Efficient simulation and graphical modeling of Covid-19 spread

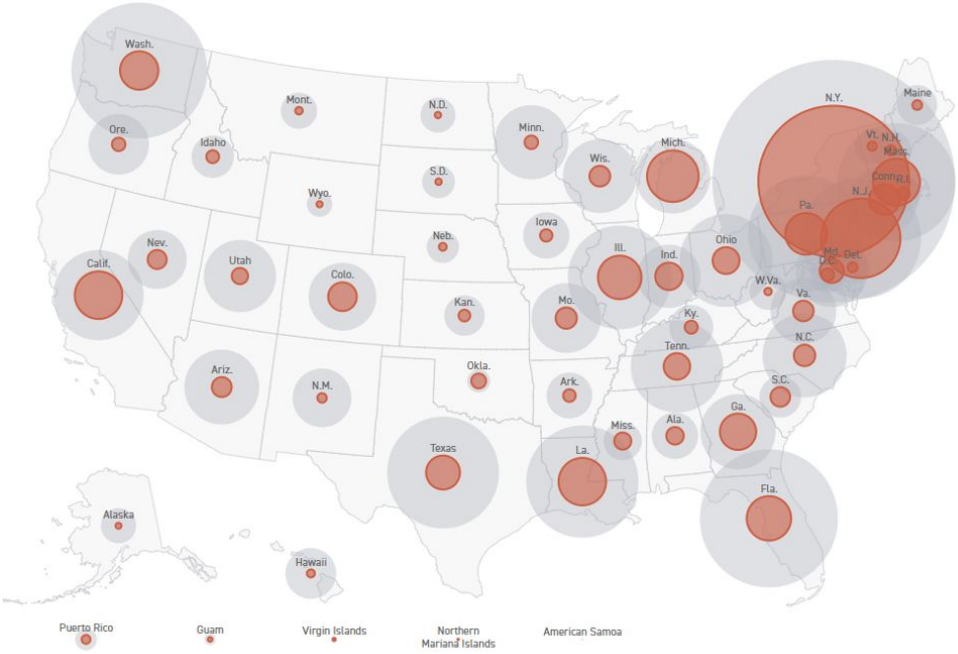
Group 13
Camille Bean, Intekhab Hossain, Hui Li and Stephanie Wu

Project Goals

Efficient simulation and graphical modeling of the spread of Covid-19 in a local community, using real Covid-19 incidents data and human interaction network data online.

We want to apply our simulation model to:

- ➔ evaluate the impact of preventative measures on public health outcomes;
- ➔ study the progression of Covid-19 in a new demographic context.



Coronavirus by state map. *"Live tracker: How many coronavirus cases have been reported in each U.S. state?" Politico.* March, 16, 2020.

Models/Data Sources : (1) Population Network

- ★ Obtained population network (list of edges) from an HIV survey done in the 1990s across **eight** different North American cities
- ★ Eight disjoint sub-networks which can be modelled parallelly
- ★ Overall: 49,355 nodes and 64,276 unique edges
- ★ Aim to augment this graph by adding random edges and nodes

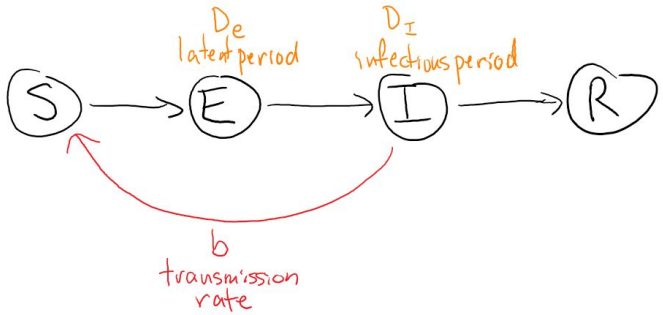
HIV Transmission Network Metastudy Project: An Archive of Data From Eight Network Studies, 1988--2001 (ICPSR 22140)

	ID1	ID2	TIETYPE	SEX1	SEX2	AGE1	AGE2	STUDYNUM
1:	1_1	106_1	sexual	female	female	27	37	1
2:	1_1	236_1	social	female	female	25	27	1
3:	1_1	266_1	social	female	female	26	28	1
4:	1_1	283_1	social	female	female	26	20	1
5:	1_1	337_1	social	female	female	25	29	1

64272:	143798_8	36441_8	sexual	female	male	21	17	8
64273:	143798_8	36442_8	sexual	female	male	21	-9	8
64274:	143811_8	36429_8	sexual	unknown	male	98	23	8
64275:	143829_8	36447_8	sexual	male	female	21	13	8
64276:	143829_8	143774_8	sexual	male	female	21	14	8

Models/Data Sources : (2) Epidemic Model

SEIR model characterizing disease states and transmission



The COVID Tracking Project

States:

- ★ Susceptible
- ★ Exposed
- ★ Infectious
- ★ Removed (recovered or deceased)

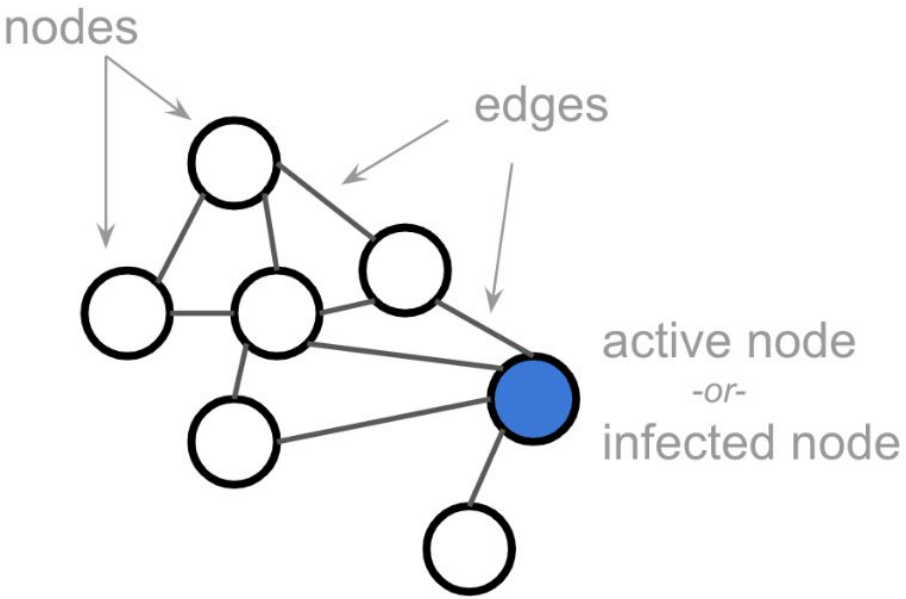
Parameters to estimate:

- ★ Transmission rate, b
- ★ Latent period, D_e
- ★ Infectious period, D_I

A need for big compute & big data parallel processing

- ★ Big compute will enable us to ***jointly*** model disease propagation in our network

- ★ Big data parallel processing will allow us to conduct multiple Monte Carlo simulations ***simultaneously***, while doing probabilistic sensitivity analyses.



Source: https://www.meltingasphalt.com/interactive/going-critical/

Infrastructure

- The project will be written in Python and run using AWS instances
- **Networks** will be supported by Spark-integrated GraphFrames on Hadoop
- To handle **big data** requirements, MPI will be used for parallelization
- OpenMP will be used to handle **big compute** requirements