

Human Mobility and Disease Spread Simulation

Madotto Andrea

Hong Kong Baptist University

M.Sc. Practicum

20-05-2015

Overview

Introduction

SIR Models

Compartment model

Social Contact Networks

Datasets

Infection, vaccination and result

Hive Plot

Human mobility

Conclusion

Overview

Introduction

SIR Models

Compartment model

Social Contact Networks

Datasets

Infection, vaccination and result

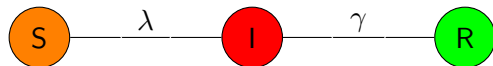
Hive Plot

Human mobility

Conclusion

Introduction

- ▶ Disease spread simulation based on SIR model
Susceptible **I**nfectives **R**ecovered
 - ▶ SIR compartment model
 - ▶ SIR on contact social networks
- ▶ Human mobility
 - ▶ Flight passengers
- ▶ Data visualizations
 - ▶ Hive plot
 - ▶ Data map



Overview

Introduction

SIR Models

Compartment model

Social Contact Networks

Datasets

Infection, vaccination and result

Hive Plot

Human mobility

Conclusion

SIR: Compartment model

- ▶ System of differential equations to simulate the disease propagations
- ▶ Three different variables (S,I and R) that represents the status of each compartment
- ▶ Where β represent the disease transmission rate
- ▶ Where γ describes the recovery rate

$$\frac{dS}{dt} = -\lambda \cdot S$$

$$\frac{dI}{dt} = \lambda \cdot S - \gamma \cdot I$$

$$\frac{dR}{dt} = \gamma \cdot I$$

$$\lambda = \beta \cdot \frac{I}{N}$$

Compartment model: R_0

Another important parameter, called basic reproduction number, is given by:

$$R_0 = \frac{\beta}{\lambda}$$

and it has well known value, useful to have a correct parametrization.

Disease	Transmission	R_0
Diphtheria	Saliva	6-7
Smallpox	Airborne droplet	5-7
Mumps	Airborne droplet	4-7
HIV/AIDS	Sexual contact	2-5
SARS	Airborne droplet	2-5
Influenza(1918 pandemic strain)	Airborne droplet	2-3
Ebola(2014 Ebola outbreak)	Bodily fluids	1.5-2.5

Table: Values of R_0 of well-known infectious diseases

Compartment model: implementation and result

- ▶ Python with scipy and numpy library
- ▶ Implement of the Algorithm 1
- ▶ Basic setting:
 $S = 100000$, $I = 3$,
 $R = 0$, $\beta = 0.5$,
 $\gamma = 0.112$ and $t = 100$

Data: Set of parameters:

$$\{S, I, R, N, \gamma, \beta\}$$

Result: Result statistic of the simulation

SetParameters($S, I, R, N, \lambda, \beta$);

foreach $t \in T$ **do**

$$S(t) = -\beta \cdot \frac{I}{N} \cdot S(t-1);$$

$$I(t) = \beta \cdot \frac{I}{N} \cdot S(t-1) - \gamma \cdot I(t-1);$$

$$R(t) = \gamma \cdot I(t-1);$$

end

Algorithm 1: Pseudocode of compartmental model simulation

Compartment model: implementation and result (cont)

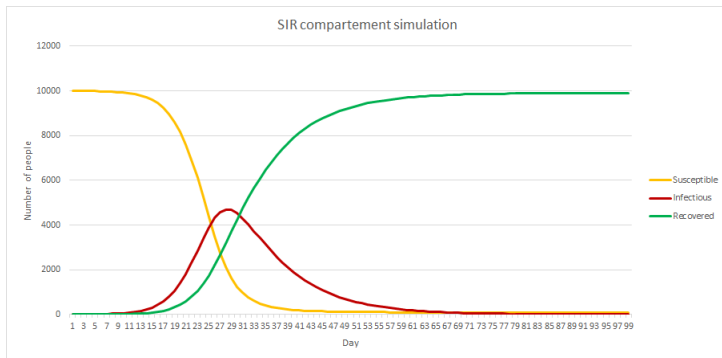


Figure: SIR compartment simulation with $R_0 = 4$

Overview

Introduction

SIR Models

Compartment model

Social Contact Networks

Datasets

Infection, vaccination and result

Hive Plot

Human mobility

Conclusion

SIR: Social Contact Networks

- ▶ Weighted graph $G(V, E)$ describe a contact social network
- ▶ $V = \{v_1, \dots, v_n\}$ a set of nodes that represents the different individuals
- ▶ Set of label L in the node for representing the state of the nodes (S, I or R)
- ▶ $E = \{e_1, \dots, e_n\} \subseteq (V \times V)$ a list of edges that represent the connection between different individuals
- ▶ Each edge $e = (v_i, v_j)$ is weighted by the duration of the contact between v_i and v_j
- ▶ Probability of infection (depend on the status of the node, the state of its neighbour and the duration of the contact)

$$w_i = \sum_{e \in adj_I(v_i)} l_e \quad l_e \in duration(e)$$
$$p(v_i) = 1 - (1 - r)^{w_i} \quad (1)$$

SIR: Social Contact Networks

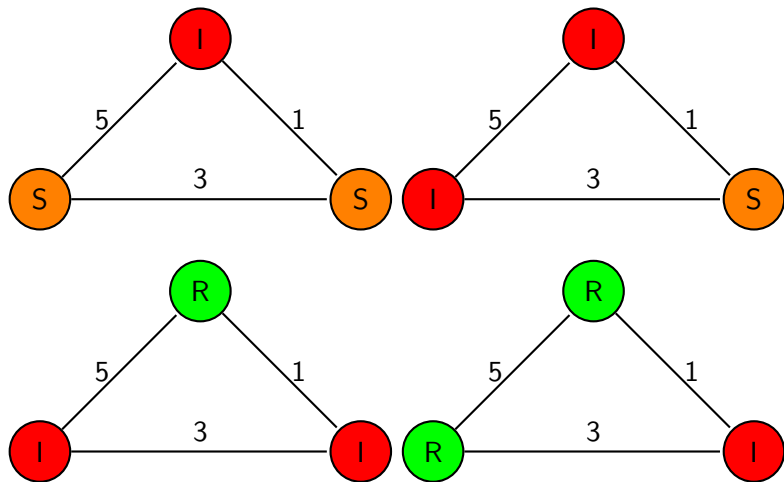


Figure: Example with 3 nodes, 3 edges and $\Delta_{t_i} = 2$

Social Contact Networks: details and implementation

- ▶ Graph database to implement the network (Neo4j)
 - ▶ Py2neo library to connect Python to the database
 - ▶ Implementation of the Algorithm 2 with Python
-
- ▶ A node can remain infect for a Δ_{t_i} period
 - ▶ `calculateProbInfect(v_i)` use the Equation 1
 - ▶ `graphUpdate(V)` update all the labels just after the finish of the For
-
- ▶ Note that a susceptible node can become infected just if it has at least one infected node
 - ▶ Δ_{t_i} represent the infection duration, after this time the node become recovered

Data: Social network graph

Result: Statistic on the simulation and effectiveness of the tested intervention

```
graphCreate();  
foreach  $t \in T$  do  
    applyIntervention( $G$ );  
    foreach  $v_i \in V$  do  
         $p_i = \text{calculateProbInfect}(v_i)$ ;  
  
         $r_i = \text{random}(0,1)$ ;  
        if  $r_i \geq p_i$  then  
             $v_i.\text{status} = \text{INFECTED}$ ;  
        end  
    end  
    graphUpdate( $V$ );  
end
```

Algorithm 2: Pseudocode of simulation

Social Contact Networks: Technologies

- ▶ Neo4j: ensures the atomicity, consistency, isolation and durability during the graph manipulations. In addition, it provides flexibility, scalability and high performance.



- ▶ Py2neo: is a client library and comprehensive toolkit for working with Neo4j from within Python applications and from the command line. The core library has no external dependencies and has been carefully designed to be easy and intuitive to use.



Overview

Introduction

SIR Models

Compartment model

Social Contact Networks

Datasets

Infection, vaccination and result

Hive Plot

Human mobility

Conclusion

Datasets: Scientific collaboration networks

- ▶ This dataset represents the collaborations between the authors and co-authors in the scientific papers
- ▶ Connection between two authors if they have their names in the same publication
- ▶ Duration of the contact is represented by the number of the papers that the two authors have in common
- ▶ The authors spend time, during the days, with the co-authors and the duration of the contact is proportional to the number of the papers they work together

Datasets: Student from University of California

- ▶ Represents the relationships between the students in the University of California
- ▶ Connection between the students if they exchange messages
- ▶ exchange a lot of message, then they have higher contact duration
- ▶ We consider the duration of the contact as proportional to the number of the message exchanged between the students

Datasets: Summary

Dataset name	Number of nodes	Number of edge	Density
Scientific collaboration networks	1589	2741	2.97
Student from University of California	1899	8695	20.97

Table: Summary of the datasets

Note that, this dataset comes from previous article and web repository. Hence, before import this data to Neo4j we have clean and sometimes summarized. To do that, we imported all the data in MySQL.

Overview

Introduction

SIR Models

Compartment model

Social Contact Networks

Datasets

Infection, vaccination and result

Hive Plot

Human mobility

Conclusion

Experiment: infections

First of all, we run the simulation in all the datasets using the Algorithm 2. This because we need to have a referent measurements in order to classify the countermeasure effectiveness. For reference we show the result of the Scientific collaboration network. The basic setting are: $r = 0.05$ and $t = 20$.

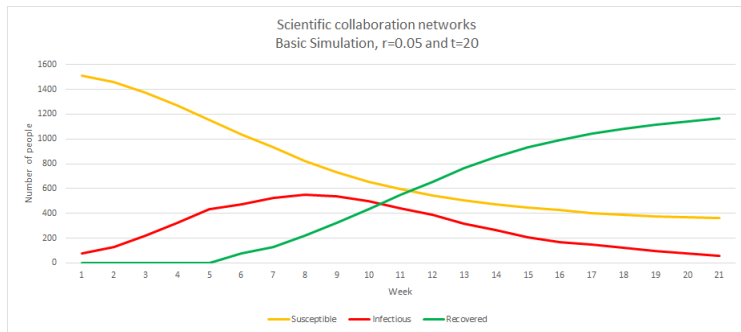


Figure: Scientific collaboration networks, basic simulation

Experiment: vaccinations

- ▶ To contain the spread of the virus
- ▶ The application of vaccine shot to the network will start after the ratio I/S is greater 0.1
- ▶ The effectiveness of the vaccination is not 100% but just 70-80%
- ▶ We would like to represent the vaccination decision making of an single node in the network.
 - ▶ Depending on the cost of the vaccination
 - ▶ Depending on the behaviours of the neighbour

Vaccinations decision making

To model the decision making we implement the following formula

$$\lambda_i = \frac{N_i^{non}}{(N_i^{non} + N_i^{vac})}$$

$$r = \frac{c_{vac}}{c_{inf}}$$

where i is an index of a node in the network, c_{vac} represent the cost of the vaccination and c_{inf} the costs of disease infection. The number r is called reproduction number and more smaller means cheaper. Then a node take the vaccine if

$$\begin{cases} r < \lambda_i & \text{take the vaccine} \\ r \geq \lambda_i & \text{don't take the vaccine} \end{cases} \quad (2)$$

Vaccinations result

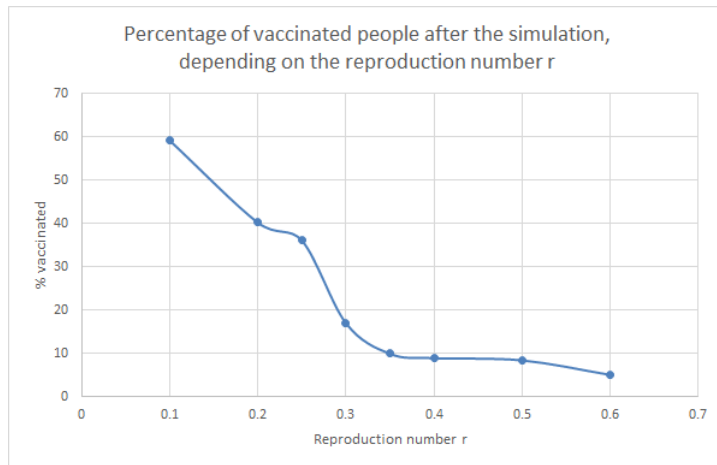


Figure: Percentage of vaccinated people and their reproduction number on the Scientific collaboration network dataset

Vaccinations result (cont)

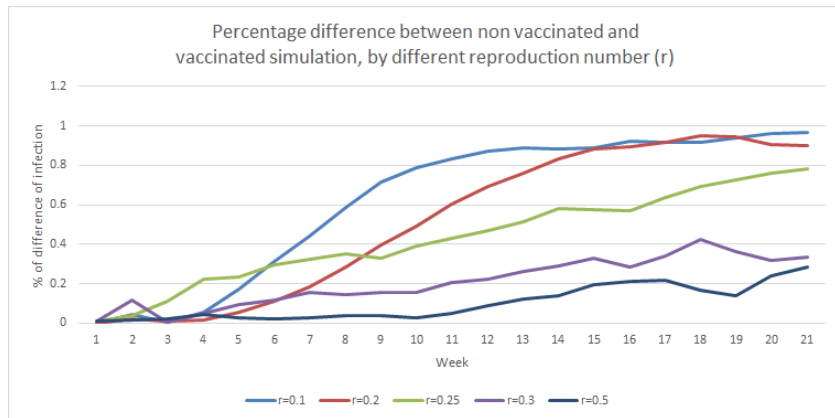


Figure: Percentage difference infected and vaccinated on the Scientific collaboration network dataset

Vaccinations result (cont)

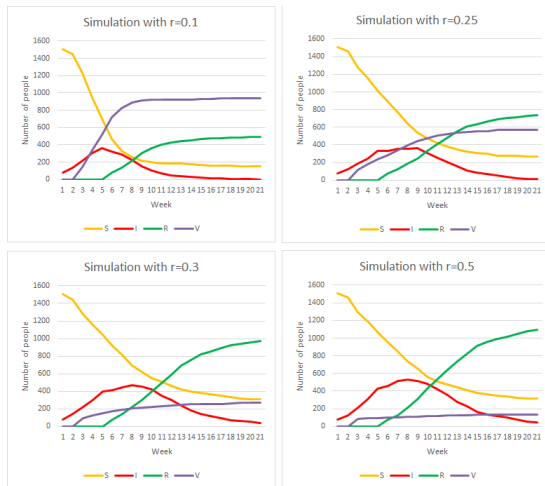


Figure: Plot of four run with different r on the Scientific collaboration network dataset

Overview

Introduction

SIR Models

Compartment model

Social Contact Networks

Datasets

Infection, vaccination and result

Hive Plot

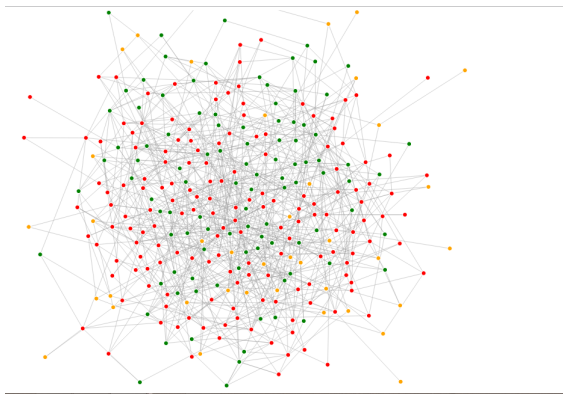
Human mobility

Conclusion

Why use Hive Plot and what is an Hive Plot

Why?

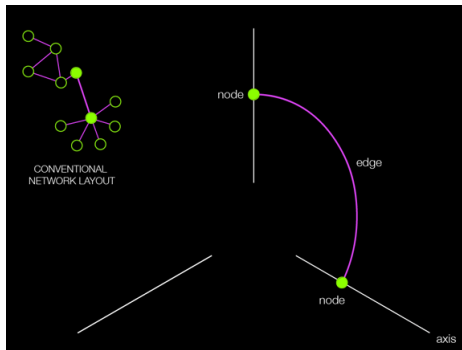
The network graphs in the datasets have a lot of nodes and edges, thus it is easy to have the hairball effect when we plot the entire graph in just one figure. The following picture is just a network with 300 nodes and 600 edges.



Why use Hive Plot and what is an Hive Plot (Cont)

What is an Hive Plot?

This visualization fixes the position of each node among three or more different axis and plots only the useful edges.



Hive plot in SIR model

- ▶ Not show the edge between the node which have the same label
- ▶ Be able to sort the nodes in some ways (e.g. sort by age, degree of outgoing edges)
- ▶ The orange nodes represents susceptible individuals (in the S axis)
- ▶ The red nodes represents infected one (in the I axis)
- ▶ The green represents recovered (in the R axis)
- ▶ The purple represents vaccinated (in the V axis)

We have create two type of visualization: one represent the current state of the graph and another one represent the evolution of the graph. This means show which nodes pass between different status.

Hive plot SIR model: current state

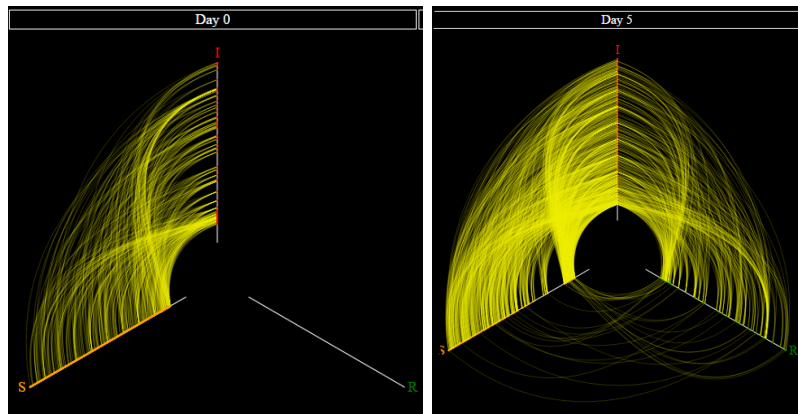


Figure: Current state of the graph. The hive on the left is the day 0 and the right the day 5

Hive plot SIR model: evolution of the graph

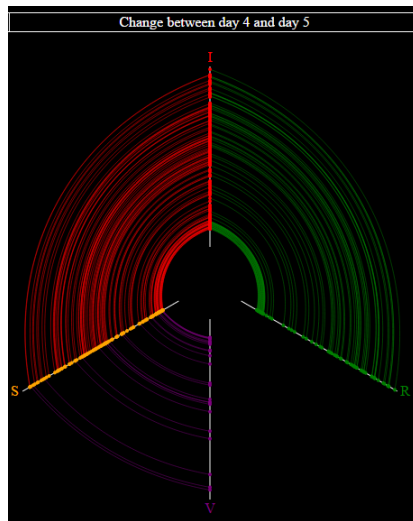
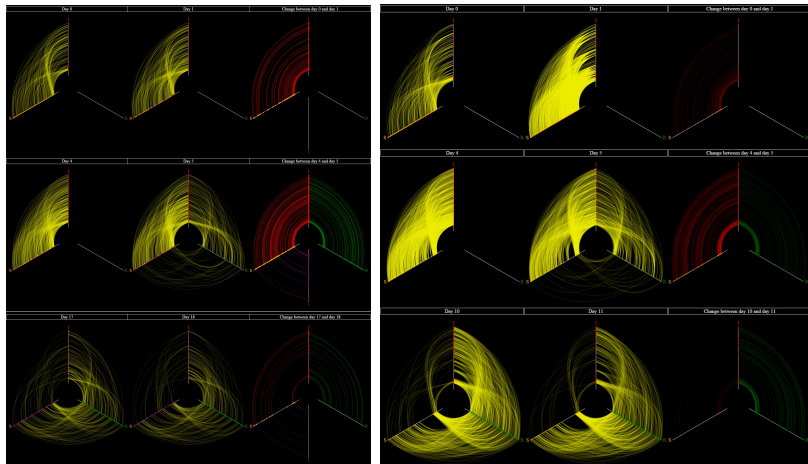


Figure: Represent the transaction between the node during the simulation

Hive plot SIR model: Final result



Overview

Introduction

SIR Models

Compartment model

Social Contact Networks

Datasets

Infection, vaccination and result

Hive Plot

Human mobility

Conclusion

Human mobility

- ▶ The human mobility is an important aspect in the world disease transmission
- ▶ Analyze this aspect through a flight travel simulation

We have two different layer of simulation:

- ▶ Passenger between the country, that is the flight network. Then, we use the Algorithm 2 (SIR Social contact network)
- ▶ Inside every country. Then, we use the Algorithm 1 (SIR compartment model)

Therefore we have 200 node (country) in the graph, and for each country we run the compartment model simulation.

Human mobility: datasets

We combine three different datasets:

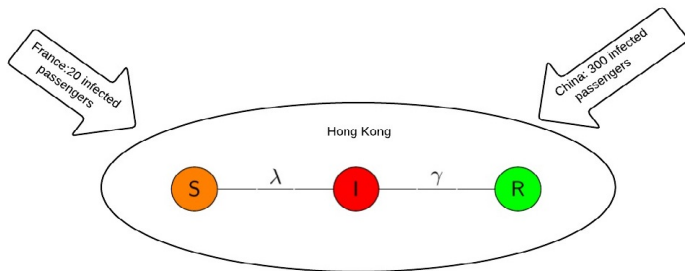
- ▶ Open flight, here we take all the information regarding the airport, such as: Name, City, Country, IATA/FAA, Longitude and Latitude
- ▶ Vbd-air, here we found information about the number of passengers for each route
- ▶ CIA, coordinate of the center of every country.

To combine this information we used Mysql and Python. The first to join different table, and the second to standardize and clean the data.

Just after this process we import that in Neo4j for the simulation.

Human mobility: implementation

- ▶ Use Neo4j for the network simulation
- ▶ Each node maintain three variable S,I and R
- ▶ Compartment simulation in each node
- ▶ Global clock

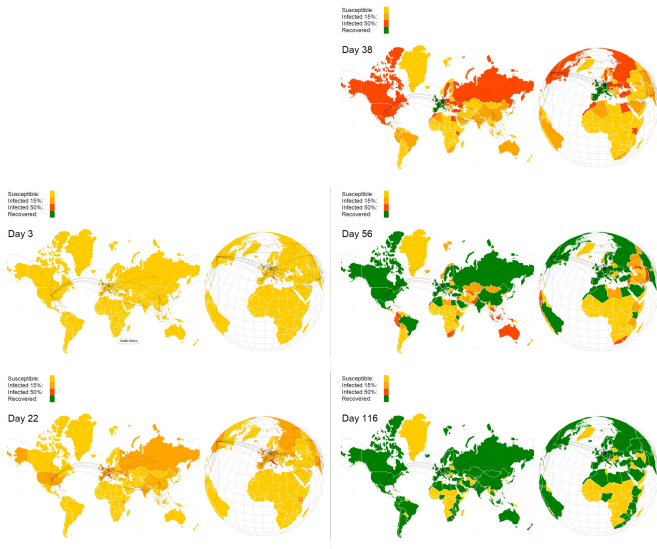


Human mobility: data visualization

To visualize the flight network and the disease propagation we use an info graphic map.

- ▶ D3js map framework
- ▶ Display a flat map and a globe map to better visualize the map
- ▶ Orange for the susceptible countries
- ▶ Dark orange for countries with at least 15% of the population is infected
- ▶ Red for countries with at least 50% of the population is infected
- ▶ Green for countries with at least 50% of the population is recovered

Human mobility: data visualization (cont)



Overview

Introduction

SIR Models

Compartment model

Social Contact Networks

Datasets

Infection, vaccination and result

Hive Plot

Human mobility

Conclusion

Conclusion

- ▶ We have analysed two different techniques to simulate the infection process
 - ▶ based on a social contact network
 - ▶ based on a compartment model
- ▶ We saw how the decision of an individual, depending on the cost of the vaccination and the neighbours' decision, influences the spread of the disease
- ▶ How human movement by flight influence the propagation of a virus
- ▶ we implement two methods to display the result in clear and understandable way
 - ▶ hive plot
 - ▶ info graphic map

Q & A