Activity-based modeling and simulation of epidemics

C.Cortes Balcells, R. Krueger, M. Bierlaire

PhD Candidate,
Ecole Polytechnique Federale de Lausanne (EPFL)

STRC Conference

Motivation

Challenges

- Lockdown across the world due to SARS-CoV-2 manifest the need of **robust** and **dynamic** models, to guide decision making.
- Accounting for individual behaviour through an epidemic outbreak by using **large scale models**.
- Datasets are growing in size and are becoming available in continuous streams.
- Difficulty of finding **disaggregated data** to **validate** the model.
- Capturing **spread of the disease** through **public transportation**.
- Allows to **assess the impact** that a certain policy has on **different segments of the population**.

Research gaps

Limitations

- Lack of data leads to add aggregated parameters inside the agent-based models, [TYK+20].
- Agent-based models in order to define more targeted and less disruptive interventions. Results are achieved using real-time data-driven analysis, [AMCB+20].
- Clear methodology to know which variables are meaningful inside an epidemiological model, for example income or residence place, [CPK+21].
- Make the probabilities time dependant, since an early adoption can potentially allow to contain the epidemics, [MBCV20].

Agent-based epidemiological models

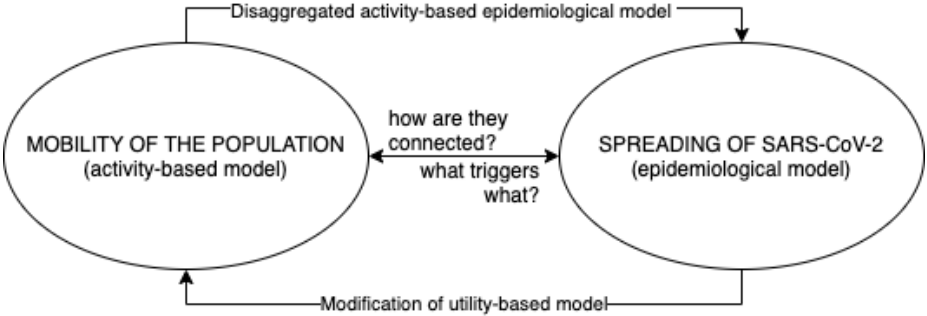

Figure: Schematic connection cycle between mobility and epidemics.

Outline of this talk

1. Added value of using disaggregate models for modelling SARS-CoV-2 spreading. [1]
2. Description of the preliminary considerations and presentation of a model that accounts for virological and socio-economic variables.[2]
3. Potential of these models to study SARS-CoV-2 policy decision making.[3]

Literature:

[1] A. Aleta, D. Martin-Corral, M. Bakker, A. Piontti, M. Ajelli, M. Litvi-nova, M. Chinazzi, N. Dean, M. Halloran, I. Longini, A. Pentland,A. Vespignani, Y. Moreno, and E. Moro. Quantifying the importance and location of sars-cov-2 transmission events in large metropolitan areas, 12 2020.

[2] S. Chang, E. Pierson, P. Koh, J. Gerardin, B. Redbird, D. Grusky, andJ. Leskovec. Mobility network models of covid-19 explain inequities and inform reopening. Nature, 589:1-6, 01 2021.

[3] M. Mancastroppa, R. Burioni, V. Colizza, and A. Vezzani. Active and inactive quarantine in epidemic spreading on adaptive activity-driven networks. Physical Review E, 102, 08 2020.

Activity-based models

Activity-based model [AG92]

- They allow more complex policies to be evaluated.
- The phenomena are understood as the result of the interaction of multiple agents, each guided by individual norms or intelligence.
- This interaction results in a complex system, consisting of many sub-systems and agents (applicable to many disciplines: ecology, economics, computer simulation...).
- It uses microscopic simulation.
- Examples: MATsim, TRANSIMS...

Utility-based model example formulation

See [CN05], based on [AdPL93] and [CSH75].

Utility of a plan S_{plan}:

$$S_{\text{plan}} = \sum_{q=0}^{N-1} S_{\text{act},q} + \sum_{q=0}^{N-1} S_{\text{trav,mode}\,(q)}$$

The utility of an activity q:

$$S_{\text{act},q} = S_{\text{dur},q} + S_{\text{wait},q} + S_{\text{late.ar},\,q} + S_{\text{early.dp},q} + S_{\text{short.dur},q}$$

Epidemiological models

Compartmental models

The SIR epidemic model can be written in the following way(c.f [KMS17]):

- The transitions at each time step Δt are:

$$\frac{\partial S}{\partial t}(t) = -\beta I(t)\frac{S(t)}{N}$$
$$\frac{\partial I}{\partial t}(t) = \beta I(t)\frac{S(t)}{N} - \gamma I(t)$$
$$\frac{\partial R}{\partial t}(t) = \gamma I(t)$$

- S: Susceptible
- I: Infected
- R: Recovered

$$N - R_\infty = S(0)\exp\left(-\frac{\beta}{N\gamma}R_\infty\right)$$

Compartmental models

Issues:

- SEIR models work on an aggregate level: neglects the imperfect mixture.
- Not transferable to different epidemics.
- Crutial parameters might not be available.
- Exponential is a strong assumption.

Network models

The graph G is defined by n vertices, and m edges:

$\{G_1, G_2, \ldots, G_n\}$, where $n = \begin{pmatrix} M \\ m \end{pmatrix}$ with $M = N(N-1)/2$.

The probability of picking each graph is the same:

$1/n$.

Network models

Issues:

- Complex to find the correct adjacency matrix.
- Difficult to use them in densely populated areas.
- Quality of the contacts between two individuals. The adjacency matrices are binary.
- Static character of network models.

Individual-centric models

If an individual is susceptible and it has contact with an infected agent, it becomes infected with a probability: p.

This probability can be defined as desired. For example in [Smi09], the probability for person n to become infected by this process in a time step t is defined as:

$$P_{n,t} = 1 - \exp\left[-\Theta \sum_{m} q_{m,t} \cdot ci_{nm,t} \cdot in_{n,t} \cdot \tau_{nm,t}\right]$$

Main issues of this probability:

- Parameters unknown for COVID-19, so set to $value = 1$,
- All multiplying so it might be independent modifying one or another,
- We want to create dependence on not only individual and time but also on the location.

Individual-centric models

- $\rightarrow$ insight on transmission and intervention that will complete what can be obtained with usual compartmental models (SIR).
- Added value of these models ...
  - The interactions between agents are nonlinear, discontinuous or complex..
  - When the space is crucial and we do not have fixed positions.
  - Population is heterogenous with different socioeconomic characteristics.
  - Agents have complex behaviour.
  - Topology of interactions is complex.

Methodology

Notation	Description Variables
S	Susceptible population.
I	Infected population.
R	The population who recovered from the disease and got immunity.
Δt	The time-step of the simulation.
X_m	Explanatory variables from the dataset.
H	Total number of individuals in the population.
I_i^{met}	Number of individuals crossed by individual i.
H_i^{met}	Number of total crossings between two individuals.
ϵ_i	Error term explanatory variables of β_i.
μ_i	Error term explanatory variables of γ_i.
α_m	Parameters of the explanatory variables.
θ_m	Parameters of the explanatory variables.
β_i	Contagion rate between S and I.
$1/\gamma_i$	Length of the infectious period for population I.

Table: Table of notation

For each individual i we define:

Every susceptible individual i, at time t, has a probability of becoming infected:

$$\dot{P}_{S \to I}(t) = 1 - \exp(-\beta_i \frac{I}{H} dt)$$

Every infected individual i, at time t, has a probability of becoming recovered:

$$\dot{P}_{I \to R}(t) = 1 - \exp(-\frac{1}{\gamma_i} dt)$$

Every non-recovered individual i, at time t, has a probability of dying:

$$\dot{P}_{NR \to D}(t) = 1 - \exp(-\lambda dt)$$

The β_i' is defined as the sum of the different factors that make an agent recover:

$$\beta_i' = \sum_{j=1}^{m} \alpha_m X_m + \epsilon_i$$

The β_i depends on the number of people that the agent has contact with:

$$\beta_i = \beta_i' \frac{I_{met}}{H_{met}}$$

The $1/\gamma_i$ is defined as the sum of the contagion risks coming from the different sources of infections:

$$\gamma_i = \sum_{j=1}^{m} \theta_m X_m + \mu_i$$

Vaccination

We introduce vaccination in our model by adding the effectiveness of the different vaccines against SARS-CoV-2, [Roa21]:

Vaccine	Effectiveness in %
NVX-CoV2373	96.0
Comirnaty	95.0
mRNA-1273	94.1
Sputnik V/Gam-Covid-Vac	91.6
BBIBP-CorV	79.0
AZD1222/Covishield	76.0
Ad26.COV2.S	72.0

Table: Effectiveness of vaccines against SARS-CoV-2.

- We generate random assignments of the different types of vaccines in each canton and the total number of doses.
- In Switzerland, three of the vaccines have been distributed: BioNTech, Pfizer, Moderna and Johnson & Johnson.

MATSim-Episim

- Episim is an open source framework which can be used to simulate, based on MATSim events, an epidemic outbreak.
- The model has states exposed, infectious, showing symptoms, seriously sick (should be in hospital), critical (needs intensive care), and recovered.
- The durations from one state to the next follow log-normal distributions.

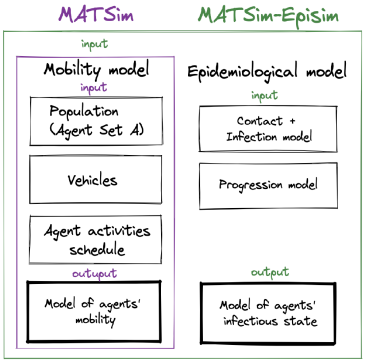

Figure: High-level description of the models' hierarchy.

Dataset

Descriptions of the available variables are:

Variable	Description
Individual	Id of the individual.
Age	Age of the individual.
Gender	Gender of the individual.
Home	Coordinates of the individual home.
Infected	If the SARS-CoV-2 test was positive or not.
Vaccinated	If the individual was vaccinated.
Vaccinationdate	When was the individual vaccinated.

Table: Description of respondent specific variables

Conclusions and future work

- **Lack** of a **consensus** on how to best represent the **infectiousness** of a disease in a **given population**.
- Most existing research focuses on an **aggregated approach** to estimate the various parameters that define the spread of an infectious disease. It is important to account for **heterogeneity**.
- Epidemiological models are a **crucial element** for **public transportation planning** and **activity-travel behavior**. Lack of research focused on **evaluating public transportation policies** for a **targeted group**.

Thank you

References I

[AdPL93] Richard Arnott, André de Palma, and Robin Lindsey. A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *American Economic Review*, 83:161–79, 02 1993.

[AG92] Kay Axhausen and Tommy Gärling. Activity-based approaches to travel analysis: Conceptual frameworks, models and research problems. *Transport Reviews - TRANSP REV*, 12:323–341, 10 1992.

[AMCB+20] Alberto Aleta, David Martín-Corral, Michiel Bakker, Ana Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie Dean, M. Halloran, Ira Longini, Alex Pentland, Alessandro Vespignani, Yamir Moreno, and Esteban Moro. Quantifying the importance and location of sars-cov-2 transmission events in large metropolitan areas, 12 2020.

[CN05] David Charypar and Kai Nagel. Generating complete all-day activity plans with genetic algorithms. *Transportation*, 32:369–397, 07 2005.

[CPK+21] Serina Chang, Emma Pierson, Pang Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec. Mobility network models of covid-19 explain inequities and inform reopening. *Nature*, 589:1–6, 01 2021.

References II

[CSH75] Jeffrey Chapman, Thomas Slobko, and Joseph E. Haring. Congestion theory and transport demand. *Regional Science and Urban Economics*, 5(4):493–501, 1975.

[KMS17] Istvan Kiss, Joel Miller, and Péter Simon. *Mathematics of Epidemics on Networks*, volume 46. 01 2017.

[MBC+20] Sebastian Müller, Michael Balmer, Billy Charlton, Ricardo Ewert, Andreas Neumann, Christian Rakow, Tilmann Schlenther, and Kai Nagel. Using mobile phone data for epidemiological simulations of lockdowns: government interventions, behavioral changes, and resulting changes of reinfections, 07 2020.

[MBCV20] Marco Mancastroppa, Raffaella Burioni, Vittoria Colizza, and Alessandro Vezzani. Active and inactive quarantine in epidemic spreading on adaptive activity-driven networks. *Physical Review E*, 102, 08 2020.

[Roa21] Mónica Mena Roa. Infografía: ¿qué tan eficaces son las vacunas contra la covid-19, May 2021.

[Smi09] Timo Smieszek. A mechanistic model of infection: Why duration and intensity of contacts should be included in models of disease spread. *Theoretical biology and medical modelling*, 6:25, 11 2009.

[TYK+20] J. T. Tuomisto, Juha Yrjölä, M. Kolehmainen, Juhani Bonsdorff, J. Pekkanen, and Tero Tikkanen. An agent-based epidemic model reina for covid-19 to identify destructive policies. *medRxiv*, 2020.

Model Formulation

Waiting:

$$S_{\text{wait},q} = \beta_{\text{wait}} \cdot t_{\text{wait},q}$$

Performing an activity:

$$S_{dur,q} = \beta_{\text{perf}} \cdot t_{\text{typ},q} \cdot \ln\left(t_{\text{dur},q}/t_{0,q}\right)$$

Late arrival :

$$S_{\text{late.ar},q} = \begin{cases} \beta_{\text{late.ar}} \cdot \left(t_{\text{start},q} - t_{\text{latest.ar},q}\right) & \text{if } t_{\text{start},q} > t_{\text{latest.ar},q} \\ 0 & \text{else} \end{cases}$$

Early departure :

$$S_{\text{early.}\,dp} = \begin{cases} \beta_{\text{early.dp}} \cdot \left(t_{\text{earliest.}\,dp,q} - t_{\text{end},q}\right) & \text{if } t_{\text{end},q} < t_{\text{earliest.dp},q} \\ 0 & \text{else} \end{cases}$$

SIR model used in Episim, [MBC$^+$20]

- The probability for person n to become infected by this process in a time step t in [MBC$^+$20]:

$$P_{n,t} = 1 - \exp\left[-\Theta \sum_m q_{m,t} \cdot ci_{nm,t} \cdot in_{n,t} \cdot \tau_{nm,t}\right]$$

- **Main issues of this probability:**
  - Parameters unknown for COVID-19, so set to $value = 1$,
  - All multiplying so it might be independent modifying one or another,
- We want to create dependence on not only individual and time but also on the location.