 reveals that the average shortest path length decreases with the value of R_0, either by increasing the number of connections

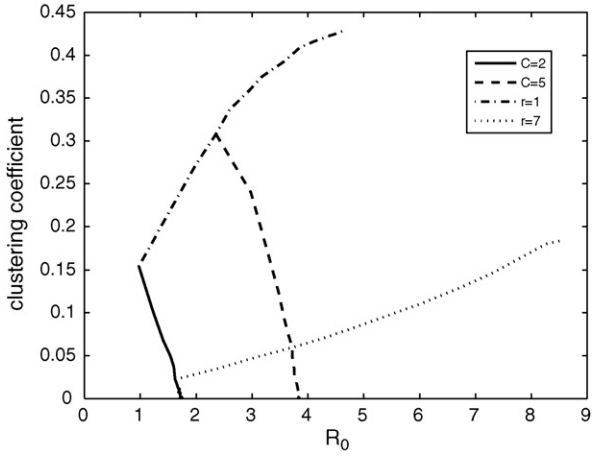

Fig. 18. Clustering coefficient in function of R_0 for some fixed values of r and C. By keeping C constant, this coefficient decreases with R_0 because higher values of R_0 correspond to higher values of r. By keeping r constant, this coefficient increases with R_0 because higher values of R_0 correspond to higher values of C.

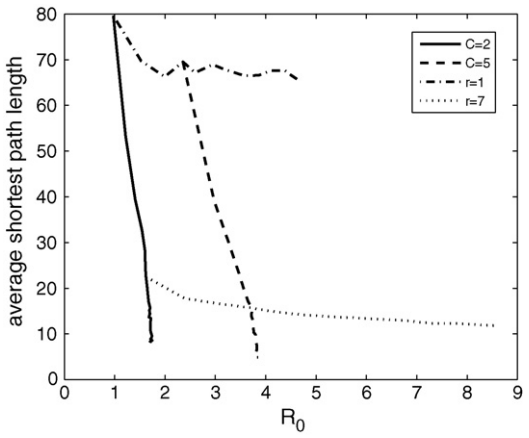

Fig. 19. Average shortest path length in function of R_0 for some fixed values of r and C. By keeping C constant, this length decreases with R_0 because higher values of R_0 correspond to higher values of r. By keeping r constant, this length also tends to decrease with R_0 because higher values of R_0 correspond to higher values of C.

or the neighborhood area. Higher values of r and/or C imply smaller path lengths between two cells.

Figs. 18 and 19 show that a same value of R_0 is compatible with distinct values of clustering coefficient (cc) and average shortest path length (l) for a given disease. For instance, a disease characterized by $R_0 = 3$ is compatible with a network with $cc \simeq 0.36$ and $l \simeq 68$ and with another network with $cc \simeq 0.05$ and $l \simeq 17$.

5. Discussion

The random coupling topology described by Eq. (1) is function of the parameters r (related to the area where connections between cells can be made) and C (expressing the number of connections starting from each cell). Fig. 8 shows that R_0, a bifurcation parameter in the dynamical system jargon, increases with r and/or C. When this parameter is above a critical value (1, by definition), then the model used for deriving the analytical expression for R_0 predicts that the disease will spread in the population. There are reports about severe epidemics occurred in slave ships during the 19th century (e.g. Watts, 1999). This is a historic example illustrating the influence of C on disease spreadings. Slaves had a very limited value of r, but the increase of C caused by gathering them in a ship increased the value of R_0 of such a population, which made easier the propagation of diseases. Cases illustrating the influence of r in disease propagations can also be easily found. For instance, sailing ships developed mainly from the 13th century amplified the human mobility on the world; consequently, the value of r and the incidence of epidemics. A classical example is the devastation of the Aztec population by smallpox in the 16th century after the Spanish invasion (e.g. Watts, 1999).

Beyond the predictive power of R_0 (questioned by some authors, e.g. Alves et al., 2003; Cross et al., 2007), its numerical value is important because it gives hints of the effort required to contain a disease spreading. For instance, Bauch and Earn (2004) proposed a game theoretical framework to analyze, in function of R_0, how the morbidity risk perception associated to voluntary vaccination policies influence the course of epidemics. Shulgin et al. (1998) used a SIR model based in ODE to determine, in terms of R_0, the minimum percentage p of susceptible individuals that must be vaccinated to eradicate a contagious disease; they found $p = (R_0 - 1)/R_0$. This expression was also derived by Huang (2008) from a SEIR model (where E represents the exposed (latent) group), who applied it for calculating the critical vaccination level associated to 21 well-known infectious diseases. A similar expression was obtained by Farrington (2003) considering the vaccine efficacy.

The value of R_0 can be unambiguously determined from the asymptotical stable stationary concentrations (as shown in Fig. 7); however, public health departments cannot wait for the SIR system reaching its permanent regime to specify the control strategies. Thus, there are attempts of estimating the value of R_0 from the transient features of $I(t)$ (e.g., Roberts and Heesterbeek, 2007). For instance, the classical SIR model can be obtained from Eq. (2) by imposing $c = 0$ and $e = 0$. In this model based on differential equations by supposing that the groups of S, I and R are homogeneously mixed, $R_0 \equiv aS(0)/b$ and the maximum peak of $I(t)$, $I_{\max}$, is given by (e.g. Murray, 2003):

$$\frac{I_{\max}}{N} = 1 - \frac{S(0)}{NR_0}[\ln(R_0) + 1] \simeq 1 - \frac{1}{R_0}[\ln(R_0) + 1] \quad (8)$$

Thus, in this model R_0 is directly related to the normalized value of the maximum peak (assuming that $S(0) \simeq N$, i.e., almost all population is initially susceptible). By using a model taking into account the spatial relations among the individuals, here we found that the maximum peak of $I(t)$, the corresponding peak instant and the settling time may not be appropriate for estimating R_0.

We also showed that a same value of R_0 can be associated to networks with distinct values of clustering coefficient and average shortest path length. This result can affect the evaluation of the effectiveness concerning different strategies employed for controlling a disease spreading. For instance, Lipsitch et al. (2003) and Zhang (2007) studied the transmission dynamics of the severe acute respiratory syndrome (SARS) by using mathematical models that do not explicitly consider the geographical localizations of the individuals, and they concluded that $R_0 \approx 3$ in the absence of control measures. Because distinct values of topological properties can produce the same value of R_0 in a model considerating the spatial structure of the contact network, it is difficult to evaluate the effective contribution of each control measure, like isolations of SARS cases, quarantine of their asymptomatic contacts and restrictions on long-range population movement. Of course, all these interventions decrease (r and C, and consequently) R_0 and help to reduce the transmission. However, the weight of each intervention in the global result is far from being understood, possibly because the correspondences among R_0 and the topological properties of the contact network are not one-to-one. Another example: Chen et al. (2006) quantified the impact of distinct control measures (such as filtration, isolation, vaccination) in containing the propagation of indoor airborne infections (like chickenpox, influenza, measles, SARS); however, they did not take into account the contact network. These authors proposed to explicitly consider this network in their model, in order to get more accurate predictions. Based in our results, we do agree with them.

Let a contagious disease and a host population be characterized by the epidemiological and topological parameters k, b, c, e, r and C (according to our SIR models). Future works should try to answer the following questions:

(1) Can this disease be eradicated by isolating a fraction F of I-individuals? Or, is there a critical value $F_c < 1$ so that if the fraction $F = F_c$ of the I-individuals is isolated, then the disease is eradicated? In other words, is this unique procedure enough to reduce the effective value of R_0 (that is, the value of R_0 for a disease propagating in a population employing control measures) below 1? Observe that isolation means strong reduction in the values of r and C for I-individuals and, consequently, alteration in the topological properties of the contact network.
(2) Preventive attitudes (as the use of masks against airborne infections or condoms against sexually transmitted diseases) alter the coupling topology of the S-individuals. Can this disease be eradicated if a fraction F of S-individuals have preventive attitudes? Is there a critical value $F_c < 1$ so that if the fraction $F = F_c$ of the S-individuals takes precaution measures, then the disease is eradicated? In other words, the effective value of R_0 can be decreased below 1 only with these attitudes?
(3) Our simulations showed that random contact networks with distinct topological parameter values can result in the same value of R_0 for the same disease. One can wonder if the two control strategies mentioned in the items (1) and (2) when separately applied against a particular disease would lead to networks with distinct topological features but with similar values of R_0.
(4) Which of these two strategies can reduce R_0 below 1 requiring the lowest value of F_c?
(5) Vaccination is frequently modelled either as a conversion of S-individuals to R-individuals or as a decreasing in the birth rate of S-individuals (e.g. Bauch and Earn, 2004; Shulgin et al., 1998). Thus, this control strategy seems not be directly related to changes in the topological properties of the contact network. However, if there exists a vaccine against the disease under analysis, what is the optimum combination among these strategies (isolation, prevention, vaccination) in order to eradicate such a disease? Here "optimum" can be interpreted as lowest cost and/or highest quality of life.

Acknowledgement

LHAM is partially supported by CNPq.

References

Ahmed, E., Agiza, H.N., Hassan, S.Z., 1998. On modeling Hepatitis B transmission using cellular automata. J. Stat. Phys. 92, 707–712.
Alves, D., Haas, V.J., Caliri, A., 2003. The predictive power of R_0 in an epidemic probabilistic system. J. Biol. Phys. 29, 63–75.
Anderson, R.M., May, R.M., 1992. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford.
Barabasi, A.L., 2003. Linked: How Everything Is Connected to Everything Else and What It Means. Plume, London.
Bauch, C.T., Earn, D.J.D., 2004. Vaccination and the theory of games. Proc. Natl. Acad. Sci. U.S.A. 101, 13391–13394.
Bollobàs, B., 2001. Random Graphs. Cambridge University Press, Cambridge.
Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U., 2006. Complex networks: structure and dynamics. Phys. Rep. 424, 175–308.
Chen, S.C., Chang, C.F., Liao, C.M., 2006. Predictive models of controls strategies involved in containing indoor airborne infections. Indoor Air 16, 469–481.
Cross, P.C., Johnson, P.L.F., Lloyd-Smith, J.O., Getz, W.M., 2007. Utility of R_0 as a predictor of disease invasion in structured populations. J. R. Soc. Interface 4, 315–324.
Doran, J.R., Laffan, S.W., 2005. Simulating the spatial dynamics of foot and mouth disease outbreaks in feral pigs and livestock in Queensland, Australia, using a susceptible-infected-recovered cellular automata model. Prev. Vet. Med. 70, 133–152.
Farrington, C.P., 2003. On vaccine efficacy and reproduction numbers. Math. Biosci. 185, 89–109.
Fuentes, M.A., Kuperman, M.N., 1999. Cellular automata and epidemiological models with spatial dependence. Physica A 267, 471–486.
Hartvigsen, G., Dresch, J.M., Zielinski, A.L., Macula, A.J., Leary, C.C., 2007. Network structure, and vaccination strategy and effort interact to affect the dynamics of influenza epidemics. J. Theor. Biol. 246, 205–213.
Huang, S.Z., 2008. A new SEIR epidemic model with applications to the theory of eradication and control of diseases, and to the calculation of R_0. Math. Biosci. 215, 84–104.
Kansal, A.R., Torquato, S., Harsh, G.R., Chiocca, E.A., Deisboeck IV, T.S., 2000. Simulated brain tumor growth dynamics using a three-dimensional cellular automaton. J. Theor. Biol. 203, 367–382.
Kleczkowski, A., Grenfell, B.T., 1999. Mean-field-type equations for spread of epidemics: the "small world" model. Physica A 274, 355–360.
Lipsitch, M., Cohen, T., Cooper, B., Robins, J.M., Ma, S., James, L., Gopalakrishna, G., Chew, S.K., Tan, C.C., Samore, M.H., Fisman, D., Murray, M., 2003. Transmission dynamics and control of severe acute respiratory syndrome. Science 300, 1966–1970.
Monteiro, L.H.A., Chimara, H.D.B., Berlinck, J.G.C., 2006a. Big cities: shelters for contagious diseases. Ecol. Model. 197, 258–262.
Monteiro, L.H.A., Paiva, D.C., Piqueira, J.R.C., 2006b. Spreading depression in mainly locally connected cellular automaton. J. Biol. Syst. 14, 629–671.

Monteiro, L.H.A., Sasso, J.B., Berlinck, J.G.C., 2007. Continuous and discrete approaches to the epidemiology of viral spreading taking into account the delay of incubation time. Ecol. Model. 201, 553–557.
Murray, J.D., 2003. Mathematical Biology I: An Introduction. Springer, New York.
Newmann, M.E.J., 2002. Spread of epidemic disease on networks. Phys. Rev. E 66, 016128.
Nishiura, H., Inaba, H., 2007. Discussion: emergence of the concept of the basic reproduction number from mathematical demography. J. Theor. Biol. 244, 357–364.
Pautasso, M., Jeger, M.J., 2008. Epidemic threshold and network structure: the interplay of probability of transmission and of persistence in small-size directed networks. Ecol. Compl. 5, 1–8.
Piqueira, J.R.C., Castaño, M.C., Monteiro, L.H.A., 2004. Modeling the spreading of HIV in homosexual populations with heterogeneous preventive attitude. J. Biol. Syst. 12, 439–456.
Roberts, M.G., Heesterbeek, J.A.P., 2007. Model-consistent estimation of the basic reproduction number from the incidence of an emerging infection. J. Math. Biol. 55, 803–816.
Shulgin, B., Stone, L., Agur, Z., 1998. Pulse vaccination strategy in the SIR epidemic model. Bull. Math. Biol. 60, 1123–1148.
Sirakoulis, G.Ch., Karafyllidis, I., Thanailakis, A., 2000. A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. Ecol. Model. 133, 209–223.
Ward, M.P., Laffan, S.W., Highfield, L.D., 2007. The potential role of wild and feral animals as reservoirs of foot-and-mouth disease. Prev. Vet. Med. 80, 9–23.
Watts, S., 1999. Epidemics and History: Disease, Power and Imperialism. Yale University Press, Yale.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of small-world networks. Nature 393, 440–442.
Wolfram, S., 1994. Cellular Automata and Complexity: Collected Papers. Westview Press, New York.
Xu, X.J., Wang, W.X., Zhou, T., Chen, G., 2006. Geographical effects on epidemic spreading in scale-free networks. Int. J. Mod. Phys. C 17, 1815–1822.
Yakowitz, S., Gani, J., Hayes, R., 1990. Cellular automaton modeling of epidemics. Appl. Math. Comput. 40, 41–54.
Zhang, Z., 2007. The outbreak pattern of SARS cases in China as revealed by mathematical model. Ecol. Model. 204, 420–426.