The structure of coevolving infection networks

S. Wieland¹, T. Aquino¹,² and A. Nunes¹(a)

¹ Centro de Física da Matéria Condensada and Departamento de Física, Faculdade de Ciências da Universidade de Lisboa - P-1649-003 Lisboa, Portugal, EU
² Centro de Matemática e Aplicações Fundamentais, Faculdade de Ciências da Universidade de Lisboa P-1649-003 Lisboa, Portugal, EU

received 3 September 2011; accepted in final form 21 November 2011
published online 3 January 2012

PACS 89.75.Fb – Structures and organization in complex systems
PACS 87.10.Mn – Stochastic modeling
PACS 05.10.Gg – Stochastic analysis methods (Fokker-Plank, Langevin, etc.)

Abstract – Disease awareness in infection dynamics can be modeled with adaptive contact networks whose rewiring rules reflect the attempt by susceptibles to avoid infectious contacts. Simulations of this type of models show an active phase with constant infected node density in which the interplay of disease dynamics and link rewiring prompts the convergence towards a well-defined degree distribution, irrespective of the initial network topology. We develop a method to study this dynamic equilibrium and give an analytic description of the structure of the characteristic degree distribution and other network measures. The method applies to a broad class of systems and can be used to determine the steady-state topology of many other adaptive networks.

Copyright © EPLA, 2012

Introduction. – Research on the influence of network topology on the global dynamics of systems composed of many interacting units has been very active for the last ten years. More recently, adaptive and coevolving networks, that is, networks whose topology changes along with the dynamics, have begun to be considered from two points of view. The first one is to explore the consequences of this interplay from the viewpoint of global dynamics [1]. The second is to understand what sort of interactions may give rise to certain network architectures [2]. Understanding the origin of degree distributions and correlations of real networks was first addressed by considering networks that evolve independently of the dynamics [3], and it is natural to extend this idea to networks whose evolution is determined by the dynamics they support.

Infection dynamics is a natural setting to investigate the coupled evolution of the interacting elements and the network of interactions, because disease awareness translates into susceptibles that try to evade infection by changing their contact patterns according to the disease status of their neighbors. The SIS (susceptible, infected, susceptible) and the SIR (susceptible, infected, recovered) models are the simplest models for the spread of epidemic infections [4]. They are based on the idea that an infected individual infects susceptibles at a certain rate for a certain period, and then recovers and becomes susceptible again (if the disease does not confer immunity), or recovers and becomes immune. Stochastic versions of these models on networks can be mapped onto bond percolation [5].

Simple epidemic models like these exhibiting trivial global dynamics on static networks have been shown to display interesting behavior on adaptive networks, in which links between nodes may be created or removed according to rewiring rules dictated by disease awareness. In the network scheme of SIS dynamics, susceptible nodes (S-nodes, a fraction $[S]$ of the total number of nodes) are infected with rate $p$ along links connecting them to infected nodes (I-nodes, the remaining fraction $[I] = 1 - [S]$), whereas the latter recover with rate $r$. Gross et al. [6] incorporated disease awareness into the conventional SIS model by introducing a topology-changing, yet link-number preserving, rewiring mechanism: $S$-nodes try to evade infection by retracting links from infected neighbors with rate $w$, and rewiring them to randomly selected $S$-nodes. The description of the model in the framework of the standard pair approximation yields a complex phase diagram. In addition to a frozen phase (where the disease dies out) and to a stationary active phase (where $[I] \neq 0$ is constant and $[AB]$, the number per node of links connecting nodes of type $A$ and $B$, $A,B \in \{S,I\}$, is also
constant), other active phases emerge that are absent in the SIS model without rewiring, namely an oscillatory phase and a bistable phase [7]. Simulations in the stationary active phase show the states of the nodes and the links coevolving to produce and maintain a dynamic network topology characterized by well-defined degree distributions not only for the global network, but also for the subsets of S- and I-nodes. This remarkable fact and the structure of the characteristic degree distribution were left mostly unexplained. In the wake of this first paper based on SIS infection dynamics, several contributions have addressed extensions and modifications [8] of the original model, focusing mostly on the global network topology, but it is not amenable to an alternative to simulations to produce approximately steady-state distributions from the observed network topology, but it is not amenable to analytic treatment. Overall, the description of the network topology underlying the steady states of these models has received comparatively little attention. Here we develop an analytic method to compute degree distributions and other measures that characterize this network topology.

The node cycle model and its solution. — We illustrate the method taking the original model [6] as an example of an adaptive network, and focusing on the stochastic process that each node and its links follow as the former undergoes status change from susceptible to infected, and gains (in the susceptible S-stage) or loses (in the infected I-stage) links. This process can be treated analytically, and the stationary degree distributions for the full node cycle (NC), from susceptible to infected and back again, can be obtained and compared with the degree distributions of the susceptible and infective subnetworks of the original network model. The NC approach is based on the ergodic properties of the network stationary state, which are a consequence of the random rewiring process that ensures mixing. Indeed, we observed in simulations of the steady state of the network model that recording ensemble statistics yields the same distributions as sampling a single node over a sufficiently long time.

In each stage of the NC we keep track of the numbers \( x \) and \( y \) of susceptible and infected neighbors of the node (its joint degree \((x, y)\)). Due to infection and rewiring processes, these numbers change at certain rates, defining a time-homogeneous Markov process that can be described in each stage as a finite random walk on a degree grid spanned by \( x \) and \( y \) (fig. 1). The time-continuous random walks performed in each stage are one-step processes in both coordinates. They are coupled by stage transition, which takes place at constant rate \( r \) from the I-stage to the S-stage, and at rate \( py \) from the S-stage to the I-stage.

![Fig. 1: (Color online) Status and joint degree evolution of a node going through the S-stage (blue dashed line) and the I-stage (red solid line), starting with joint degree \((x_0, y_0)\) in the S-stage. Bold arrows show all possible transitions and their rates.](image)

The probability \( x, y \equiv P_S(x, y, t|x_0, y_0) \) of a walker in the S-stage, having started at coordinates \((x_0, y_0)\), being at \((x, y)\) at time \( t \) obeys the master equation

\[
\frac{d[x,y]}{dt} = (w + r)\{(y + 1)[x - 1, y + 1] - y[x, y]\}
- p y[x, y] + \tilde{w} ([x - 1, y] - [x, y])
+ \tilde{p}_S [x + 1, y - 1] - x[x, y]
\]

with the boundary conditions \([-1, y] = [x, -1] = 0\). The first term on the right-hand side of eq. (1) takes into account the swapping of an infected neighbor with a susceptible neighbor due either to rewiring or to recovery of the infected neighbor. The second term on the right represents the transition to the I-stage and an overall probability mass loss as in the active phase all susceptible nodes eventually become infected. The stochastic dynamics of the local random variables \( x \) and \( y \) is coupled to the global network dynamics through the total degree gain rate \( \tilde{w} \) and through the force of infection \( \tilde{p}_S \) that determines the infection rate of susceptible neighbors of the node. These two transitions are represented by the third and fourth terms on the right-hand side of eq. (1). A similar master equation holds for the I-stage, where instead of \( \tilde{p}_S \) the rate \( \tilde{p}_I \) accounts for the force of infection on susceptible neighbors of an I-node. We assume the correspondence parameters \( \tilde{w}, \tilde{p}_S \) and \( \tilde{p}_I \) to be constant and will assign values later on.

The master equation (1) can be solved using the generating function formalism [10]. Multiplying both sides of eq. (1) by \( \chi^x y^y \) and summing over \( x \) and \( y \) yields for the probability generating function

\[
F_S(\chi, \gamma, t|x_0, y_0) = \sum_{x,y=0}^{\infty} \chi^x y^y P_S(x, y, t|x_0, y_0)
\]

a linear first-order PDE

\[
\frac{\partial F_S}{\partial t} = \{(w + r)\chi - (w + r + p)\gamma\} \frac{\partial F_S}{\partial \gamma}
+ \tilde{p}_S(\gamma - \chi)(\partial F_S / \partial \chi) + \tilde{w}(\chi - 1).
\]

18003-p2
Its solution, given that $F_S(\chi, \gamma, 0|x_0, y_0) = \chi x_0 \gamma y_0$, is

$$F_S(\chi, \gamma, t|x_0, y_0) = e^{-\gamma t} \left[ c_1(t) + c_2(t) e^{\gamma t} \right]$$

\begin{align*}
&\times \left[ c_3(t) \right] + c_4(t) e^{\gamma t}, \\
&= \left[ c_3(t) \right] + c_4(t) e^{\gamma t}, \\
&= \left[ c_3(t) \right] + c_4(t) e^{\gamma t},
\end{align*}

(3)

where $c_j(t), j \in \{1, \ldots, 6\}$ consist of linear combinations of exponential functions of time and $c_\gamma > 0$.

The Taylor expansion of eq. (3) up to arbitrary order is easy to compute, yielding $P_S(x, y, t|x_0, y_0)$. With the remaining probability mass at time $t$ being $F_S(1, 1, t|x_0, y_0)$, the normalized total occupation time at $(x, y)$, having started at $(x_0, y_0)$, reads as

$$P_S(x, y|x_0, y_0) = \frac{\int_0^\infty P_S(x, y, t|x_0, y_0) dt}{\int_0^\infty F_S(1, 1, t|x_0, y_0) dt}.$$

Given that the rate of reinfection of a $S$-node with $y$ infected neighbours is $p y$, we calculate

$$\tilde{P}_S(x, y|x_0, y_0) = py \int_0^\infty P_S(x, y, t|x_0, y_0) dt$$

to obtain the probability of, having started at $(x_0, y_0)$, ending the walk in the $S$-stage at $(x, y)$. In a similar way, a closed-form expression for the $I$-stage can be obtained, where $\tilde{P}_I = p I$ because of recovery being coordinate-independent.

The elements of $\tilde{P}_A, A \in \{S, I\}$, are the transition probabilities that map a random walker’s initial coordinate distribution (ICD) in stage $A$, $\Phi_A(x, y)$, onto the ICD of the subsequent stage according to

$$\Phi_{I,S}(x, y) = \sum_{x_0, y_0} \tilde{P}_{S,I}(x, y|x_0, y_0) \Phi_{S,I}(x_0, y_0).$$

(4)

The ICD of the $S$-stage after a full node cycle changes according to

$$\Phi_{S}(x, y) = \sum_{x_0, y_0} \tilde{P}(x, y|x_0, y_0) \Phi_S(x_0, y_0),$$

where the $\tilde{P}(x, y|x_0, y_0)$ are given by the Chapman-Kolmogorov identity

$$\tilde{P}(x, y|x_0, y_0) = \sum_{x', y'} \tilde{P}_{I,S}(x, y|x', y') \tilde{P}_{S}(x', y'|x_0, y_0).$$

Setting a cutoff for the maximum degree, the elements of $\tilde{P}$ encode a finite ergodic Markov chain, and therefore an arbitrary ICD in the $S$-stage will converge upon iteration of the node cycle to a unique stationary state $\Phi^*_S(x, y)$ which, according to the Perron-Frobenius theorem, can be found as the right eigenvector of $\tilde{P}$ associated with the eigenvalue 1. Once $\Phi^*_S(x, y)$ are known, the normalized total occupation times of coordinates in the stationary state can be computed in either stage as

$$P^*_S(x, y) = \sum_{x_0, y_0} P_{S,I}(x, y|x_0, y_0) \Phi_{S,I}(x_0, y_0),$$

(5)

where, in particular,

$$P^*_I(x, y) = \Phi^*_S(x, y)$$

because of eq. (4) and $\hat{P}_I = P_I$. In each stage, $A \in \{S, I\}$, the average number of infected and susceptible neighbors, $\langle x_A \rangle$ and $\langle y_A \rangle$, and the average degree, $\langle k_A \rangle = \langle x_A \rangle + \langle y_A \rangle$, can be computed from the distributions $P^*_S(x, y)$.

The lifetime distributions in each stage $T_{S,I}(t, x, y)$ of random walkers with initial coordinates $(x, y)$ and the generating functions $F_{S,I}$ are related through

$$F_{S,I}(1, 1, t|x, y) = 1 - \int_0^t T_{S,I}(t', x, y) dt',$$

with overall lifetime distributions $T_{S,I}(t)$

$$T_{S,I}(t) = \sum_{x,y} \Phi^*_S(x, y) T_{S,I}(t, x, y)$$

and survival functions $L_{S,I}(t)$ in the steady state

$$L_{S,I}(t) = 1 - \int_0^t T_{S,I}(t') dt' = \sum_{x,y} \Phi^*_S(x, y) F_{S,I}(1, 1, t|x, y).$$

(6)

The average duration $\tau_{S,I}$ of each stage in the stationary state is thus given by

$$\tau_{S,I} = \sum_{x,y} \Phi^*_S(x, y) \int_0^\infty F_{S,I}(1, 1, t|x, y) dt.$$

It then follows from the solution of the analogue of eq. (2) for the $I$-stage that $\tau_I = \frac{1}{r}$ as expected, because there random walkers switch to the $S$-stage at a coordinate-independent constant rate $r$.

Correspondence with the network model and comparison with simulations. — The stationary state of the NC process was solved analytically for any choice of the parameters $w$, $r$, $p$ and $\tilde{w}$, $\tilde{p}_S$, $\tilde{p}_I$. These solutions encompass a much larger set of systems than the original network model of [6], because the correspondence parameters, tying network to NC dynamics, have yet to be determined. A first estimate for these parameters may be obtained from the equilibrium link number per node predicted by the pair approximation model of [6]. Combining this approximation with the NC description yields
\[ \tilde{w} = w \frac{|SI|}{(1 - |I|)}, \] and similar expressions hold for the other correspondence parameters. The results for the subensemble degree distributions obtained through this approach reproduce the qualitative behaviour of the network simulations but the quantitative agreement is poor (results not shown). An alternative way of constraining the correspondence parameters comes out of embedding the NC description in a given network of fixed degree. On the one hand, the correspondence parameters must be such that link depletion in the \( S \)-stage and link addition in the \( S \)-stage balance out for that degree. On the other hand, they should ensure that computing the numbers of links between nodes of type \( S \) and \( I \) does not depend on whether one sums over the infected neighbors of all the \( S \)-nodes or over the susceptible neighbors of all \( I \)-nodes. These consistency conditions will provide a self-contained determination of the correspondence parameters within the NC framework which is independent of the pair approximation and does not involve network simulations for parameter fitting.

For the NC process to represent the stationary dynamics of the network, so that the \( P_{S,I}(x,y) \) correspond to the network subensemble degree distributions and \( \tau_S \) and \( \tau_I \) are proportional to the fractions \([I]\) and \([S]\) of infected and susceptible nodes in the network, we shall therefore impose the following two conditions:

1) The average node degree must be equal to the average degree \( \langle k \rangle \) that is chosen \textit{a priori} in the network model,

\[ \tau_S \langle k_S \rangle + \tau_I \langle k_I \rangle = \langle k \rangle (\tau_S + \tau_I), \]

where \( \langle k_{S,I} \rangle \) are the average degrees computed from eq. (5).

2) The \( S \)- and \( I \)-stages represent nodes that are each other’s neighbors, and therefore \( \tau_S \langle y_S \rangle = \tau_I \langle x_I \rangle \) for the mean number of susceptible and infected neighbors in the two stages, with the averages again computed from eq. (5).

Computing the NC stationary state and choosing values of the correspondence parameters such that the consistency conditions i) and ii) hold, the NC description can be mapped onto the stationary state of a particular network model. However, the constraints set by i) and ii) may be impossible to satisfy exactly. Because of the mean field approximation involved in the assumption of constant \( \tilde{p}_S \) and \( \tilde{p}_I \), the stochastic process we have dealt with does not capture the weak correlations present in the network. This translates into i) and ii) not being exactly fulfilled, and so the description of the stationary state given by the NC is only approximate. In order to compare the analytic results with simulations on networks, we have used the free parameters \( \tilde{w}, \tilde{p}_S \) and \( \tilde{p}_I \) to minimize the squared distance of the NC output to the target values set by conditions i) and ii), bringing the stochastic model to represent as accurately as possible the state and degree evolution of a network node. We shall refer to the analytic stationary degree distributions obtained through the method outlined above for the optimal choice of \( \tilde{w}, \tilde{p}_S \) and \( \tilde{p}_I \) as the NC degree distributions. Results for a high and a low rewiring regime in a network of average degree \( \langle k \rangle = 7 \) are presented in fig. 2. For different choices of \( w, r \) and \( p \) in the stationary active phase, the NC degree distributions as well as \( \Phi^*_y(x,y) \) show very good quantitative agreement with those obtained from long Monte Carlo (MC) simulations of the network model of [6] and [9], as do the cumulative node lifetime distributions given by eq. (6) (fig. 3).

Our method predicts, through combinatorial handling of \( P_{S,I}(x,y) \), steady-state densities of any star motif. With steady-state node densities being given through

\[ \tau_S / \tau_I = (1 - [I]_{NC}) / [I]_{NC}, \]

Fig. 2: (Color online) Subensemble steady-state degree distributions for susceptibles (circles), infected (diamonds) as well as the steady-state ICD for the \( I \)-stage (squares) for two different rewiring regimes. Insets: linear plots of the distributions in their dominating degree range; comparison of steady-state prevalence \([I]_{MC}\) taken from MC simulations and \([I]_{NC}\) computed by the NC framework. Solid lines are predictions by the node cycle. Parameters \( p = 0.008, r = 0.005 \). (a) \( w = 0.025, \tilde{w} = 0.12, \tilde{p}_S = 0.044, \tilde{p}_I = 0.049 \). (b) \( w = 0.050, \tilde{w} = 0.22, \tilde{p}_S = 0.042, \tilde{p}_I = 0.045 \). Cutoff for overall degree in NC matrices is \( k_{\text{max}} = 80 \). MC simulations according to [11] with \( N = 5 \times 10^4 \) nodes, \( k = 7 \), initial Erdös-Rényi graph and initial prevalence \([I]_0 = 0.6 \). Statistics were recorded at \( \tau = 3 \times 10^4 \) for \( 10^3 \) network realizations. Error bars are smaller than markers.
The density of triplets with a central I-node connected to a S- and another I-node, normalized by the system size, is illustratively computed as

\[ [IIS]_{\text{NC}} = \sum_{x,y=0}^{\infty} \frac{x y}{\tau_{\text{NC}}} \cdot \frac{p_{x,y}^{\text{NC}}}{p_{x,y}^{\text{I}}} \]

For the parameters of fig. 2(a), this yields \([IIS]_{\text{NC}} = 3.844\), with \([IIS]_{\text{MC}} = 3.824 \pm 0.004\) from MC simulations. The small discrepancies (see fig. 2) between steady-state moment densities observed in MC simulations and those predicted by the NC method are another consequence of the constant \(\tilde{p}_S\) and constant \(\tilde{p}_I\) approximation. An improvement on this approximation would be to consider \(\tilde{p}_S\) and \(\tilde{p}_I\) to be linear functions of time chosen to represent status and degree correlations of the network model. For instance, \(\tilde{p}_I\) should increase in time along the I-stage, because older (lower degree) I-nodes have susceptible neighbors that are younger than average S-nodes, and younger S-nodes in turn have more than the average number of infected neighbors. Although it would no longer be possible in this case to find a simple solution like eq. (3) for the generating functions \(F_{S,I}\), the whole NC procedure could in principle be carried out numerically for any set of positive rate functions to yield an irreducible transition matrix \(\tilde{P}\), a unique stationary degree distribution \(\tilde{P}_{S,I}(x,y)\) for each stage, and improved approximations for all the moment densities that characterize the network’s steady state.

**Discussion and conclusions.** – Apart from giving a complete description of the network’s steady state that is in excellent quantitative agreement with the results of MC simulations, the NC method also shows that convergence of an arbitrary ICD to a unique steady-state topology determined by the parameters \(u, r, p\) and by the network’s fixed degree will take place once the rate functions associated with the correspondence parameters \(u, \tilde{p}_S\) and \(\tilde{p}_I\) are defined. These may be simply chosen as the constants that optimize the consistency conditions i) and ii), or, as mentioned above, may be taken as functions of time with more free parameters, allowing a closer match of the NC model with the network-based system. In either case, \(\tilde{P}\) encodes a time-homogeneous Markov chain, meaning that the method applies only to the network model’s stationary state and not to its transients. However, the correspondence parameters embody the mean-field approximations involved in the NC construction, and in the simplest case they are related only with two steady-state moment densities, \([S]\) and \([I]\). If, as predicted in [6] and confirmed by simulations, in the stationary active phase these densities reach equilibrium independently of network structure, then the NC approach shows that this implies that the network topology, as described by various probability distributions, will reach equilibrium as well.

Hence the pairwise model of [6] and our NC approach should be considered complementary analytic frameworks to describe approximately adaptive networks. The former allows for a complete global treatment of system dynamics, at the cost of confining the level of description to low-order densities. The latter is a local model, in that it identifies active steady states and gives a comprehensive description of their equilibrium dynamics through various characteristic distributions.

In conclusion, taking the SIS network model with rewiring of Gross et al. as an example, we developed a self-sufficient method to derive analytically the steady-state subensemble degree distributions, and other measures that characterize the unique steady state in the simple endemic phase of the model. The only requirements of the construction are that we have a cyclic two-state system, that the state changes are due to node or contact processes with constant rates, and that the network dynamics is due to link processes also with constant rates. The requirement of constant rates may be relaxed at the expense of an analytic solution for the probability generating functions not being available in this case. As for the improved approximations briefly discussed in the previous section, dealing with non-linear transition rates would require the whole NC procedure to be carried out numerically for any stage degree distribution \(\tilde{P}_{S,I}(x,y)\) for each stage, and improved approximations for all the moment densities that characterize the network’s steady state.

**Discussion and conclusions.** – Apart from giving a complete description of the network’s steady state that is
The authors would like to thank A. Parisi for many helpful discussions. Financial support from the Portuguese Foundation for Science and Technology (FCT) under Contract POCTI/ISFL/2/261 is gratefully acknowledged. SW and TA were also supported by FCT under Grant No. SFRH/BD/45179/2008 (SW) and CFTC-618-BII-02/08 (TA).

REFERENCES

[1] Holme P. and Newman M. E. J., Phys. Rev. E, 74 (2006) 056108; Fefferman N. H. and Ng K. L., Phys. Rev. E, 76 (2007) 031919; Van Segbroeck S., Santos F. C., Lenaerts T. and Pacheco J. M., Phys. Rev. Lett., 102 (2009) 058105; Szolnoki A. and Perc M., EPL, 86 (2009) 30007.

[2] Zimmermann M. G., Eguíluz V. M. and San Miguel M., Phys. Rev. E, 69 (2004) 065102; Gräser O., Xu C. and Hui P. M., EPL, 87 (2009) 38003; Vazquez F., Eguíluz V. M. and San Miguel M., Phys. Rev. Lett., 100 (2008) 108702; Iniguez G., Kertész J., Kaski K. K. and Barrio R. A., Phys. Rev. E, 80 (2009) 066119.

[3] Barabási A.-L. and Albert R., Science, 286 (1999) 509.

[4] Murray J. D., Mathematical Biology (Springer, New York) 1993.

[5] Grassberger P., Math. Biosci., 63 (1983) 157; Newman M. E. J., Phys. Rev. E, 66 (2002) 016128; Sander L. M., Warren C. P. and Sokolov I. M., Physica A, 325 (2003) 1.

[6] Gross T., Dommar D’Lima C. J. and Blasius B., Phys. Rev. Lett., 96 (2006) 208701.

[7] Gross T. and Kevrekidis I. G., EPL, 82 (2008) 38004.

[8] Bagnoli B., Lió P. and Sguanci L., Phys. Rev. E, 76 (2007) 061904; Shaw L. B. and Schwartz I. B., Phys. Rev. E, 77 (2008) 066101; Risau-Gusman S. and Zanette D. H., J. Theor. Biol., 257 (2009) 52; Funk S., Gilad E., Watkins C. and Jansen V. A. A., Proc. Natl. Acad. Sci. U.S.A., 106 (2009) 6872; ; 81 (2010) 046120; Lagorio C., Dickson M., Vazquez F., Braunstein L. A., Macri P. A., Migueles M. V., Havlin S. and Stanley H. E., Phys. Rev. E, 83 (2011) 026102.

[9] Marceau V., Noël P. A., Hébert-Dufresne L., Allard A. and Dubé L. J., Phys. Rev. E, 82 (2010) 036116.

[10] van Kampen N. G., Stochastic Processes in Physics and Chemistry (Elsevier, Amsterdam) 1981.

[11] Gillespie D. T., J. Comput. Phys., 22 (1976) 403.

[12] Demirel G., Prizak R., Reddy P. N. and Gross T., to be published in Eur. Phys. J. B (2011) http://dx.doi.org/10.1140/epjb/e2011-10844-4.