Would humanity survive a Spanish flu pandemic today?

NTDS project

EPFL

Doenz Jonathan   Esguerra Martin   Vojinovic Stefano

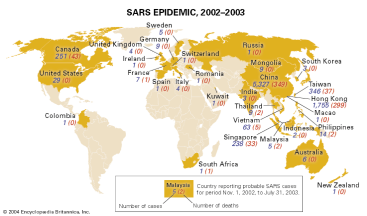
Overview

Introduction

Human existence has always been affected by diseases. As the world is more and more connected, the spread of a pandemic can be extremely quick and hard to track. We want to explore how a new illness could spread through the world and what control strategies could be implemented.

Spanish Flu

SARS epidemic



Dataset

OpenFlights/Airline Route Mapper Database

- 4 different datasets. (Airports, airlines, routes and planes)
- 7000+ airports initially, 3000+ after filtering
- 67000+ routes, a lot of routes repeated.
- Graph constructed using airports as nodes and routes as edges
- 11 connected in components resulting graph, 1 giant component.
- Graph follows a power law

EPFL

Airports network properties

- Only the largest component of the network is kept
- Only the *undirected* network is considered
- We refer to this network as the *airports network*

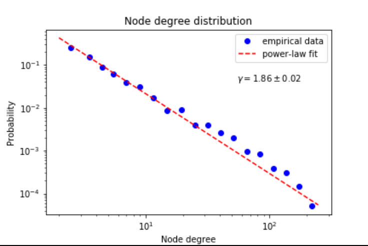
Properties of the airports network

Feature	**Value**
number of nodes	3186
number of edges	18832
average node degree $\langle k \rangle$	11.8
average squared node degree $\langle k^2 \rangle$	763.7
median node degree	4
average shortest path	3.958
diameter	12

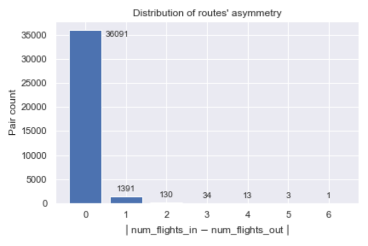
EPFL

Airports network properties

Degree distribution



Routes' asymmetry distribution

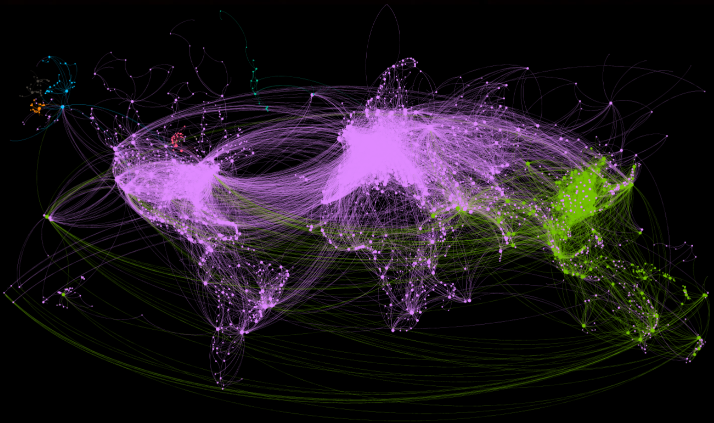


- The network is scale-free with $\gamma = 1.86 \pm 0.02$
- 95.8% of the routes are symmetric

EPFL

Spectral Clustering (SC)

- Clusters are based on graph properties
- One cluster contains the majority of the airports for any value of k



Dataset visualization performed by Spectral Clustering with 8 clusters

KMeans Clustering (KM)

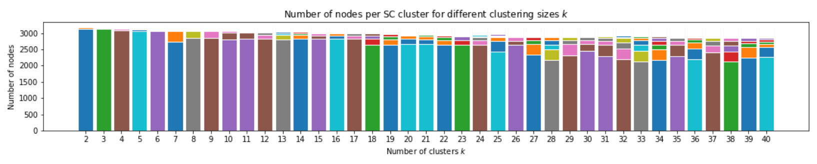
- Clusters are based on latitude and longitude
- Clusters sized are balanced

Dataset visualization performed by KMeans Clustering with 8 clusters

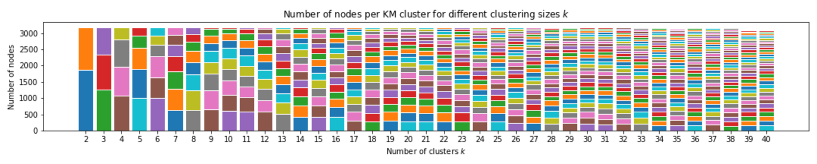
EPFL

SC and KM Clusters Size Comparison

Spectral clustering



KMeans clustering



EPFL

Epidemics Model

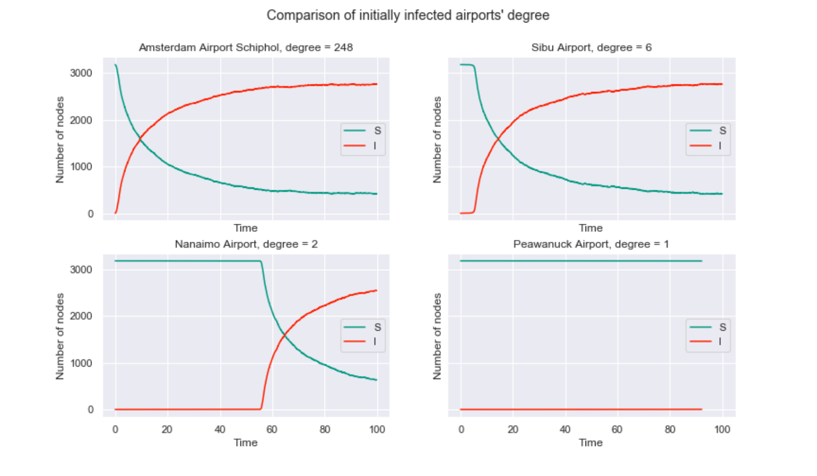
We use the *SIS* epidemics model

$$S \underset{\mu}{\overset{\beta}{\rightleftarrows}} I$$

We set the parameters β and μ in accordance with the Spanish-flu $R_0 = \frac{\beta}{\mu} = 2.5$
Model's interpretation:

- A susceptible airport i can be infected by an infected airport j with probability $\beta w_{i,j}$
- An infected airport can become susceptible at a rate μ

Spread Depends on Degree of Initially Infected Airport



Spread Depends on Degree of Initially Infected Airport

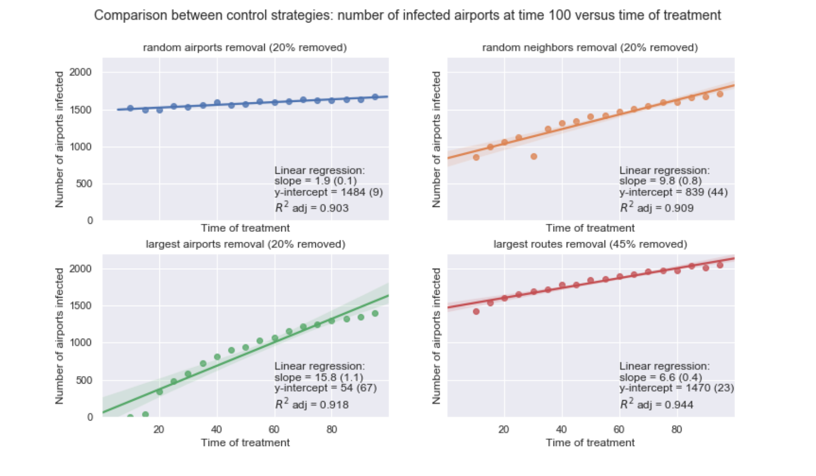
Endemic state reached: 87% of airports are infected

Control Strategies

1. `random_airports_removal`
   Shutdown of 20% of airports of the network. Airports are randomly chosen.
2. `random_neighbors_removal`
   Select 20% of airports at random, then shutdown a random neighbor for each of these.
3. `largest_airports_removal`
   Shutdown the top 20% connected airports.
4. `largest_routes_removal`
   Remove the top 45% connected routes.

EPFL

Control Strategies Comparison: Time of Treatment



EPFL

Control Strategies Comparison: Time of Treatment

- Random airports removal is quite ineffective (large y-intercept) and is not sensitive to the time of treatment either (small slope)
- Random neighbors removal is more effective (y-intercept = 839 vs 1484), and quite sensitive (slopte = 9.8)
- Remove largest airports is best: acting early completely disrupts the epidemics. It is the most sensitive too
- Remove (45% of) routes is not very effective. It is more sensitive than random airports removal

EPFL

Summary

- Most of the world's airports would be contaminated (87%)
- Remove most connected airports is the most effective strategy, it can disrupt the pandemics if applied early enough
- If a large fraction of the world's population succumbs, it is hard to predict whether humanity would survive

EPFL

Epidemics Simulation Example

Play epidemics spread animation if time allows!