

Lecture 15.
 Thursday 19th
 November, 2020.
 Compiled: Tuesday
 8th December,
 2020.

0.1 Application of metapopulation models

In the last lecture we have introduced the so called **SIR metapopulation model**, deriving it and understanding it *analytically*. Now we want study it further, focusing on the **spatial propagation** and **predictability** under the assumption of Markovian mobility. Note that beside it, there might be many other assumptions we can make to model mobility.

0.1.1 Spatial propagation dynamics

Let us discuss the **dynamics** of spatial spreading when we are *above* the **epidemic threshold**. Given that an epidemic has started in a given city i , we want to understand how it will spread to j , *h...* Let us define the **seeding time** (or *arrival time*) as it follows: it is the time of arrival of the *first* case in patch j . Let us focus on a 2 patches model (see fig 1). Recalling that p is the travelling probability from i to j , $I(t)$ is the number of infectious in patch i , the **probability** that an infected individual arrives in j at time t is $[1 - (1 - p)^{I(t)t}]$.

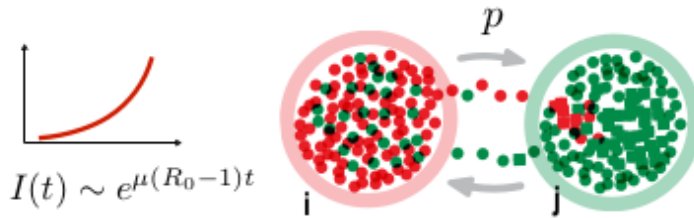


Figure 1: Dynamics of spatial spread: p is the traveling probability $i \rightarrow j$, and $I(T)$ is the prevalence in patch j .

We can now try to compute the **probability** that the **first** infectious arrives in j at time t , and it is given by:

$$P(t_{\text{seeding}} = t) = \prod_{s=1}^{t-1} (1-p)^{I(s)t} \cdot [1 - (1-p)^{I(t)t}] \xrightarrow{p \rightarrow 0} P(t_{\text{seeding}} = t) = pI(t)e^{-p \int_0^t I(s)ds} \quad (1)$$

Defining $a = \mu(R_0 - 1)$, we can rewrite the probability as:

$$P(t_{\text{seeding}} = t) = pI(t)e^{at}e^{-pae^{at}} \quad (2)$$

That is a **Gumbel distribution**, whose expected value is:

$$\langle t_{\text{seeding}} \rangle \simeq \frac{1}{a} \ln(pa) \quad (3)$$

This is the expected time needed for the infection, starting in i , to end up in j for a 2 patches model. Let us consider now the generalization to a **chain** of identical patches (see fig 2), denoting the population as N and the traveling weight among patches p .

Now it occurs that there is **correlation** among patches, since the time at which the patch i is infected does depend on the previous one, that in turn depends on the one before and so forth. Infected people that travel between patches are actually important only at the starting of each infection process for every patch, but later these can be neglected since their contribution is small compared to the normal spreading

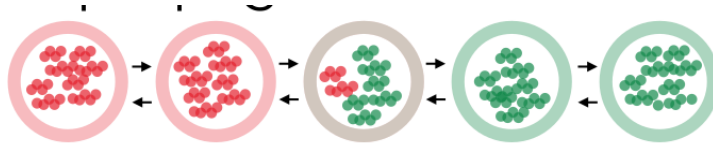


Figure 2: Dynamics of spatial spread: p is the traveling probability among patches, that are now more than two.

dynamics. We can define as Δ_i the interval that interoccurs between two consecutive seedings, obviously related to different patches:

$$\langle t_{\text{seeding},i} \rangle - \langle t_{\text{seeding},i-1} \rangle = \Delta_i, \quad \langle t_{\text{seeding},n} \rangle = \sum_{i=1}^n \Delta_i \quad (4)$$

Consequently Δ_i are correlated and not identically distributed: incidence dynamics in city i is a combination of the introductions from $i - 1$ (travelling rate, infected individual travelling...) and infection transmission with patch i . However, the simple approximations one can think of is:

$$\langle \Delta \rangle = \langle t_{\text{seeding},1} \rangle$$

does not work so bad¹.

It is straightforward to introduce a **metric** that describes how nodes are close to each other. It is function of connection “weights” (see Fig. 3) that are in turn given by the number of travellers, of flights and connections are present between a pair of nodes. In this way we are able to introduce an **effective distance**² $\ln(p_{ij})$ between nodes i and j : despite the geographical distance is way larger, it comes that New York and London are actually closer than New York and any other rural place in the MidWest using this metric.

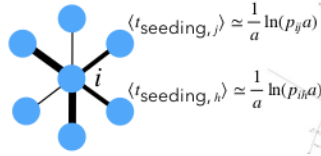


Figure 3: Effective distance between node i and other nodes connected to it. The larger the edge, the stronger the connection weight.

So, the existence of **pathways**³, through which it is more likely that the spreading of a disease can occur, comes in help, but we are requested to make risk assessment analysis for some regions. Indeed, it allows us to better **predict** the possible path through which the evolution of a spreading might occur. We now introduce the *overlap function* $\Theta(t) \in [0, 1]$ that describes the similarity between 2 outbreak realizations among different numerical simulations: the closer to 1, the more similar they are. Obviously, the **higher** the **overlap**, the **higher** the **predictability**.

As one can clearly understand from Fig. 4, introducing the **degree heterogeneity** decreases the predictability: in this way we are including hubs in our model. On the other hand, if we consider also **pathways** hence bringing in **weight fluctuations** numerical simulations are more likely to return similar outputs (the predictability is increased).

¹Gautreau et al JTB 2008.

²Brockmann, Helbing, Science 2013.

³Colizza, Barrat, Barthélemy et Vespignani, PNAS (2006).

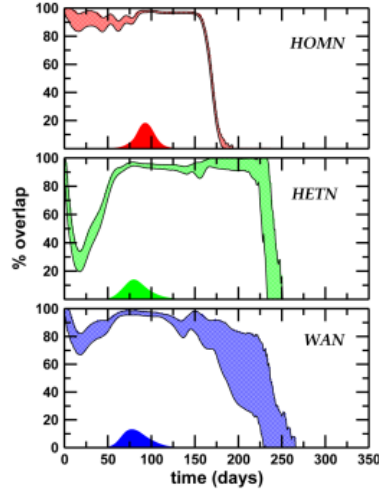


Figure 4: **Top:** numerical simulations with no degree fluctuations and no weight fluctuations. **Middle:** numerical simulations with only degree fluctuations and no weight fluctuations. **Bottom:** numerical simulations with both degree fluctuations and weight fluctuations.

One of the most important questions when dealing with epidemics and mobility is whether introducing **travel bans** would allow us to gain some time wrt the spreading of the disease. In other words we want to see if a restriction of traffic would be **effective** in either containing or delaying the propagation⁴. We therefore *rescale* the travelling probability by a factor of ω :

$$\langle t_{\text{seeding}, T.R.} \rangle \simeq \frac{1}{a} \ln(p\omega) \quad (5)$$

We can now compute how much time we “gain” when we introduce some traffic restrictions:

$$\langle t_{\text{seeding}, T.R.} \rangle - \langle t_{\text{seeding}} \rangle \simeq \frac{1}{a} \ln(p\omega) - \frac{1}{a} \ln(pa) = \frac{1}{a} \ln(\omega) \quad (6)$$

This result is indeed really important: to have a **consistent delay** in the propagation, we must **cut** the **flights** about **order of magnitudes**, with all possible consequent social and economical implications. Therefore, the moral is that this is the way *better to help the metapopulation where the epidemic started*, rather than cutting flights, being it quite useless. This would mean to try to decrease the slope of the exponential growth in the patch where the disease originated. For instance, during H1N1 pandemic in 2009, traffic volume from Mexico to Europe decreased by a factor of 40%, and the only gain in delay was really poor (order of days). Indeed, the same result is obtained thanks to numerical simulations with a global spreading model for influenza and it turn out that the delay is negligible.

We want now to briefly discuss how these arguments can be applied to **epidemic assessment**, in particular wrt **COVID-19** situation⁵. One of the first early warnings emanated by the Wuhan Municipal Health Committee was the following: “*urgent notice on the treatment of pneumonia of unknown cause*”. This was actually picked up by ProMED-mail, which is an independent program for emerging diseases, that raised an alert to the situation in Wuhan at the beginning of January. Later, some cases of pneumonia were discovered far from the epicentre among travellers, namely 2 in Thailand and 1 in Japan: compared to the 40 *local* cases and to the traffic

⁴Gautreau et al JTB 2008; Hollingsworth et al Nature Med 2006; Scalia Tomba et al Math Biosci 2008.

⁵<https://www.nytimes.com/interactive/2020/03/22/world/coronavirus-spread.html>.

volume (i.e. p to travel) this was really out of scale. This argument was actually developed by the Imperial College, stating that⁶ this new pneumonia had already started spreading worldwide and provided an estimate for the number of cases. We want now to follow the argument developed by the just mentioned report. Clearly, the number of infected that travel is $I_{trav} = I_{Wuhan} \cdot p_{travel}$, where the probability that an infectious travels p_{travel} is the product between the daily probability to travel p_{daily} and the time to detect a case T_d :

$$p_{travel} = p_{daily} T_d \quad (7)$$

where p_{daily} is the probability of travelling out of Wuhan, namely $p_{daily} = \frac{w}{N}$ and the time to detect a case T_d is the sum of two terms: *incubation period* and *time to hospitalization* (first cases were detected only after hospitalization, i.e. when symptoms started to be severe). Using real parameters, namely the passenger per day at Wuhan airport $w = 3301$, the catchment population of the just mentioned airport $N \sim 10^6$ *individuals* and the incubation period 5–6 *days* and time to hospitalization 4–5 we could infer that the estimated cases were around 1800 (95 %C.I. : 427–4471). This is much more than the 40 cases officially detected. After that particular moment, with some delay, WHO raised an alert, too, and surveillance was heightened in foreign countries. Since mobility data among countries is actually more reliable, we can use them in order to infer the epidemiological situation in the seed country.

A tool used in **numerical simulations**, for the worldwide spread of epidemics, we may ever encounter is the website <http://www.gleamviz.org/>. GLEaM⁷ is indeed a site that allows us to both visualize and simulate numerically the evolution of a pandemic using different data mobility and networks.

0.1.2 SIR metapopulation model with memory

Let us now discuss the **SIR metapopulation** model in a different regime: **commuting**^{8 9}, and not air travel any more. We therefore **drop** the **assumption** that the traveling rate is *negligible* with respect to the epidemic time scale, i.e. $p_{ij} \ll \mu$. The interoccurrence time between two trips, from now on, will be of order of *hours*: much less than the recovery rate.

In general, treating mathematically the interplay between mobility and transmission is very difficult. The problem can be solved by using **time scale separation**, that works as a sort of *mean field approximation*. Hence, either the **epidemic unfolds faster than mobility**, that is the case of flu ($\mu \sim 10^{-1} \text{ days}^{-1}$) and air traveling (rate $\sim 10^{-3} \text{ days}^{-1}$) or alternatively **mobility is faster than the epidemic**, that is the case of commuting (usually we travel twice a day, so the rate is $\sim 0.5 \text{ days}^{-1}$) and flu ($\mu \sim 10^{-1} \text{ days}^{-1}$).

We want now to take into account that, when **commuting**, we usually **return** to the place where we started. A typical trip in this case would be $i \rightarrow j \rightarrow i$. Nothing prevents us to leave once again as soon as we are back in i for k -th patch, but this is not commuting any more. The first modification we make is to divide every patch (here for simplicity we show only two in Fig. 5) distinguishing on the individuals and whether they are resident in i or j . In this way we can consider people that, despite they live in another patch i , they have travelled to patch j with a certain **leaving rate** σ_{ij} and will return home according to a **returning rate** $\rho^{-1} = \tau \approx 8h$ that is independent of the destination. $N_{ij}(t)$ are the individuals resident in i and

⁶<https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-1-case-estimates-of-covid-19/>

⁷Global Epidemic and Mobility Model.

⁸Sattenspiel, L. & Dietz, K. Math. Biosci. 128, 71–91 (1995)

⁹Keeling, M. J. & Rohani, P. Ecol. Lett. 5, 20–29 (2002)

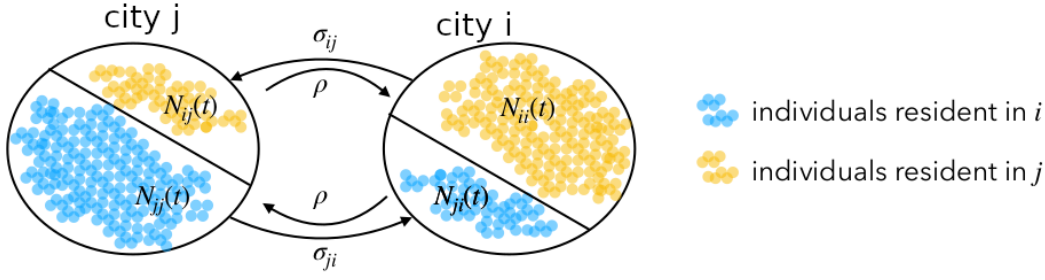


Figure 5: Graphical representation for the SIR metapopulation model with memory, this is actually more realistic since we usually back and forth in such trips home \leftrightarrow workplace. We distinguish the population inside patches on their residence place.

traveling to j , while we approximate the **population** for a given patch i as *constant*: $N_i = N_{ii}(t) + \sum_j N_{ij}(t)$. In other words we are assuming no immigration and no births nor deaths. In this way we can keep track of from where and individual left.

We want now to consider the the change in time of people resident in i and either staying in i or leaving for j :

$$\partial_t N_{ii} = - \sum_j \sigma_{ij} N_{ii}(t) + \rho \sum_j N_{ij}(t) \quad (8a)$$

$$\partial_t N_{ij} = \sigma_{ij} N_{ii}(t) - \rho N_{ij}(t) \quad (8b)$$

The linear ordinary **differential equation** of order 1 to be solved can be rewritten as:

$$\partial_t N_{ii}(t) + (\rho + \sigma_i) N_{ii}(t) = N_i \rho \quad (9)$$

The first solution is:

$$\begin{aligned} N_{ii}(t) &= e^{(\rho+\sigma_i)t} \left(C_{ii} + N_i \rho \int_0^t e^{-(\rho+\sigma_i)s} ds \right) = \\ &= \frac{N_i}{1 + \sigma_i/\rho} + \left(N_{ii}(0) - \frac{N_i}{1 + \sigma_i/\rho} \right) e^{-\rho(1+\sigma_i/\rho)t} \end{aligned} \quad (10)$$

While for the individuals resident in i and travelling to j :

$$\begin{aligned} N_{ij}(t) &= \frac{\sigma_{ij} N_i / \rho}{1 + \sigma_i / \rho} - \frac{\sigma_{ij}}{\sigma_i} \left(N_{ii}(0) - \frac{N_i}{1 + \sigma_i / \rho} \right) e^{-\rho(1+\sigma_i/\rho)t} + \\ &+ \left[N_{ii}(0) - \frac{\sigma_{ij} N_i / \rho}{1 + \sigma_i / \rho} - \frac{\sigma_{ij}}{\sigma_i} \left(N_{ii}(0) - \frac{N_i}{1 + \sigma_i / \rho} \right) \right] e^{-\rho t} \end{aligned} \quad (11)$$

More important quantity to be defined is the **time of relaxation** to the equilibrium, which is defined as τ and dominated by:

$$[\rho(1 + \sigma_i/\rho)]^{-1} \sim \rho^{-1} = \tau \quad \text{since } \rho \gg \sigma_i \quad (12)$$

where the probability of leaving i regardless the destination is $\sigma_i = \sum_j \sigma_{ij}$. Therefore, the **equilibrium solutions** one may find are:

$$N_{ii}(t) = \frac{N_i}{1 + \sigma_i/\rho}, \quad N_{ij}(t) = \frac{\sigma_{ij} N_i / \rho}{1 + \sigma_i/\rho} \quad (13)$$

Since also in the **steady state** number of people that are *resident* in i is conserved, the we can compute the number of people **present** in i as:

$$N_i^* = N_{ii} + \sum_j N_{ji} = \frac{N_i}{1 - \sigma_i/\rho} + \sigma_j \frac{N_j \sigma_{ji} / \rho}{1 - \sigma_j/\rho} \quad (14)$$

The ratio actually σ_i/ρ quantifies the proportion of time spent outside and in the residence population.

Let us discuss now some **limiting cases**:

- $\sigma_i \rightarrow 0 \implies N_{ii}(t) \rightarrow N_i; N_{ij}(t) \rightarrow 0; N_i^* \rightarrow N_i$ in this case people rarely leave their residence, thus the non travelling individuals approaches the population of residents;
- $\rho \rightarrow \infty \implies N_{ii}(t) \rightarrow N_i; N_{ij}(t) \rightarrow 0; N_i^* \rightarrow N_i$ people return home immediately thus the non travelling individuals approaches the population of residents;
- $\rho \rightarrow 0 \implies N_{ii}(t) \rightarrow 0; N_{ij}(t) \rightarrow \frac{\sigma_{ij}}{\sigma_i} N_i; N_i^* \rightarrow \sum_j \frac{\sigma_{ji}}{\sigma_j} N_j$ here the case collapses to **migration** process: people never get back and the population of resident in i is distributed among the neighbouring destinations j .

One should recall now that the time scale of commuting $\tau \approx 8$ h, while the duration of an acute infection (i.e. flu) is way more: $\mu^{-1} \approx [1 - 3]$ days. We could adopt, as said, a *mean field* paradigm: we assume that the person can be partially in a place and partially in another one: on average it happens that he might be at the same time in different places. **Transmission dynamics** is therefore **slower than mobility**: we can assume that compartments occupation numbers are at the equilibrium with respect to mobility dynamics. Formally, the **occupation numbers** are:

$$X_{ii}^{[m]} = \frac{X_i^{[m]}}{1 + \sigma_i/\rho}, \quad X_{ij}^{[m]} = \frac{\sigma_{ij} X_i^{[m]}/\rho}{1 + \sigma_i/\rho}, \quad X^{[m]} = S, I, R \quad (15)$$

We want now to understand how many infected individuals a susceptible person resident in i is exposed to, that is to say we want to compute the **force of infection** λ :

$$\partial_t I = \lambda S(t) - \mu I(t), \quad \lambda = \beta \frac{I(t)}{N(t)} \quad (16)$$

Instead of explicitly modelling mobility, we now directly compute the effects of the other patches on the risk of infection, i.e. we break down the force of infection in its contributions. S_i is distributed among patch i and all possible destination j according to the proportions:

$$\left\{ \frac{1}{1 + \sigma_i/\rho}, \dots, \frac{\sigma_{ij}/\rho}{1 + \sigma_i/\rho}, \dots \right\}$$

and the risk of infection has the form:

$$\lambda_i = \frac{\lambda_{ii}}{1 + \sigma_i/\rho} + \sum_j \frac{\lambda_{ij} \sigma_{ij}/\rho}{1 + \sigma_i/\rho} \quad (17)$$

where the first contribution is the contribution of people that stay at home, and the second is the contribution of people infected that visit patch i from any other patch j . In the formula above we have that λ_{ii} is the force of infection that a person resident i experiences from infected people that are in i , formally:

$$\lambda_{ii} = \frac{\beta_i}{N_i^*} \left[I_{ii} + \sum_j I_{ji} \right] = \frac{\beta_i}{N_i^*} \left[\frac{I_i}{1 - \sigma_i/\rho} + \sum_j \frac{I_j \sigma_{ji}/\rho}{1 - \sigma_j/\rho} \right] \quad (18)$$

Whereas λ_{ij} susceptible resident in j that are travelling and so are exposed in other patch i to infected people:

$$\lambda_{ij} = \frac{\beta_j}{N_j^*} \left[I_{jj} + \sum_l I_{lj} \right] = \frac{\beta_j}{N_j^*} \left[\frac{I_j}{1 - \sigma_j/\rho} + \sum_l \frac{I_l \sigma_{li}/\rho}{1 - \sigma_l/\rho} \right] \quad (19)$$

Note as some of the terms are similar to the steady state ones in equation 14. These last expressions are actually useful in **numerical simulations**, and they allow us to understand the relative role of mobility and infection parameters on the epidemic dynamics in order to speed up the simulations, too. One should actually remember that both *infections* and *travels* are **stochastic processes**, and last results hold only when mobility is faster than the epidemic time scale.