An analytical formula of the basic

reproduction number on cellular SIR networks

Aboubakar Sidiki [(*,**)] and Maurice Tchuente [(*,**)]

[(*)]IRD-UMI 209 UMMISCO ; [(**)]LIRIMA, équipe IDASCO
Université de Yaoundé I
Yaoundé
Cameroun
abous44@yahoo.fr, Maurice.Tchuente@ens-lyon.fr

ABSTRACT. In mathematical epidemiology, the basic reproduction number R_0 is the average number of new infections produced by an infective individual also called of generation 1, introduced in a completely susceptible population. If $R_0 < 1$, then the disease dies, whereas for $R_0 > 1$, the infection can invade the host population and persist. For SIR contact networks, one generally approximates R_0 by the average number $R_{2,3}$ of infective individuals of generation 3 produced by an infective of generation 2. We give here a simple analytic formula for $R_{2,3}$ on cellular networks. Simulations on two dimensional networks with von Neumann and Moore neighbourhoods, show that $R_{2,3} = 1$ corresponds to an epidemic threshold, and this confirms the good quality of $R_{2,3}$ as approximation of R_0 .

RÉSUMÉ. En épidémiologie mathématique, le taux de reproduction de base R_0 est le nombre moyen d'infections secondaires produites par un individu infecté introduit dans une population susceptible. Si $R_0 < 1$, alors la maladie disparait, alors que pour $R_0 > 1$, l'infection peut envahir une fraction non nulle de la population. Pour les réseaux de contact de type SIR, R_0 est généralement approximé par le nombre moyen $R_{2,3}$ d'infectés de génération 3 produits par un infecté de génération 2. Nous proposons ici une formule analytique simple de $R_{2,3}$ pour les réseaux cellulaires. Les simulations faites sur des grilles de dimension 2, à voisinage de von Neumann et de Moore, montrent que $R_{2,3} = 1$ correspond à un seuil épidémique, ce qui confirme la bonne qualité de $R_{2,3}$ comme approximation de R_0.

KEYWORDS : Basic reproduction number, cellular network, SIR model.

MOTS-CLÉS : Taux de reproduction de base, réseaux cellulaires, modle SIR.

1. Introduction

The classical SIR epidemiological model subdivides the population into three groups called compartments. Compartment S contains susceptible individuals. Compartment I corresponds to the class of infective individuals who can transmit the disease when they have adequate contacts with susceptible individuals. Compartment R corresponds to recovered individuals, i.e. who are immunized, dead or are no longer in contact with susceptible individuals. A presentation of SIR and related models may be found in [4].

Basic compartmental or mean field epidemiological models based on systems of ordinary differential equations, assume that individuals in the population are uniformly mixed. Such models usually have a Disease Free Equilibrium (DFE) corresponding to the state of the population in the absence of the disease. In most cases, there is a threshold parameter R_0 such that :

– if $R_0 < 1$, then the DFE is locally asymptotically stable, and the disease cannot invade the population, as the consequence of the introduction of a small number of infected individuals when the population is at DFE ;

– on the contrary if $R_0 > 1$, then the DFE is unstable and invasion is possible.
The mathematical definition of R_0 has a nice interpretation in terms of epidemiology. Indeed, R_0 coincides with the average number of secondary infections produced by a typical infective individual introduced in a completely susceptible population.

There are several approaches for incorporating heterogeneity in epidemiological models. In some models, the space is discrete and consists of a number of patches which are supposed to be well mixed. These patches represent social units with movements of individuals between patches. This leads to a multi-patch, multi-compartment approach that has been used for instance to propose a generalization of the classical Ross-MacDonald model which describes the dynamics of malaria [1].

In this paper, we consider contact networks, where an edge (i,j) between two vertices means that the corresponding individuals make frequent contacts with each other. In this context, the question is to study how the properties of the infectious disease and the topological properties of the contact network combine to determine the propagation of the infectious disease in the population. Andersson [5] has studied SIR epidemics on random networks. He assumed that the neighbours of a given node were randomly distributed and that, with high probability, two neighbouring nodes did not have common neighbours. With these assumptions, he showed that $R_0 = \rho(\frac{<d^2>}{<d>} - 1)$, where $< d >$ is the average vertex degree, $< d^2 >$ is the average value of the square of the vertex degrees and ρ is the spreading rate of the infectious agents.

In such SIR networks, the qualitative behaviour of the dynamics is captured by the total number of individuals who experience the disease before the end of the epidemic. The threshold phenomenon observed on the total size of the epidemic is then related to the basic reproduction number as follows : if $R_0 < 1$, then following the introduction of few infective individuals in a susceptible population, the total size of the epidemic remains small, while for $R_0 > 1$, the proportion of this total size is significant with positive probability. In this context the basic reproduction number R_0 is generally approximated by the average number $R_{2,3}$ of new infections produced by a second generation infective individual, following the introduction of one infectious individual in a completely susceptible network.

Pastor-Satorras and Vespignani [7] have studied SIS scale-free networks, i.e. networks where the distribution of probabilities that a node has exactly i neighbours follows a power law $P(i) = i^{-\alpha}$, with $2 < \alpha \leq 3$. They have shown that, as the size of the network tends to infinity, such networks are very weak in face of SIS infectious diseases, and present an effective epidemic threshold that is vanishing in the limit $n \longrightarrow \infty$, whatever is the spreading rate ρ of the infectious agents.

Later, Piccardi and Casagrandi [2] have shown that the properties exhibited by Pastor-Satorras and Verpignani were strongly linked to the underlying simple SIS compartmental model. More precisely, they have shown that scale-free networks can be unable to support diseases "with non linear force of infection whose prevalences can abruptly collapse to zero while decreasing the transmission parameters".

We are interested here in the particular case of cellular networks, i.e. regular and locally connected arrays with nodes indexed by $\mathbb{Z}^s$, and for which there is a finite set $V = \{v_1, ..., v_n\}$ called the neighbourhood index, such that each node u is connected to $u + v_1, ..., u + v_n$.

This paper is organized as follows. Section 2 presents the method proposed recently by Aparicio and Pascual for the computation of an approximate value of $R_{2,3}$, together with a refinement. Section 3 gives a simple analytical formula for $R_{2,3}$ on cellular SIR networks. Section 4 presents, for the von Neumann and Moore neighbourhoods on two-dimensional arrays, experiments that confirm the good quality of $R_{2,3}$ as approximation of R_0. In section 5 we conclude and give some directions for future work.

2. Approximate formulas for $R_{2,3}$ on finite cellular networks

A finite cellular network is a couple N=($\mathbb{Z}_L^s, V$), where s and L are positive integers, $\mathbb{Z}_L$ is the set of integers modulo L and $V = \{v_1, v_2, ..., v_n\}$ is a subset of $\mathbb{Z}_L^s$ called the neighbourhood index. Two particular neighbourhoods have been extensively studied in the literature: the von Neumann neighbourhood defined by $V_1 = \{(e_1, e_2, ..., e_s) :\mid e_1 \mid + \mid e_2 \mid + ... + \mid e_s \mid = 1\}$, and the Moore neighbourhood defined by $V_\infty = \{(e_1, e_2, ..., e_s) : \max_{1 \leq i \leq s} |e_i| = 1\}$.

Hereafter we assume that $0 \notin V$ and $s = 2$. N is then a two-dimensional array. In the von Neumann neighbourhood V_1, a node u is linked to its four nearest nodes on the left, right, up and down directions (see Figure 1a). The Moore neighbourhood V_∞ is obtained from V_1 by adding the four nearest nodes in the diagonal directions (see Figure 1b).

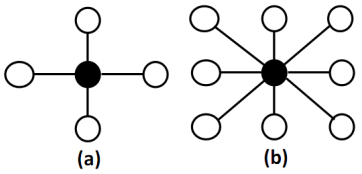

Figure 1. *von Neumann and Moore neighbourhoods*

In such networks, two types of conflicts may occur between infected individuals trying to infect a common neighbour. Two infected individuals A and B of the second generation may compete to infect a common neighbour C (see Figure 2a). This situation corresponds to cycles of length 4 in N. On the other hand, the initial infected individual may compete with a second generation infective A, for the infection of a common neighbour B (see Figure 2b). This situation happens when the network contains a cycle of length 3.

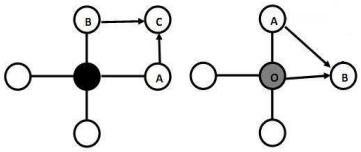

Figure 2. *Infection conflicts*

Recently, Aparicio and Pascual [8] have derived a very simple approximate formula for $R_{2,3}$ on two-dimensional arrays with Moore neighbourhood. They assume that the probability for a susceptible individual (that is in contact with infective neighbours) to become infected is ρ, irrespective of the number of infected neighbours. On the other hand, they only take into account the competition between infected individuals of generation 2 for the infection of $u \notin \Gamma(0)$, and assume that when a susceptible node v has an infected neighbour u and exactly k other neighbours with positive probabilities of infection $p_1, p_2, ..., p_k$, it becomes infected with probability ρ, and the contribution of u for its infection is $\rho/(1 + p_1 + p_2 + ... + p_k)$.

The contribution of a secondary case $u \in \Gamma(0)$ to the infection of a neighbour $v \in \Gamma(u) - \{0\}$ is then computed as follows: u is supposed to be infected, hence it has 'level' 1, all other nodes of $\Gamma(0) - \{v\}$ are supposed to be infected with 'level' ρ, and the contribution of u to the infection of v that has exactly k neighbours with 'level' ρ is $\rho/(1 + k\rho)$.

The second row of Table 1 gives the contribution of node X for the infection of nodes A, B, C and D in Figure 3a where the initial infected individual is '$\bullet$'.

	A	B	C	D
X (see Figure 3a)	ρ	$\frac{\rho}{1+\rho}$	$\frac{\rho}{1+2\rho}$	$\rho(1-\rho)$
Y (see Figure 3b)	$\frac{\rho}{1+2\rho}$	$\frac{\rho}{1+\rho}$	$\rho(1-\rho)$	$\rho(1-\rho)$

Table 1.

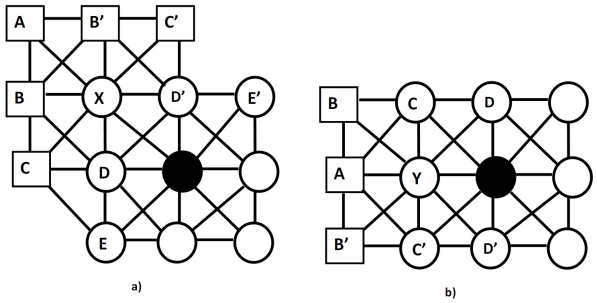

Figure 3. $\Gamma(0) \cup \Gamma(X)$ *and* $\Gamma(0) \cup \Gamma(Y)$ *for the Moore neighbourhood*

Note that for X to infect D, it is necessary that D has not been infected by the first generation infective. Since the approach does not take into account the competition for the infection of $D \in \Gamma(0)$, it follows that the contribution of X for the infection of D is $\rho(1-\rho)$, hence

$$R_{2,3,X}^{AP} = \rho + 2\frac{\rho}{1+\rho} + 2\frac{\rho}{1+2\rho} + 2\rho(1-\rho)$$

Similarly, one can compute the contribution of Y to $R_{2,3}$ (see the third row of Table 1):

$$R_{2,3,Y}^{AP} = \frac{\rho}{1+2\rho} + 2\frac{\rho}{1+\rho} + 4\rho(1-\rho)$$

In V_∞ there are 4 nodes of type X and 4 nodes of types Y. Hence [8]:

$$R_{2,3}^{AP} = \frac{\rho}{2}(7 - 6\rho + \frac{4}{1+\rho} + \frac{3}{1+2\rho})$$

Comment: Let us consider a node $v \in \Gamma^2(0) - \Gamma(0), v \neq 0$, whose neighbours $u_1, u_2, ..., u_k$ can be infected with probabilities $p_1, p_2, ..., p_k$. The probability that v is not infected is $\prod_{i=1}^{k}(1-\rho p_i)$. Hence v is infected with probability $1 - \prod_{i=1}^{k}(1-\rho p_i)$.

By removing the two first hypotheses stated by Aparicio and Pascual, the contributions of X and Y to $R_{2,3}$ are given in Table 2.

	A	B	C	D
X (see Figure 4a)	ρ	$\frac{1-(1-\rho)(1-\rho^2)}{1+2\rho}$	$\frac{1-(1-\rho)(1-\rho^2)^2}{1+2\rho}$	$\frac{(1-\rho)(1-(1-\rho)(1-\rho^2)^3)}{1+3\rho}$
Y (see Figure 4b)	$\frac{1-(1-\rho)(1-\rho^2)}{1+2\rho}$	$\frac{1-(1-\rho)(1-\rho^2}{1+\rho}$	$\frac{(1-\rho)(1-(1-\rho)(1-\rho^2))}{1+\rho}$	$\frac{(1-\rho)(1-(1-\rho)(1-\rho^2)^3)}{1+3\rho}$

Table 2.

This leads to the refined formula

$$\tilde{\mathrm{R}}_{2,3}^{AP} = \frac{\rho}{2} + \frac{2(1-(1-\rho)(1-\rho^2))}{1+\rho} + \frac{2(1-\rho)(1-(1-\rho)(1-\rho^2)^3)}{1+3\rho} + \frac{3(1-(1-\rho)(1-\rho^2)^2)}{2(1+2\rho)} + \frac{(1-\rho)(1-(1-\rho)(1-\rho^2))}{1+\rho}$$

In section 4 we will see that $\tilde{\mathrm{R}}_{2,3}^{AP}$ approximates $R_{2,3}$ better than $R_{2,3}^{AP}$.

3. Analytical formula for $R_{2,3}$ on cellular networks

Let us now consider the problem of computing $R_{2,3}$ for a general cellular network N=($\mathbb{Z}_L^s, V$). Note that V=$\Gamma(0)$. We can proceed as follows: compute, for each node u, the average number N_u of infections of generation 3 produced at u following the infection of the single node O of a completely susceptible network. Since the only sites that can experience third generation infections correspond to $W = \Gamma(V) - \{0\}$, we derive $R_{2,3} = \sum_{u \in W} N_u / (\mid V \mid \rho)$.

We are now ready to state the main result of this paper.

Theorem (Site theorem) In a cellular network N=($\mathbb{Z}_L^s, V$) with an infectious disease of transmission rate ρ, the average number of ternary cases produced by a secondary case, following the introduction of one infective individual in a completely susceptible population is:

$$R_{2,3} = \frac{\sum_{u \in W}(1-\rho)^{\delta_u}(1-(1-\rho^2)^{\alpha_u})}{\mid V \mid \rho}$$

Where $\delta_u = \begin{cases} 1 & \text{if } u \in V \\ 0 & \text{otherwise.} \end{cases}$ and $\alpha_u = \mid \Gamma(u) \cap V \mid$

Proof: In order to compute N_u for $u \in W$, we consider all possible configurations of the neighbourhood of u. We only need to consider the configurations of $\Gamma(u) \cap V$ which corresponds to potential second generation infectious that can infect u. So let us assume that $\mid \Gamma(u) \cap V \mid = r$.

Case 1: $u \in W - V$ (see for instance nodes u, v and w in Figure 4). Note that $\delta_u = 0$ and $\alpha_u = r$ Clearly, k elements of $\Gamma(u) \cap V$ are infected with probability $C_r^k \rho^k (1-\rho)^{r-k}$. Given such a configuration, the probability that u is an infective of generation 3 is $1 - (1-\rho)^k$. As a consequence

$$\begin{aligned} N_u &= \sum_{k=1}^{r} C_r^k \rho^k (1-\rho)^{r-k} (1 - (1-\rho)^k) \\ &= \sum_{k=1}^{r} C_r^k \rho^k (1-\rho)^{r-k} - \sum_{k=1}^{r} C_r^k (1-\rho)^{r-k} [\rho(1-\rho)]^k \\ &= [1 - (1-\rho)^r] - [((1-\rho) + \rho(1-\rho))^r - (1-\rho)^r] \\ &= 1 - (1-\rho)^r - [(1-\rho^2)^r - (1-\rho)^r] \\ &= 1 - (1-\rho^2)^r \\ &= (1-\rho)^{\delta_u} (1 - (1-\rho^2)^{\alpha_u}) \end{aligned}$$

Case 2: $u \in W \cap \Gamma(0)$ (see nodes u' and v' in Figure 4). Note that $\delta_u = 1$ and $\alpha_u = r$. In order to obtain N_u, we just need to multiply the expression obtained in case 1, by $(1-\rho)$ in order to express the fact that u must not be infected by node 0. This leads to $N_u = (1-\rho)(1-(1-\rho^2)^r) = (1-\rho)^{\delta_u}(1-(1-\rho^2)^{\alpha_u})$.

This shows that:

$$R_{2,3} = \frac{\sum_{u \in W} (1-\rho)^{\delta_u} (1 - (1-\rho^2)^{\alpha_u})}{\mid V \mid \rho}$$

Where $\delta_u = \begin{cases} 1 & \text{if } u \in V \\ 0 & \text{otherwise.} \end{cases}$ and $\alpha_u = \mid \Gamma(u) \cap V \mid$

Comment: We have called this "site theorem", because it is based on a method that evaluates the average number of ternary cases produced on each site $u \in W$.

Application to two-dimensional arrays

For the von Neumann neighbourhood (see Figure 4a) $\delta_u = \delta_v = 0$, $\alpha_u = 1$ and $\alpha_v = 2$. Since $\Gamma^2(0) - \{0\}$ contains 4 nodes of type u and 4 nodes of type v, it follows that: $R_{2,3} = \frac{4\rho^2 + 4(1-(1-\rho^2)^2)}{4\rho} = 3\rho - \rho^3$

For the Moore neighbourhood (see Figure 4b), we can compute Table 3.

node	u	v	w	u'	v'
δ	0	0	0	1	1
α	3	2	1	4	2
nb of nodes of this type	4	8	4	4	4

Table 3.

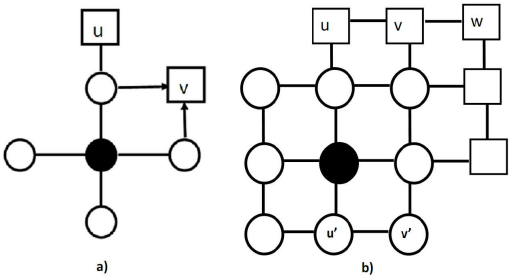

Figure 4.

Hence

$$R_{2,3} = \frac{4(1-(1-\rho^2)^3)+8(1-(1-\rho^2)^2)+4(1-(1-\rho^2))+4(1-\rho)(1-(1-\rho^2)^4)+4(1-\rho)(1-(1-\rho^2)^2)}{8\rho}$$
$$= \frac{\rho}{2}(14 - 12\rho^2 + 5\rho^4 - 6\rho + 7\rho^3 - 4\rho^5 - \rho^6 + \rho^7)$$

4. Simulations

For two-dimensional networks with von Neumann and Moore neighbourhoods, we let ρ vary in the interval $[0, 0.45]$ because, for the arrays considered, this interval contains the critical value ρ_c such that $R_{2,3}(\rho_c) = 1$. We then draw the curve $R^*_{2,3}$ obtained by simulation and compare it to R_0^{AP}, $\tilde{R}_{2,3}^{AP}$ and $R_{2,3}$. Simulations are done on a network of size 30×30. Figure 5 shows that for both networks, $R_{2,3}^{AP} \leq \tilde{R}_{2,3}^{AP} \leq R_{2,3}$, and $R_{2,3}$ is very close to $R^*_{2,3}$.

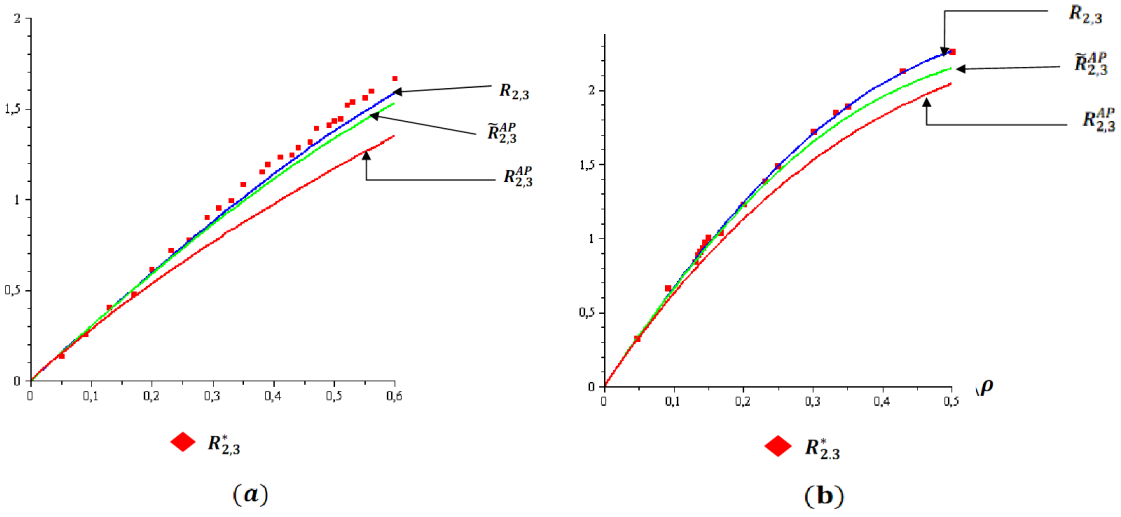

Figure 5.

Table 4 gives, for the total size of the epidemic introduced in section 1, the three maximal values obtained during the 500 simulations, for 2D arrays on von Neumann and Moore neighbourhoods. It appears clearly that, for $R_{2,3} < 1$ the total size is small, whereas for $R_{2,3} > 1$ a large outbreak occurs. This confirms the quality of $R_{2,3}$ as approximation of R_0.

$R_{2,3}$	0.6	0.7	0.8	0.9	0.98	1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
von Neumann	27	35	51	65	75	115	237	292	426	453	827	882	897	899	900
	20	33	35	61	71	96	232	258	381	445	817	879	897	898	900
	20	30	35	58	71	94	182	253	359	441	812	879	894	897	900
Moore	24	49	65	99	103	133	187	235	446	524	640	768	808	863	900
	13	41	64	90	91	120	185	234	441	524	624	734	803	850	898
	19	40	53	83	89	118	169	207	441	522	604	719	798	848	898

Table 4.

5. Discussion

In [11] it has been shown how, using the disorder parameter proposed by Watts and Strogatz [11] to rewire the connections of a cellular network, one can derive $R_{2,3}$ for some small-world networks. In social networks, communities are groups of nodes that have a high density of edges within them and low density of edges between groups [10, 9]. A great challenge is to show the impact of communities on the spread of infectious diseases. Indeed, following the introduction of an infective individual u in a susceptible network, the infection will invade primarily the community of u.

6. References

[1] P. AUGER, E. KOUOKAM, G. SALLET, M. TCHUENTE, B. TSANOU, "The Ross-Macdonald model in a patchy environment", *Mathematical Biosciences*, vol. 216, 2008.

[2] C. PICCARDI, R. CASAGRANDI, "Inefficient epidemic spreading in scale-free networks.", *Physical Review*, vol. E77, num. 026113, 2008.

[3] P. VAN DEN DRIESSCHE, J. WATMOUGH, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission", *Mathematical Biosciences*, num. 180, 2002.

[4] H.W. HETHCOTE, "The mathematics of infectious diseases", *SIAM Review*, vol. 4, num. 4, 2000.

[5] H. ANDERSSON, "Epidemics in a population with social structures", *Mathematical Biosciences*, num. 140, 1997.

[6] R. M. MAY, A.L. LLOYD, "Infection dynamics on scale free networks", *Physical Review E*, 2001.

[7] R. PASTOR-SATORRAS, VESPIGNANI, "Handbook of Graphs and Networks: From the Genome to the Internet", *S. Bornholdt and H. G. Schuster*, 2003.

[8] J. P. APARICIO, M. PASCUAL, "Building epidemiological models from R0: an implicit treatment of transmission in networks", *Proc. Roy. Soc.*, num. B 274, 2007.

[9] B. NGONMANG, M. TCHUENTE, E. VIENNET, "Local communities identification in social networks (to appear)", *Parallel Processing Letters*, 2012.

[10] S. FORTUNATO, "Community detection in graphs", *physics Reports*, vol. 486, 2010.

[11] D.J. WATTS, S.H. STRUGATZ, "Collective dynamics of small world networks", *Nature*, num. doi:10.1038/30918, vol. 393, 1998.