# The impact of spatial and social structure on an SIR epidemic on a weighted multilayer network 

Ágnes Backhausz

Eötvös Loránd University and Alfréd Rényi Institute of Mathematics agnes.backhausz@ttk.elte.hu Joint work with István Z. Kiss, Péter L. Simon, Edit Bognár, Villő Csiszár, Damján Tárkányi, András Zempléni

21 November 2022, POPNET Connects seminar, University of Amsterdam (online)

## Acknowledgement + outline

This research was supported by the National Research, Development and Innovation Office within the framework of the RRF-2.3.1-21 Health Security National Laboratory and the Thematic Excellence Program 2021 - National Research Sub programme: "Artificial intelligence, large networks, data security: mathematical foundation and applications".

- Introduction: random graphs and SIR processes
- parameter sensitivity on a four-layer random graph model with cliques of fixed size
- parameter estimation on a two-layer random graph with preferential attachment dynamics


## SIR process on a random graph

Agent-based epidemic model; the connections between the individuals are chosen randomly

SIR process: susceptible-infected-recovered


## SIR process on a random graph

Agent-based epidemic model; the connections between the individuals are chosen randomly

SIR process: susceptible-infected-recovered


## SIR process on a random graph

Agent-based epidemic model; the connections between the individuals are chosen randomly

SIR process: susceptible-infected-recovered


## SIR process on a random graph

Agent-based epidemic model; the connections between the individuals are chosen randomly

SIR process: susceptible-infected-recovered


## SIR process on a random graph

Agent-based epidemic model; the connections between the individuals are chosen randomly

SIR process: susceptible-infected-recovered


## Real networks



A network of work relationships Source: http://develop-project.eu/

## Epidemic spread on random networks



An SIR model on 10000 vertices, with 100 simulations, on different random graph models source: Keeling, Eames, Networks and epidemic models, 2005

## Random graph models

- Erdős-Rényi model (1959): each pair is connected independently with the same probability
- preferential attachment models (e.g. Barabási-Albert model, 1999): in a randomly growing network, "popular" vertices have higher chance to get new neighbors
- configuration model (Bollobás, 1981): we fix the degree distribution (e.g. the proportion of vertices with 5 neighbors is $3 \%$ ), then we choose a graph uniformly at random


## Random graph models

- Erdős-Rényi model (1959): each pair is connected independently with the same probability
- preferential attachment models (e.g. Barabási-Albert model, 1999): in a randomly growing network, "popular" vertices have higher chance to get new neighbors
- configuration model (Bollobás, 1981): we fix the degree distribution (e.g. the proportion of vertices with 5 neighbors is $3 \%$ ), then we choose a graph uniformly at random
- geometric random graphs: vertices are embedded in the plane, and vertices that are closer to each other have higher chance to be connected
- random hypergraphs: groups of vertices are chosen randomly, to represent groups of people with strong interactions


## A random multilayer model

Goal: to find effective social distancing strategies in a model which does not contain too many parameters, but captures the most important features of the group structure of the society.

Related work:

- F. Ball, D. Mollison, G. Scalia-Tomba, Epidemics with two levels of mixing, 1997.
- G. Bianconi, Multilayer networks: structure and function, 2018.
- A. Aleta et al., Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19, 2020.


## A random multilayer model

Imagine a city of 5000 individuals in three age groups (children, adult, retired).

## A random multilayer model

Imagine a city of 5000 individuals in three age groups (children, adult, retired).

- households: sizes chosen according to the 2011 census in Hungary
- workplaces: groups of 10 adult people, chosen randomly
- schools: groups of 200 children, partitioned into 20 "classes" of size 10 with stronger relationships


## A random multilayer model

Imagine a city of 5000 individuals in three age groups (children, adult, retired).

- households: sizes chosen according to the 2011 census in Hungary
- workplaces: groups of 10 adult people, chosen randomly
- schools: groups of 200 children, partitioned into 20 "classes" of size 10 with stronger relationships
- spatial structure of the households: a grid with an additional dense part
- casual contacts (e.g. shops, playgrounds), depending on the spatial structure (capacity: 200, size of staff: 10)


## A random multilayer model



A caricature of the multilayer network where the building blocks are households (see green sets in the left panel) which are placed on or act as nodes in a square lattice which mimics spatial proximity

## Parameters

- households in a $40 \times 40$ grid (approximately 4000 individuals)
- infection rate at home: 1
- within school classes and at workplaces: 1
- infection rate between school classes, within the same school: 0.2
- infection rate between neighbors: 0.1
- infection rate of casual contacts: 0.1

We ran an SIR model, realized by a Gillespie algorithm.

## Effect of the population size



The evolution of the proportion of infected and recovered individuals by changing the size of the population ( $n_{r}$ is the size of the grid); the size of the population is approximately $4000,6000,8500$ and 12000 in the four cases respectively.

## Comparison to compartment models



The behavior of the maximal proportion of infected individuals and the final proportion of recovered individuals for the random multilayer graph model with random parameters (dots) and for the corresponding compartment model (red curve); the graph model has higher peaks

## Comparison to compartment models

$$
\begin{align*}
\dot{S}(t) & =-\beta I(t) \frac{S(t)}{N}  \tag{1}\\
\dot{I}(t) & =\beta I(t) \frac{S(t)}{N}-\gamma I(t)  \tag{2}\\
\dot{R}(t) & =\gamma I(t) \tag{3}
\end{align*}
$$

where $\beta$ and $\gamma$ denote the infection and recovery rates and $N$ is the size of the population.

The ratio $x=R_{\infty} / N$ satisfies the equation

$$
\begin{equation*}
1-x=\mathrm{e}^{-x R_{0}} \tag{4}
\end{equation*}
$$

where $R_{0}=\beta / \gamma$ is the basic reproductive ratio.
In addition,

$$
I_{\max }=N-S-R=N-\frac{N}{R_{0}}-\frac{N}{R_{0}} \ln \left(R_{0}\right) .
$$

This leads to a connection between $I_{\max } / N$ and $R(\infty) / N$.

## The role of the schools in this model



By decreasing the infection rates at schools (both within and between school classes, with the same factor) we can significantly reduce the number of infected individuals.

## The role of the schools in this model



By decreasing the infection rates at schools (both within and between school classes, with the same factor) we can significantly reduce the number of infected individuals. Recall that children has the largest contact number in this model.

## The effect of the size of classes



Decreasing the size of the classes does not help much.

## The effect of the connections between classes



Separating classes completely significantly helps: a school with completely separated cliques in not much worse than a closed school (in this model).

## The effect of size of workplaces



Most workplaces are already separated from each other, but some of them are the staff of shops etc. $\Rightarrow$ workplace size has a larger effect than the size of the classes.

## The effect of infection intensity for neighbors



The epidemic has a higher peak on the random graph model than in the compartment model.

Parameter estimation in the case of a two-layer random graph model
A two-layer model with a different structure:
First layer: households, complete graphs of a fixed size

## Parameter estimation in the case of a two-layer random

 graph modelA two-layer model with a different structure:
First layer: households, complete graphs of a fixed size
Second layer: preferential attachment random graph, independent of the first layer (polynomial model of Ostroumova, Ryabchenko and Samosvat (2013), when the new vertex $n+2$ is added:

## Parameter estimation in the case of a two-layer random

## graph model

A two-layer model with a different structure:
First layer: households, complete graphs of a fixed size
Second layer: preferential attachment random graph, independent of the first layer (polynomial model of Ostroumova, Ryabchenko and Samosvat (2013), when the new vertex $n+2$ is added:

- With probability $p_{\mathrm{pa}}$, we choose two vertices independently from the vertices $\{1, \ldots, n\}$, with probability proportional to their in-degree (preferential attachment component), and connect a new edge to these vertices.


## Parameter estimation in the case of a two-layer random graph model

A two-layer model with a different structure:
First layer: households, complete graphs of a fixed size
Second layer: preferential attachment random graph, independent of the first layer (polynomial model of Ostroumova, Ryabchenko and Samosvat (2013), when the new vertex $n+2$ is added:

- With probability $p_{\text {pa }}$, we choose two vertices independently from the vertices $\{1, \ldots, n\}$, with probability proportional to their in-degree (preferential attachment component), and connect a new edge to these vertices.
- With probability $p_{\mathrm{u}}$, we choose two vertices independently from the vertices $\{1, \ldots, n+1\}$, with equal probability (uniform component), and connect a new edge to these vertices.


## Parameter estimation in the case of a two-layer random graph model

A two-layer model with a different structure:
First layer: households, complete graphs of a fixed size
Second layer: preferential attachment random graph, independent of the first layer (polynomial model of Ostroumova, Ryabchenko and Samosvat (2013), when the new vertex $n+2$ is added:

- With probability $p_{\text {pa }}$, we choose two vertices independently from the vertices $\{1, \ldots, n\}$, with probability proportional to their in-degree (preferential attachment component), and connect a new edge to these vertices.
- With probability $p_{\mathrm{u}}$, we choose two vertices independently from the vertices $\{1, \ldots, n+1\}$, with equal probability (uniform component), and connect a new edge to these vertices.
- With probability $p_{\mathrm{tr}}$, we choose a random edge of $G_{n}$ uniformly, and connect a new edge to both of its endpoints.


## Parameter estimation based on the number of SI edges

Suppose that the recovery rate is known, $\gamma=1$, and we would like to estimate the infection rate $\tau$. A general maximum likelihood estimate is as follows, if the number of SI edges is exactly known:

$$
\hat{\tau}=\frac{z_{l}}{\int_{0}^{T} E_{t}^{S I} d t}=\frac{z_{l}}{\sum_{t_{i}<T} E_{t_{i}}^{S I}\left(t_{i}-t_{i-1}\right)},
$$

where

- $z_{l}$ is the total number of events when a vertex gets infected;
- $E_{t}^{S I}$ is the number of SI edges (edges with one susceptible and one infected endpoint).

Parameter estimation based on the number of SI edges


The estimate converges quickly to the real parameter $\tau$, for all parameter sets. Each curve is the average of 25 simulations.

## Estimating the number of SI edges between households

Suppose that the recovery rate is known, $\gamma=1$, and we would like to estimate the infection rate $\tau$. We used the following estimate of the SI edges between households:

$$
\hat{E}_{t}^{S I, o}=I_{t} \cdot\left(d-\frac{w d}{w d+N_{\mathrm{hh}}-1}\right) \cdot \frac{S_{t}}{N}
$$

where

- $d$ is the average degree of a vertex outside its household;
- $N_{\text {hh }}$ is the size of households;
- $N$ is the size of the population;
- $I_{t}$ and $S_{t}$ are the number of infected and susceptible individuals at time $t$;
- $w$ is the weight of the edges between households.


## Parameter estimation



The estimate depends on the weight of the preferential attachment component in the random graph model. Each curve is the average of 25 simulations,

## Parameter estimation



The estimate does not really depend on the weight of the triangle component in the random graph model. Each curve is the average of 25 simulations.

## Conclusions

In our model,

- the epidemic curve in a multilayer random graph model can have a higher peak than the corresponding compartment model shows
- if everyone meets the same people every day, separating these cliques (bubbles) can be almost as efficient as a full lockdown
- the size of these cliques can have a significant effect, but not in every situation
- the structure of the graph, especially the preferential attachment component has a significant effect on the epidemic spread and on the estimates on the infection rate


## Conclusions

In our model,

- the epidemic curve in a multilayer random graph model can have a higher peak than the corresponding compartment model shows
- if everyone meets the same people every day, separating these cliques (bubbles) can be almost as efficient as a full lockdown
- the size of these cliques can have a significant effect, but not in every situation
- the structure of the graph, especially the preferential attachment component has a significant effect on the epidemic spread and on the estimates on the infection rate

Future plans

- vaccination strategy: is it worth vaccinating members of households, workplaces etc. together?
- how can we give better estimates for the parameters, e.g. using neural networks?

