

# POLITECNICO DI TORINO



## Computer aided simulations and performance evaluation

Academic year 2020/21

# Contents

<b>1</b>	<b>Epidemic Models</b>	<b>2</b>
1.1	Numerical SIR Models . . . . .	2
1.2	Simulative SIR Models . . . . .	2
1.2.1	Define all the input parameters of the simulator . . . . .	2
1.2.2	Define all the output metrics of the simulator . . . . .	2
1.2.3	Define the main data structures for the simulator . . . . .	2
1.2.4	S(t), I(t), R(t) plot . . . . .	2
1.2.5	Rt(t) plot . . . . .	3
1.3	Simulative SIR Models (extension) . . . . .	3
1.3.1	Define all the input parameters of the simulator . . . . .	3
1.3.2	Define the main data structures for the simulator . . . . .	3
1.3.3	S(t), I(t), R(t) plot . . . . .	3
1.3.4	Rt(t) plot . . . . .	3
1.3.5	Summary results . . . . .	4

# Epidemic Models

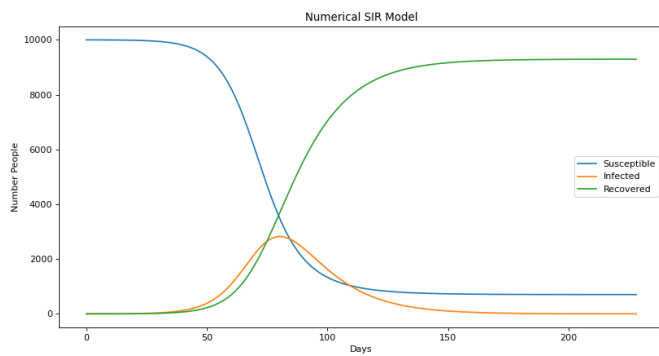
In this chapter there will be analysis of some epidemic models, in particular SIR models. In a SIR model the population is divided in 3 categories:

- **Susceptible:** 'healthy' people;
- **Infected:** people who contracted the virus;
- **Recovered:** people who contracted the virus, spent enough days infected and recovered.

The values used are:

- Population size ( $N$ ) equal to 10000 individuals;
- A contact rate, ( $\beta$ ), equal to  $0.2 \text{ day}^{-1}$ . This refers to the average number of people that an infected person meets each day;
- An recovery period equal to 14 days. So a recovery rate ( $\gamma$ ) equal to  $\frac{1}{14 \text{ days}} = 0.071 \text{ day}^{-1}$ .

## 1.1 Numerical SIR Models



The values of the S,I and R variable are calculated as follow:

$$\begin{aligned} S(t_{k+1}) &= S(t_k) - \frac{\beta}{N} \Delta t S(t_k) I(t_k) \\ I(t_{k+1}) &= I(t_k) + \frac{\beta}{N} \Delta t S(t_k) I(t_k) - \gamma \Delta t I(t_k) \\ R(t_{k+1}) &= R(t_k) + \gamma \Delta t I(t_k) \end{aligned}$$

Actually these are the solutions of differential equations. When solving differential equations, it is needed to define the initial condition, in this case:

$$\begin{aligned} S(t_0) &= N - 1 \\ I(t_0) &= 1 \\ R(t_0) &= 0 \end{aligned}$$

This is the usual condition of the beginning of an epidemic. The maximum number of infected people is 2823 and it is reached at the day 80. The pandemic ends at day 229 when the number of infected people is less than 1, at this day the number of susceptible people are 705 and the number of recovered one is 9294.

## 1.2 Simulative SIR Models

### 1.2.1 Define all the input parameters of the simulator

The input parameters of the simulation are:

- The population size;
- The total days the simulation will run for;
- The contact rate,  $0.2 [\text{days}^{-1}]$  in this case;
- The infection period,  $14 [\text{days}]$  in this case;
- The seed value used to initialize a pseudorandom number generator;
- The confidence level used to calculate the confidence interval;
- The number of runs: the number of times that we run our simulation. This is done in order to have more accurate.

### 1.2.2 Define all the output metrics of the simulator

The output metrics of the simulation are:

- For both the maximum number of infected people, the day at which this occurs and the day at which the simulation ends:
  - The average;
  - The length of the confidence interval;
  - The relative error.

### 1.2.3 Define the main data structures for the simulator

The main data structures are:

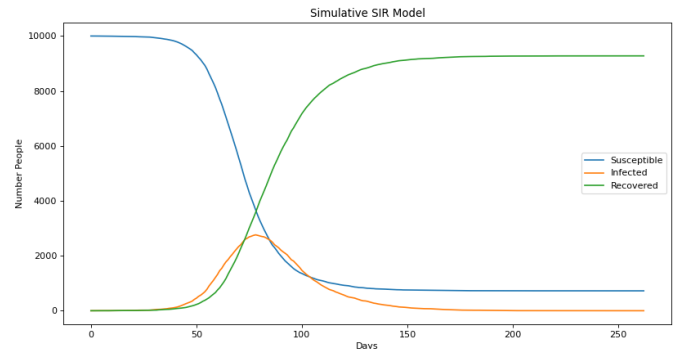
- A custom class Individual, where to store some variables (the category, the day when he is infected, ...) and some methods (UpdateState() that checks if an Individual recovered or not, ...)
- A python list with N objects of type Individual;
- A python set where to store the index of the Individual that are infected at day d. This avoid to loop over the all population since  $I < N$  (I=number of individual infected at a given day).

For every infected Individual at day  $d$  a number of contacts ( $n$ ) is chosen following a Poisson distribution (with parameter  $\lambda = \text{contact\_rate}$ ), then  $n$  Individuals are chosen to be infected.

For each new infected, the day at which they will recover is chosen using a Geometric distribution (with parameter  $p = \frac{1}{\text{infection\_period}}$ ). These distributions are chosen for 2 reasons:

- They are both discrete distributions;
- They seem to fit a real pandemic situation.

### 1.2.4 S(t), I(t), R(t) plot



The values of the S,I and R variable are calculated as follow:

$$\begin{aligned} S(t_{k+1}) &= S(t_k) - \text{new\_infections} \\ I(t_{k+1}) &= I(t_k) + \text{new\_infections} - \text{new\_recovered} \\ R(t_{k+1}) &= R(t_k) + \text{new\_recovered} \end{aligned}$$

Actually these SIR formulas are the same formulas as the numerical case but there here there is not an analytical formula for *new\_infections* and *new\_recovered* but they have to be obtained through simulation. In particular *new\_infections* is the number of new infections at day  $t_k$  and *new\_recovered* is the number of new recovered at day  $t_k$ .

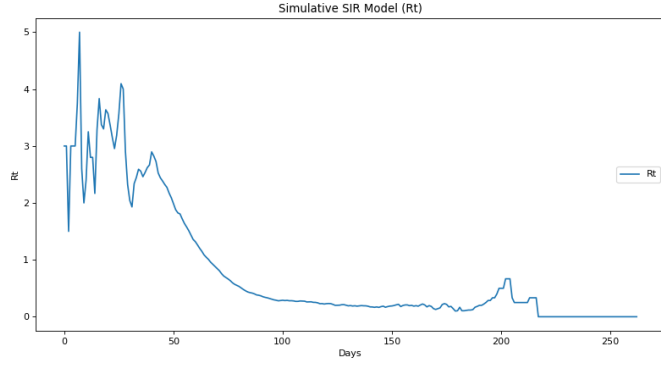
The results with 7 runs are: the maximum number of infected people is 2758.14 with a confidence interval length of 84.80 (relative error = 0.03) and it is reached at the day 76 with a confidence interval length of 9.44 (relative error = 0.12). The pandemic ends at day 265 when the number of infected people is equal to 0, with confidence interval length

44.29 (relative error = 0.16).

### 1.2.5 Rt(t) plot

The Rt value is defined as: *average number of secondary infections produced by a typical infective during the entire period of infectiousness at time t.*

It is an important value to understand the trend of the pandemic: whether the epidemic is expected to grow ( $R_t > 1$ ), or will start declining ( $R_t < 1$ ).



## 1.3 Simulative SIR Models (extension)

### 1.3.1 Define all the input parameters of the simulator

The additional input parameters of the simulation are:

- The transmission rate. This refers to the probability that an infected person infects other susceptible people, when some conditions are met: close contact, no protections and so on;
- The movement length: how much an individual can move each day;
- The contact distance: the distance at which Individuals can interact;
- The number regions in which divide the population.

### 1.3.2 Define the main data structures for the simulator

The additional main data structures are:

- A python dictionary with key, the number of the region, and as value a set with all the susceptible Individuals in that region. This avoids to loop again over the all population since the number of Individuals in a region is much smaller than N ( $n = \frac{N}{\text{number\_regions}^2}$ , on average, assuming uniform distribution)

For every infected Individual, is calculated the distance between him and all the other Individual; if this distance is less than a threshold (`contact_distance`) then they can interact and the Individual is infected with probability  $\text{transmission\_rate} = 0.2$

The major problem in checking the distance between Individuals is the complexity:  $\mathcal{O}(N^2)$ ; this is quite a large complexity and may require lot of time, especially for large number of N. For this reason, some optimizations are needed and implemented.

The main loop works as follow:

For each day, first the simulator iterates over all the infected

Individuals in that particular day ( $\mathcal{O}(I)$ ), for every infected Individual it loops over each susceptible Individual in the region where the infected Individual belongs ( $\mathcal{O}(n)$ ), it checks the distance and, if they are close enough, if the infection happened. So, the major complexity of for each day is

$$\mathcal{O}(I \times n) = \mathcal{O}(I \times \frac{N}{\text{number\_regions}^2})$$

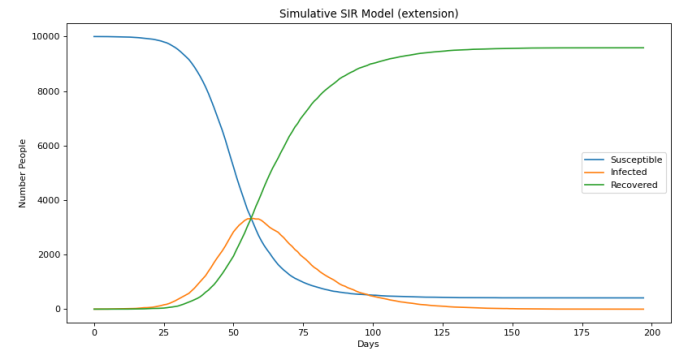
and this value is much smaller than  $\mathcal{O}(N^2)$ , for  $I$  smaller than  $N$  (as it usually is during the different day) and for `number_regions` greater than 1.

There are also some other complexity like the time in building the python dictionary for dividing the population in regions but it is quite fast.

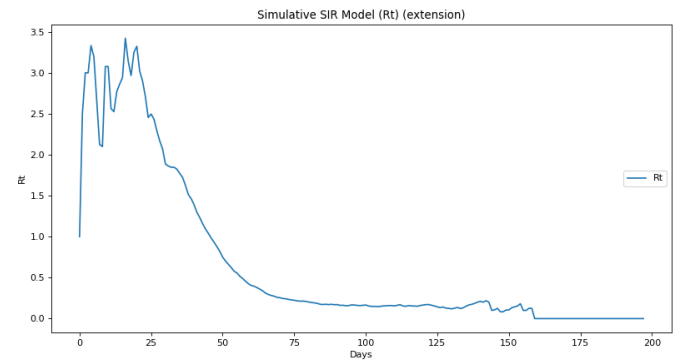
Dividing the population in regions introduce some approximation since 2 Individuals in 2 different regions (e.g. at the edge) but close enough ( $\text{distance} < \text{contact\_distance}$ ) can not be infected; but the speed up is significant!

The results with 7 runs are: the maximum number of infected people is 3204.16 with a confidence interval length of 87.77 (relative error = 0.03) and it is reached at the day 66.5 with a confidence interval length of 7.53 (relative error = 0.11). The pandemic ends at day 211.5 when the number of infected people is equal to 0, with confidence interval length 13.48 (relative error = 0.06).

### 1.3.3 S(t), I(t), R(t) plot



### 1.3.4 Rt(t) plot



The 2 models can not be compared since they have different parameters, but the trends are similar.

1.3.5 Summary results

SIR Model	Avg Max Infected	RE Max Infected	Avg Max Infected (Day)	RE Max Infected Day	Avg End Epidemic	RE End Epidemic
Numerical	2823	/	80	/	229	/
Simulative	2758.14	0.03	76	0.12	265	0.16
Simulative (extension)	3204.16	0.03	66.5	0.11	211.5	0.06