IJCSS International Journal of Complex Systems in Science

Int. J. Complex Systems in Science
vol. 1 (2011), pp. 47–54

Probabilistic framework for epidemic spreading in complex networks

Sergio Gómez[1,†], Alex Arenas[1], Javier Borge-Holthoefer[1], Sandro Meloni[2,3] and Yamir Moreno[3,4]

[1] *Departament dEnginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona (Spain)*

[2] *Department of Informatics and Automation, University of Rome "Roma Tre", Roma (Italy)*

[3] *Instituto de Biocomputación y Física de Sistemas Complejos, Universidad de Zaragoza, Zaragoza (Spain)*

[4] *Department of Theoretical Physics, Universidad de Zaragoza, Zaragoza (Spain)*

Abstract. The discovery of the important role played by the complex connectivity structure between individuals has lead to an increasing interest in the analysis of epidemic spreading in complex networks. Here we propose a discrete-time formulation of the problem of contact-based epidemic spreading, within the context of susceptible-infected-susceptible epidemic models. The proposed equations establish the relations between the probabilities of infection of individual nodes. They can be easily solved by iteration, showing an almost perfect agreement with Monte Carlo experiments throughout the whole phase diagram. This framework also allows the determination of the epidemic threshold and, unlike heterogeneous mean-field approaches, it is valid for any finite-size and weighted network.

Keywords: Epidemic spreading, Markov Chain, Heterogeneous Mean Field

1. Introduction

The problem of modeling how diseases spread among individuals has been intensively studied for many years [1, 2, 3, 4], allowing to address important issues such as immunization and vaccination policies [2, 5, 6]. The application of the theory of statistical physics, phase transitions and critical phenomena [7] has been extremely helpful to understand the macroscopic behavior of epidemic outbreaks [8, 9, 10, 11, 12]. This success relies basically on the Mean-Field (MF) approximation, where averages of the system reduce the degrees of freedom.

The study of complex networks [13, 14] has provided new grounds to the understanding of contagion dynamics. Particularly important in nature are scale-free (SF) networks, whose degree distribution follows a power law $P(k) \sim k^{-\gamma}$ for the number of connections, k, an individual has. SF networks include patterns of sexual contacts [15], the Internet [16], as well as other social, technological and biological networks [17]. The critical properties of an epidemic outbreak in SF networks can be addressed using the heterogeneous MF (HMF) prescription [8, 9, 10, 11, 12, 18]. It consists in the consideration that all nodes in a degree class have the same dynamical properties, and the neglection of fluctuations. Specifically, if β is the infection rate rate at which the disease spreads, the epidemic threshold in uncorrelated SF networks is given [8] by $\beta_c = \langle k \rangle / \langle k^2 \rangle$.

MF approaches are extremely useful to assess the critical properties of epidemic models, however they are not designed to give information about individual nodes. Then, questions concerning the probability that a given node be infected are not well posed in this framework. To obtain more details at the individual level of description, one has to rely on Monte Carlo (MC) simulations, which have also been used to validate the results obtained using MF methods. Here we present the Microscopic Markov-Chain Approach (MMCA) formalism first introduced in [19], and compare the epidemic spreading results with those obtained using the HMF, showing the outperformance of MMCA.

2. MMCA Formalism

Let us consider the dynamics of a susceptible-infected-susceptible (SIS) epidemic process over a N node complex network. We make no assumptions on the structure of the network, the MMCA formalism applies to any finite size, correlated, weighted and/or directed complex network. The main idea of MMCA is the discovery of the equations satisfied by the *microscopic* variables $p_i(t)$, which in this case represent the probabilities of nodes i being infected at time step t. Calling β the infection rate, μ the recovery probability and $q_i(t)$ the probability of node i not being infected by any neighbor, the SIS MMCA equations are

$$p_i(t+1) = (1-p_i(t))(1-q_i(t)) + (1-\mu)p_i(t) + \mu p_i(t)(1-q_i(t)) \,, \quad (1)$$

$$q_i(t) = \prod_{j=1}^{N} (1 - \beta r_{ji} p_j(t)) \,. \quad (2)$$

The three terms in the r.h.s. of eq. (1) account respectively for the probability that a susceptible node $(1-p_i(t))$ is infected by at least one neighbor $(1-q_i(t))$, an infected node does not recover $((1-\mu)p_i(t))$, and an infected

node recovers ($\mu p_i(t)$) but gets infected again by a neighbor ($1 - q_i(t)$). This third term may be removed if recovery and reinfection are not allowed at the same time step. In eq. (2) the r_{ji} parameters represent the probabilities of node j contacting node i. If only one contact is allowed at each time step, r_{ji} is proportional to the strength of the link (w_{ji}), whereas $r_{ji} = 1$ if all nodes are contacted per unit time. In general, allowing λ_j contacts per time step,

$$r_{ji} = 1 - \left(1 - \frac{w_{ji}}{w_j}\right)^{\lambda_j} , \tag{3}$$

where $w_j = \sum_i w_{ji}$ is the output strength of node j. The particular cases $\lambda_j = 1$ and $\lambda_j \to \infty$ correspond to the standard *contact process* (CP) and *reactive process* (RP) respectively. Other prescriptions for λ_j conform a family of models that can be obtained using this unified framework.

In the stationary state eq. (1) simplifies to

$$p_i = (1 - q_i) + (1 - \mu) p_i q_i \quad \text{(with reinfections)} , \tag{4}$$

$$p_i = (1 - p_i)(1 - q_i) + (1 - \mu) p_i \quad \text{(without reinfections)} . \tag{5}$$

These equations are easily solved by iteration until a fixed point (with reinfections) or a cycle (without reinfections) is found; in this second case, the average between the oscillating values must be considered. Finally, the average fraction of infected nodes in the stationary state is given by

$$\rho = \frac{1}{N} \sum_{i=1}^{N} p_i . \tag{6}$$

A first order approximation of eqs. (4) and (5) allows the determination of the epidemic threshold (see [19]),

$$\beta_c = \frac{\mu}{\Lambda_{\max}(R)} , \tag{7}$$

where $\Lambda_{\max}(R)$ is the largest eigenvalue of the contact matrix R. For the CP, $\beta_c = \mu$, and for the RP $\beta_c = \mu / \Lambda_{\max}(A)$, where A is the adjacency matrix of the network.

3. Results

To show the performance of the MMCA formalism we have carried out Monte Carlo (MC) simmulations of the RP on top of different networks, and compared them with the iteration solutions of the MMCA stationary equations. First, in fig. 1 we show the results for a SF network. The agreement between MMCA

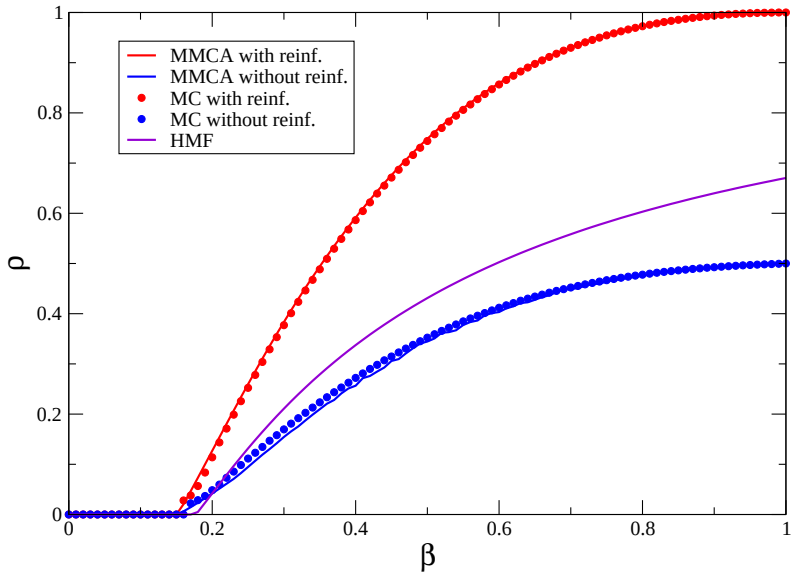

Figure 1: Comparison of MMCA, MC and HMF for a SF network of $N = 500$ and $\gamma = 2.7$, for the RP, $\mu = 1$, with and without reinfections.

and MC is matchless, both with and without reinfections. On the contrary, the numerical solutions of the HMF equations show a significant deviation in epidemic spreading, and a shift of the epidemic threshold.

Fig. 2 shows two examples where MMCA has difficulties in reproducing MC near the epidemic threshold. The first one is a regular squared grid with periodic boundary conditions, which is known to be one of the most difficult networks due to the high dynamic correlations induced by the topology. In most of the phase diagram there is good agreement between MMCA and MC, but there is an over-estimation of the epidemic spreading near the epidemic threshold. The second is a real network, the power grid [21], which is known to have an important disagreement between theory and simmulations in bond percolation processes [22]. Here the deviation between MMCA and MC is not so remarkable.

Finally, in fig. 3 we show the individual MMCA infection probabilities p_i for the air transportation network [20], using the RP and two different values of the infection rate. In this case, the agreement between MMCA and MC was already shown in [19].

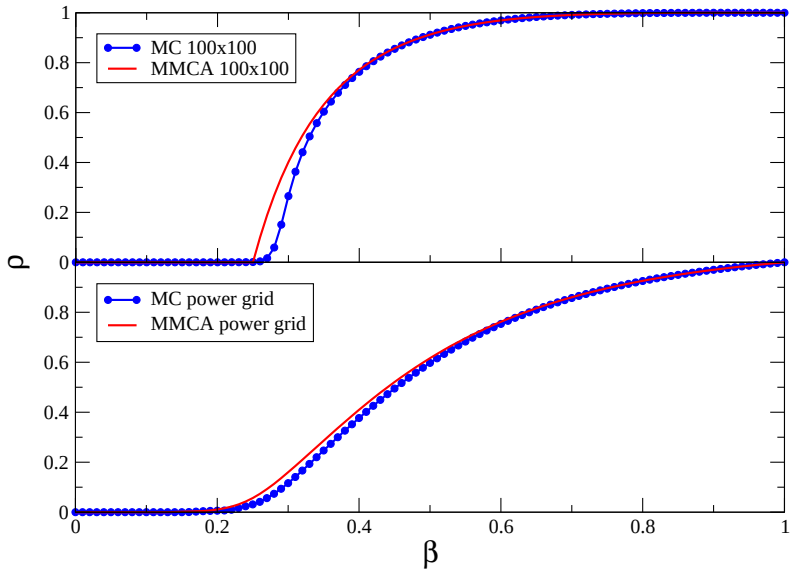

Figure 2: Comparison of MMCA and MC for a 100×100 grid (top) and the power grid (bottom), RP, $\mu = 1$ and with reinfections.

4. Conclusions

Summarizing, we have proposed a new framework to study disease spreading in networks. By defining a set of discrete-time equations for the probability of individual nodes to be infected, we construct a dynamical system that generalizes from an individual contact process to the classical case in which all connections are concurrently used, for any complex topology. Solving the equations at the stationary state, we find the whole phase diagram of the system. The numerical solution of the analytic equations overcomes the computational cost of MC simulations. Moreover, the formalism allows to gain insight on the behavior of the critical epidemic threshold for different values of the probability of contacting a fraction of neighbors per time step (λ). The method outperforms the HMF approach in the whole phase diagram.

Acknowledgements

We acknowledge the group of Prof. L. A. N. Amaral for sharing the airports data set. This work was supported by Spanish MICINN FIS2009-13730-C02-02, FIS2008- 01240 and FIS2009-13364-C02-01, and the Generalitat de Catalunya 2009-SGR-838. AA acknowledges partial support by the Director,

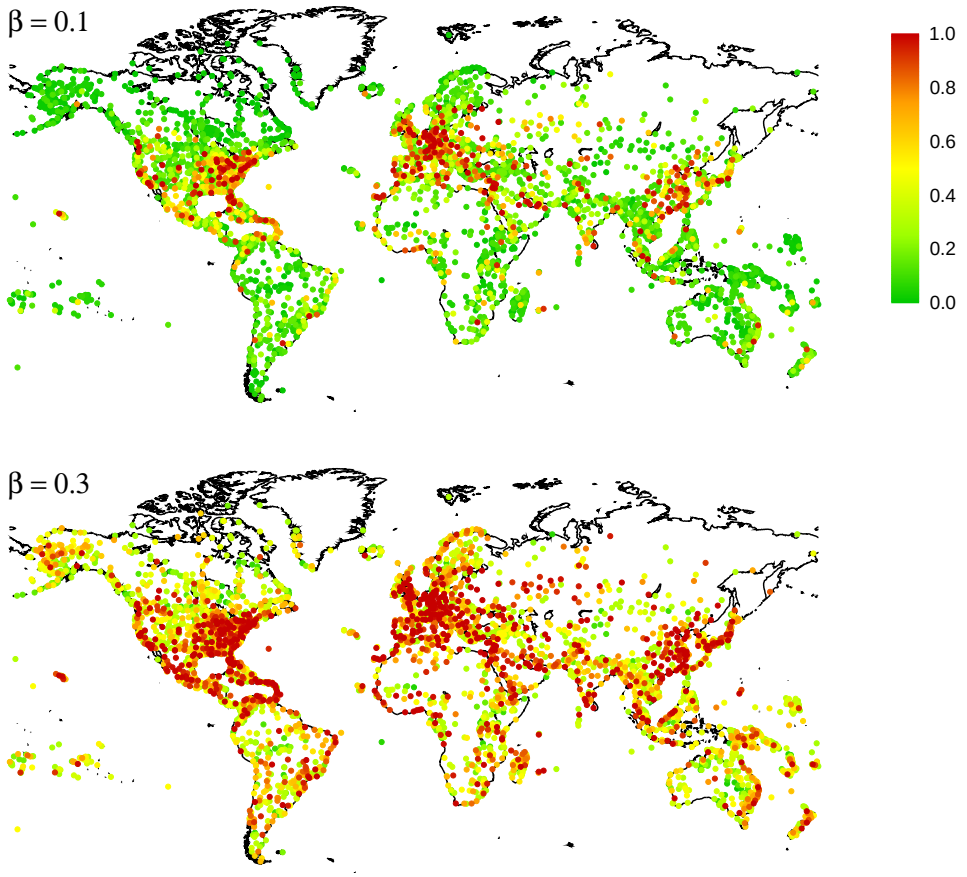

Figure 3: Individual MMCA infection probabilities for the airports network, RP, $\mu = 1$ and with reinfections.

Office of Science, Computational and Technology Research, U.S. Department of Energy under Contract DE-AC02-05CH11231. YM acknowledges support from the DGA Project PI038/08.

References

[1] H.W. Hethcote, *SIAM Review* **42**, 599 (2000).

[2] R.M. Anderson, R.M. May and B. Anderson, *Infectious diseases of humans: Dynamics and Control*, Oxford University Press, Oxford (1992).

[3] D.J. Daley and J. Gani, *Epidemic Modelling*, Cambridge University Press, UK, Cambridge (1999).

[4] J.D. Murray, *Mathematical Biology*, Springer-Verlag, Germany, Berlin (2002).

[5] L. Hufnagel, D. Brockmann and T. Geisel, *Proc. Natl. Acad. Sci. USA* **101**, 15124 (2004).

[6] S. Eubank, H. Guclu, V.S. Anil-Kumar, M.V. Marathe, A. Srinivasan, Z. Toroczkai and N. Wang, *Nature* **429**, 180 (2004).

[7] H.E. Stanley, *Introduction to Phase Transitions and Critical Phenomena*, Oxford University Press, Oxford (1987).

[8] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. Lett.* **86**, 3200 (2001).

[9] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. E* **63**, 066117 (2001).

[10] A. L. LLoyd and R. M. May, *Science* **292**, 1316 (2001).

[11] Y. Moreno, R. Pastor-Satorras and A. Vespignani, *Eur. Phys. J. B* **26**, 521 (2002).

[12] M.E.J. Newman, *Phys. Rev. E* **66**, 016128 (2002).

[13] M.E.J. Newman, *SIAM Review* **45**, 167 (2003).

[14] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).

[15] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley and Y. Aberg, *Nature* **411**, 907 (2001).

[16] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: a statistical physics approach*, Cambridge University Press, Cambridge (2004).

[17] G. Caldarelli, *Scale-Free Networks*, Oxford University Press, Oxford (2007).

[18] M. Barthélemy, A. Barrat, R. Pastor-Satorras and A. Vespignani, *Phys. Rev. Lett.* **92**, 178701 (2004).

[19] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni and Y. Moreno, *Europhys. Lett.* **89**, 38009 (2010).

[20] R. Guimerà, S. Mossa, A. Turtschi and L.A.N. Amaral, *Proc. Natl. Acad. Sci. USA* **102**, 7794 (2005).

[21] D.J. Watts and S.H. Strogatz, *Nature* **393**, 440 (1998).

[22] S. Melnik, A. Hackett, M. A. Porter, P. J. Mucha and J. P. Gleeson, arXiv:1001.1439.