Supporting Text

Air Transportation Network Properties

We use data provided by the International Air Transport Association (IATA, www.iata.org) containing the world list of airport pairs connected by direct flights and the number of available seats on any given connection for the year 2002. More precisely, each given weight $w_{j\ell}$ in the airport network is the number of available seats on direct flight connections between the airports j and ℓ. The resulting air-transportation graph comprises $V = 3,880$ vertices denoting airports and $E = 18,810$ edges accounting for the presence of a direct flight connection. The average degree of the network is $\langle k \rangle = 2E/V = 9.70$, while the maximal degree is 318. Numerical simulations consider the $3,100$ airports with highest traffic T, which are complemented by population data of the corresponding urban area obtained using different census (publicly available sources on the Web, such as www.census.gov and unstats.un.org). This fraction corresponds to 80% of the total number of airports and carries $> 99\%$ of the total traffic.

The degree probability distribution $P(k) = n(k)/V$, where $n(k)$ is the number of airports with k connections (of degree k), is broadly distributed, with a degree range covering almost two decades, over which it can be approximated (1) by a power-law with exponent close to 2. Moreover, weights and traffic display a strong heterogeneity revealed by very broad distributions $P(w)$ and $P(T)$ spanning more than 5 orders of magnitude. It is clear that such large heterogeneities will have a strong impact on the propagation of any dynamical process on the considered network.

The traffic of a node j is defined as the sum of the weights of the links starting from j

$$T_j = \sum_{\ell \in \mathcal{V}(j)} w_{j\ell} \, , \tag{1}$$

where $\mathcal{V}(j)$ denotes the set of neighbors of node j. We can average this quantity over all nodes having the same degree and obtain the average traffic $T(k)$ for a given degree class k

$$T(k) = \frac{1}{n(k)} \sum_{j|k_j=k} T_j \ . \tag{2}$$

The plot of this quantity $T(k)$ versus k reveals a power law behavior $T(k) \sim k^\beta$ with $\beta \simeq 1.5$ (2), thus pointing to very strong correlations between topology and traffic, and meaning that the *weight per link* for each node is not constant but increases with the degree of the node.

We associate to each airport a city whose population N displays a broad probability distribution. Moreover, a nonlinear relationship between population and airport traffic is obtained, of the form $N \sim T^\alpha$ with $\alpha \simeq 0.5$.

Numerical Integration Procedure, Noise

The epidemic evolution in each city is the result of a stochastic process of contact and contamination between individuals. The elementary processes are the contamination of a susceptible by an infected, at rate β

$$S + I \xrightarrow{\beta} 2I \tag{3}$$

and the spontaneous recovery of infected individuals, at rate μ:

$$I \xrightarrow{\mu} R \ . \tag{4}$$

In the limit of large populations, the master equations describing the evolution of the probabilities to find given numbers of susceptible, infected and recovered lead to a description in terms of stochastic Langevin equations (3) of the form

$$\frac{dS_j}{dt} = -\beta\frac{I_j S_j}{N_j} + \sqrt{\beta\frac{I_j S_j}{N_j}}\,\eta_{j,1}(t) \tag{5}$$

$$\frac{dI_j}{dt} = +\beta\frac{I_j S_j}{N_j} - \mu I_j - \sqrt{\beta\frac{I_j S_j}{N_j}}\,\eta_{j,1}(t) + \sqrt{\mu I_j}\,\eta_{j,2}(t) \tag{6}$$

$$\frac{dR_j}{dt} = +\mu I_j - \sqrt{\mu I_j}\,\eta_{j,2}(t)\,, \tag{7}$$

where $\eta_{j,1}$ and $\eta_{j,2}$ are independent Gaussian white noises. The most standard numerical integration procedure (4) for solving this type of equations is to consider a small time step Δt and to rewrite these evolution equations in discretized time:

$$S_j(t+\Delta t) - S_j(t) = -\beta\frac{I_j S_j}{N_j}\Delta t + \sqrt{\beta\frac{I_j S_j}{N_j}\Delta t}\,\eta_{j,1}(t) \tag{8}$$

$$I_j(t+\Delta t) - I_j(t) = +\beta\frac{I_j S_j}{N_j}\Delta t - \mu I_j\Delta t - \sqrt{\beta\frac{I_j S_j}{N_j}\Delta t}\,\eta_{j,1}(t) + \sqrt{\mu I_j\Delta t}\,\eta_{j,2}(t) \tag{9}$$

$$R_j(t+\Delta t) - R_j(t) = +\mu I_j\Delta t - \sqrt{\mu I_j\Delta t}\,\eta_{j,2}(t)\,. \tag{10}$$

A well known problem appears within this scheme especially at the beginning of the spreading, when I_j is small: since S_j is of order N_j, the deterministic terms in Eq. **9** are of order $I_j\Delta t$ while the absolute value of the noise term is of order $\sqrt{I_j\Delta t}$ and could thus be such that I_j becomes negative, an unphysical event. Various possibilities exist to avoid this. A first naive approach would consist in setting I_j to exactly zero whenever the numerical integration yields $I_j(t+\Delta t)\leq 0$. This, however, corresponds to an asymmetric truncation of the noise which may introduce uncontrolled biases. An interesting alternative has been put forward by Dickman (4). It consists in decomposing the field X_j into its integer part $[X_j]$ and its noninteger part denoted by $\tilde{X}_j$

$$X_j(t) = [X_j](t) + \tilde{X}_j(t)\,. \tag{11}$$

For small time increments, the integer part is not varying and in addition we impose that $[X_j] = 0$ is an absorbing state (and not $X_j = 0$) meaning that we impose the absorbing constraint on the integer field which we consider as the physically relevant one. These assumptions lead to the intermediary set of equations

$$\tilde{S}_j^{temp} = \tilde{S}_j(t) - \beta\frac{[I_j][S_j]}{[N_j]}\Delta t + \sqrt{\beta\frac{[I_j][S_j]}{[N_j]}\Delta t}\,\eta_{j,1}(t) \tag{12}$$

$$\tilde{I}_j^{temp} = \tilde{I}_j(t) + \beta\frac{[I_j][S_j]}{[N_j]}\Delta t - \mu[I_j]\Delta t - \sqrt{\beta\frac{[I_j][S_j]}{[N_j]}\Delta t}\,\eta_{j,1}(t) + \sqrt{\mu[I_j]\Delta t}\,\eta_{j,2}(t) \tag{13}$$

$$\tilde{R}_j^{temp} = \tilde{R}_j(t) + \mu[I_j]\Delta t - \sqrt{\mu[I_j]\Delta t}\,\eta_{j,2}(t)\,. \tag{14}$$

and the different parts of the decomposition are then updated according to

$$[X_j](t+\Delta t) = [X_j](t) + \left[\tilde{X}_j^{temp}\right] \quad \textbf{(15)}$$

$$\tilde{X}_j(t+\Delta t) = \tilde{X}_j^{temp} - \left[\tilde{X}_j^{temp}\right] \quad \textbf{(16)}$$

[and the city sizes are updated to $N_j(t+\Delta) = S_j(t+\Delta) + I_j(t+\Delta) + R_j(t+\Delta)$]. When Δt is small enough, this procedure ensures that X_j always remain positive and integer, while $\tilde{X}_j$ is after each iteration between 0 and 1.

This treatment can be extended in order to include the transport term. The model is thus represented by a compartmental system of $3,100 \times 3$ differential equations, which describe the evolution of the numbers of susceptible, infected, and recovered individuals in each city and are coupled by the transport operator Ω. The initial condition is given by the presence of one infected individual in the city j_0 where the infection starts, while all other cities are populated by susceptible individuals only.

Heterogeneity Parameter T/N

In order to gain some analytical understanding on the time evolution of the epidemics it is convenient to consider the deterministic version of the stochastic equations that reads as

$$\frac{dS_j}{dt} = -\beta \frac{I_j S_j}{N_j} + \langle \Omega_j(\{S\}) \rangle \quad \textbf{(17)}$$

$$\frac{dI_j}{dt} = +\beta \frac{I_j S_j}{N_j} - \mu I_j + \langle \Omega_j(\{I\}) \rangle \quad \textbf{(18)}$$

$$\frac{dR_j}{dt} = +\mu I_j + \langle \Omega_j(\{R\}) \rangle. \quad \textbf{(19)}$$

These equations describes only the average behavior since they contain only the average expression of the transport operator. At the early stage of the epidemics, the number of infected individuals is relatively small in all cities and it is possible to linearize the evolution equations for the number $I_j(t)$ of infected ($S_j \simeq N_j$) as

$$\partial_t I_j = \Lambda_j I_j + \sum_\ell \frac{w_{\ell j}}{N_\ell} I_\ell \,, \quad \textbf{(20)}$$

where $\Lambda_j = \beta - \mu - T_j/N_j$. The solution of this partial differential equation can be written as the solution of the following integral equation

$$I_j(t) = I_j(0)e^{\Lambda_j t} + \sum_{\ell} \frac{w_{\ell j}}{N_\ell} \int_0^t d\tau I_\ell(\tau) e^{\Lambda_j(t-\tau)} \,. \quad \textbf{(21)}$$

This integral equation shows that the ratio T_j/N_j is a relevant variable in the determination of the time behavior of I_j and thus in the level of heterogeneity of the epidemics evolution in different cities. The underlying network structure affects the epidemics evolution by the heterogeneity of the connectivity pattern and the weight distribution through the second term of the above equation. This term contains the sum over all the connections of the ratios $w_{\ell j}/N_\ell$ and determines the number and the strength of the coupling with the infection I_ℓ of city ℓ. A heterogeneous behavior for the infection behavior might therefore find its origin both in a heterogeneous connectivity pattern as well as in a heterogeneous traffic flows distribution. This is a striking evidence of the intricate nature of the interplay between the various heterogeneities in the system, that contribute simultaneously to the dynamical behavior of the epidemic.

Heterogeneity and Parameter Values

We have investigated different values of the parameters β and μ and different initial conditions in order to test the reliability of the results obtained concerning the heterogeneity level of the spreading pattern. In Fig. 6 we show the entropy profile for three epidemic diseases starting in Hong Kong and characterized by three different values of the reproductive number R_0. We consider low to moderate transmissibility, as estimated for the early epidemic stage of the Severe Acute Respiratory Syndrome (SARS) outbreak in Hong Kong (5), with values in the range 2.2 to 3.7, 2.7 being the average value. For each value, the two null models (*HOMN* and *HETN*) and the real case (*WAN*) are shown. Changes in the parameter values clearly lead to different time scales for the global spread but do not affect the overall conclusions regarding the geographical heterogeneity of the epidemic diffusion.

We have also studied the effect of different initial conditions, such as different initial infected cities and several different initial fractions of susceptible population. This is motivated by research studies about influenza epidemics that have estimated the initial fraction of susceptibles to lie between 25 and 90%, depending on the disease strain and age group (the rest of the population is initially immune, i.e., in the class R). Here we compare results obtained with three different values of the initial percentage of susceptibles in each city, namely 100%, 80%, and 60%. The propagation time scales are affected, but the entropy profiles display the same features already discussed for the absence of initial immunity $[S(t = 0) = N]$. In all cases, *HOMN* displays a strong homogeneity and sharp transitions at the early and final stages of the epidemics, while *HETN* and *WAN* profiles are characterized by long tails and shorter homogeneous phases ($H \approx 1$), thus confirming the overall results discussed in the article (Fig. 7).

1. Guimerà, R. & Amaral, L.A.N. (2004) *Eur. Phys. J. B* **38**, 381.

2. Barrat, A., Barthélemy, M., Pastor-Satorras, R. & Vespignani, A. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 3747-3752.

3. Gardiner, W.C. (2004) *Handbook of Stochastic Methods for Physics, Chemistry and Natural Sciences* (Springer, New York), 3rd Ed.

4. Dickman, R. (1994) *Phys. Rev. E* **50**, 4404.

5. 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) *Science* **300**, 1966.