The Development of Software Tools for Monitoring the Spread of Disease

By Albert Gerovitch, Andrew Gritsevskiy, and Gregory Barboy

Mentor: Dr. Natasha Markuzon

Introduction

It is of great interest to understand the spread of disease on a local level.

However, obtaining localized data and tracking and monitoring spread of disease of individuals is still a big challenge.

Goals

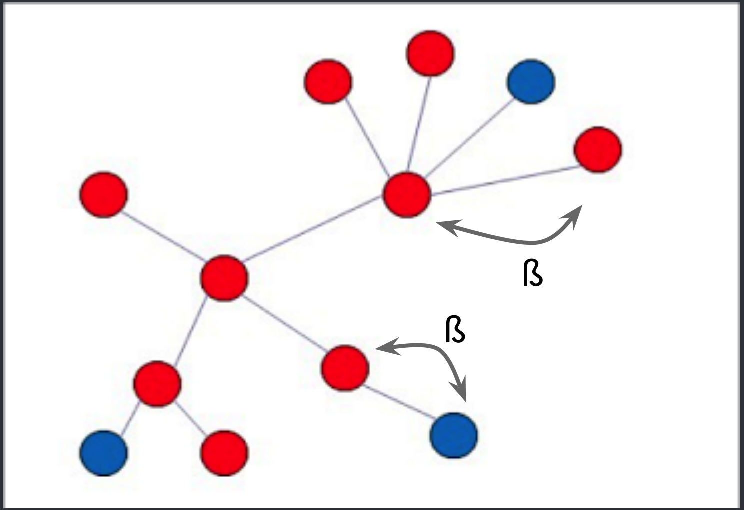
1. Enable collection of detailed local data
2. Develop models based on collected and generated data
3. Monitor and predict the spread of a disease on a daily basis
4. Perform analysis on both large and small-scale networks

Human Networks

Human Networks

Three Types of Networks:

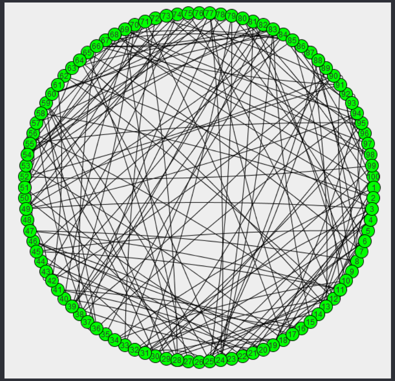
- Random
- Small World
- Scale-Free

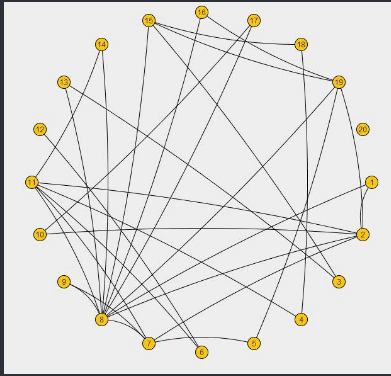


Human Networks - Random Network

- Random Connections
- Algorithm:
  - Generates a number of connections
  - Distributes them among people following distribution guidelines for each individual person

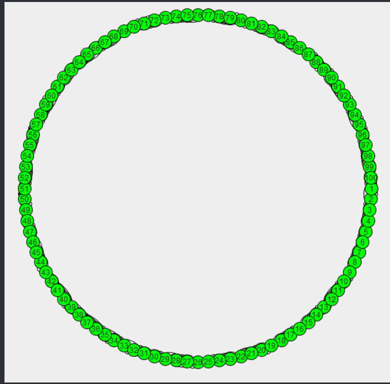
- Examples: *Handshakes at a party*

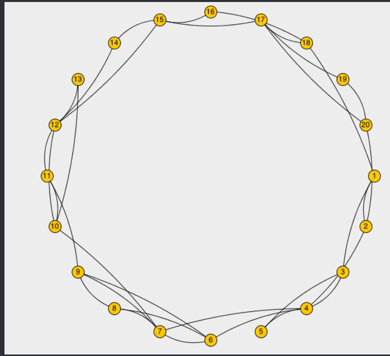




Human Networks - Small World Network

- Local Connections
  - Close, nearby on the graph
- Range (*k*) of possible friends for each person (*n*)
  - Person *n* randomly befriends people *(n+k)* to *(n-k)*



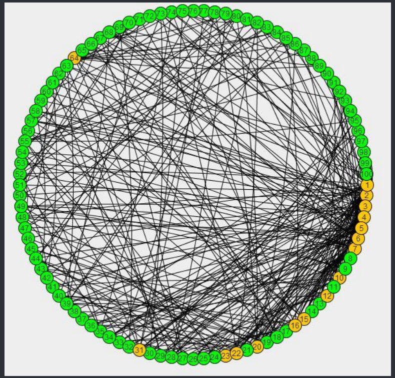


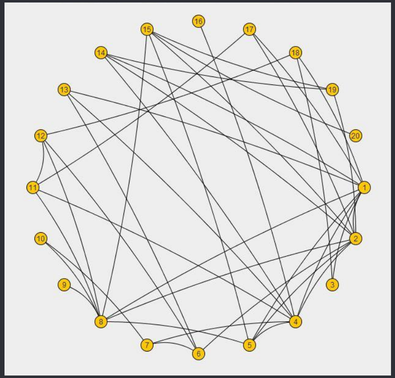
- Examples: *Wikipedia links, Food Chains*

Human Networks - Scale-Free Network

- Network built by preferential attachment
  - The more connections a node has, the more likely a new node will connect to it
  - Defining characteristic:
    - Popular nodes (hubs)

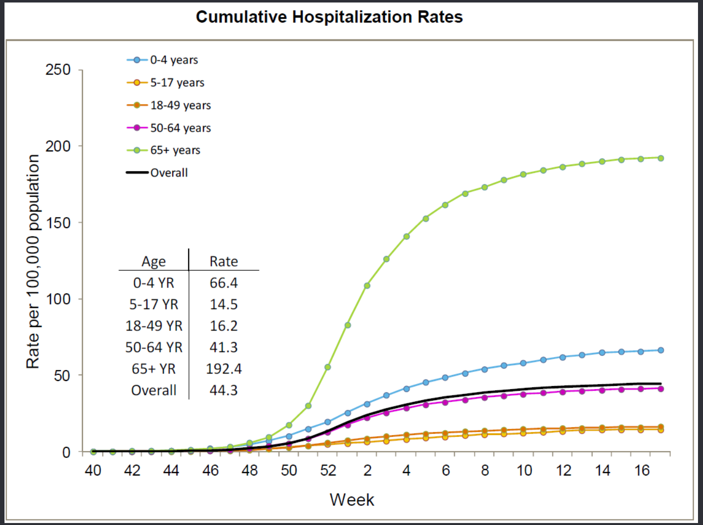
- Examples: *The Internet, Semantic Networks*





Census Data Integration

- Model Network off of town
  - Number of Households / Residents Per Household
  - Age Distributions
    - People of different ages react differently to disease



This will be discussed in more detail later

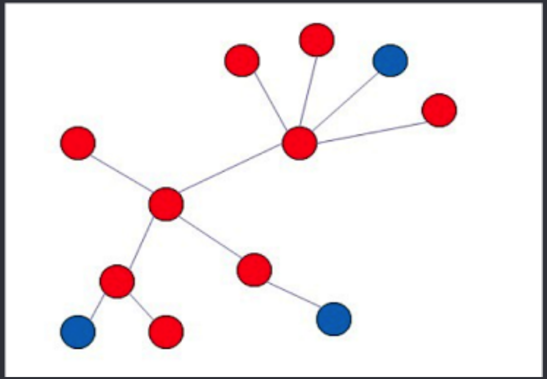
Model

Modelling disease spread

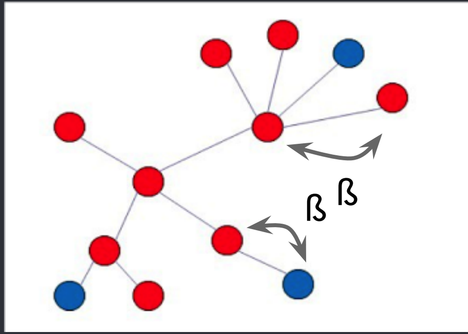
Three Components of the Model

Three components of our model:

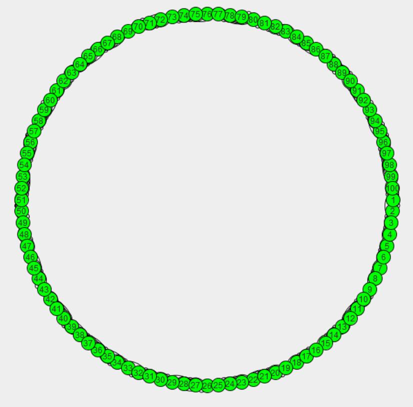
1. Undirected graph (network type)
2. State of each person (S-I-R)
3. Type of interactions between people on the graph

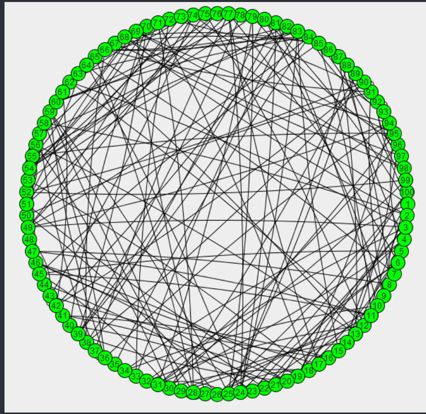


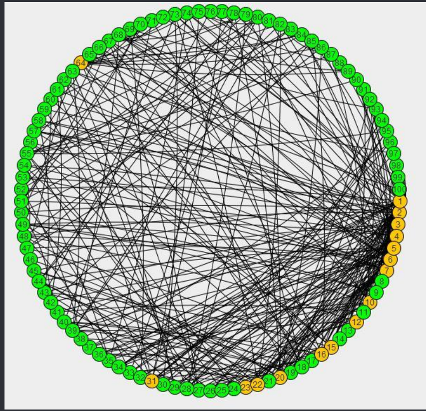
Three Components of the Model



1. Undirected graph (network type)
   a. Small-world
   b. Random
   c. Scale-free
   d. Census-based (a, b, c)
2. Type of interactions between people on the graph
   a. Disease spreads via probabilistic interactions between people



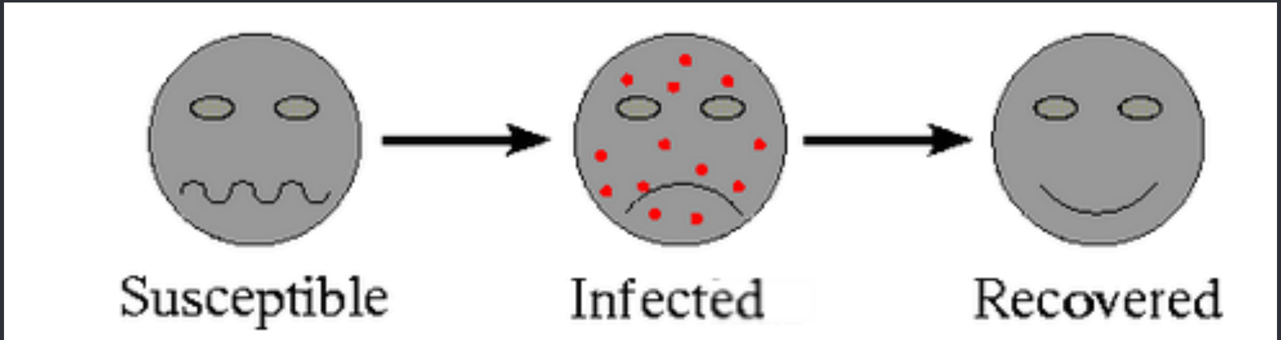




Three Components of the Model

Each person is in one of three states:

1. Susceptible *(not yet sick)*
2. Infected *(sick)*
3. Recovered *(immune)*



Simulations

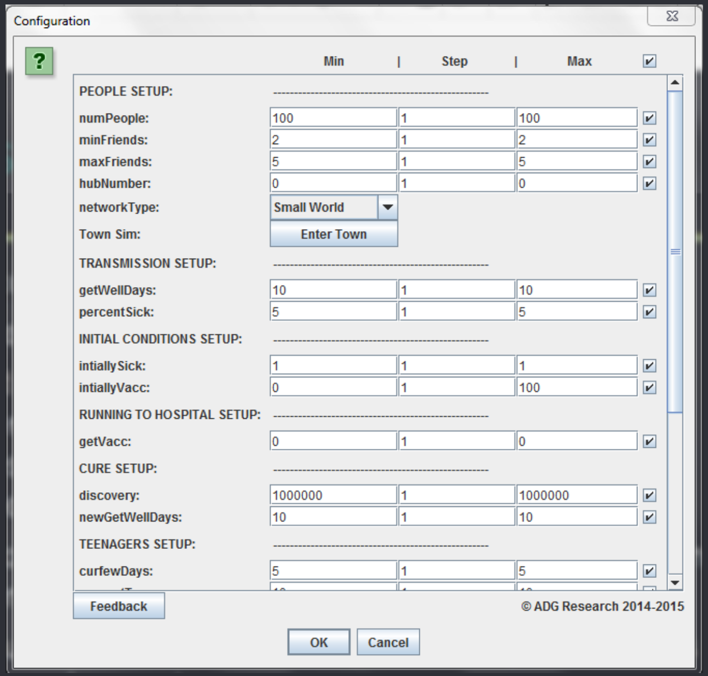
Simulations and Results

Our Simulation

User can enter a variety of parameters

- Structure of network
  - Type of network
  - Size of network, etc
- Interventions:
  - Vaccination
- Census information
  - Town or city name

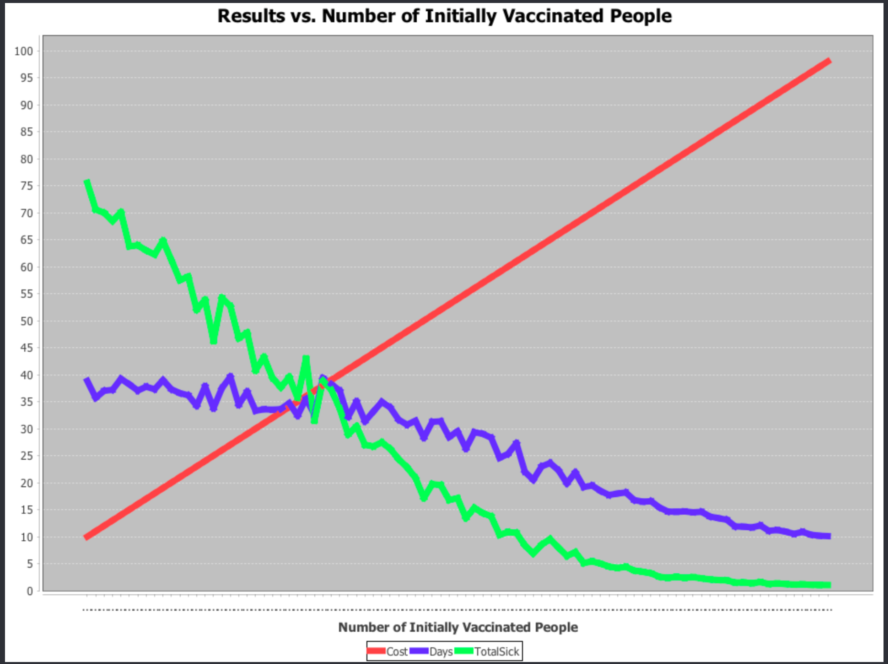
Computational duration significantly increases with number of people



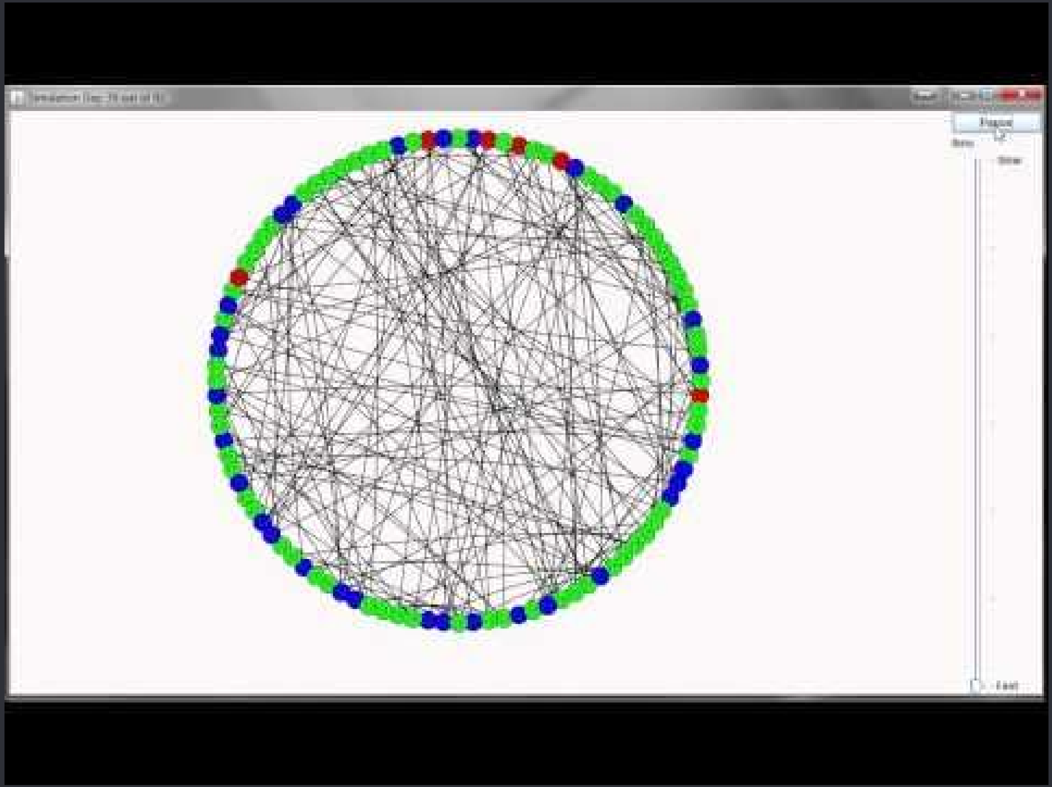
Automated Parameter Demonstration

Automated Parameter Demonstration

- The program generates a graph
- On the x-axis, there is the number of initially vaccinated people
- The red line represents cost
- The blue line represents length of the epidemic
- The green line represents total number of people infected during the epidemic



Jung Visualization Demonstration



Different Types of Networks

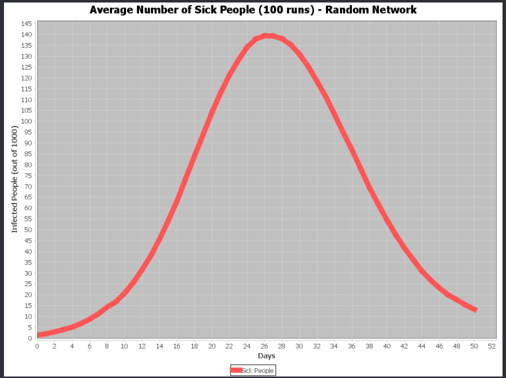
1. We analyzed simulation results using different types of networks
2. We used the following parameters:
   - 1000 people
   - 3-7 friends
   - 10% chance of infecting a friend
   - 1 person initially sick
   - 2 people initially vaccinated
   - A person recovers after 10 days of being sick
   - If your friend is sick, there is a 10% chance that you will immediately get a vaccination

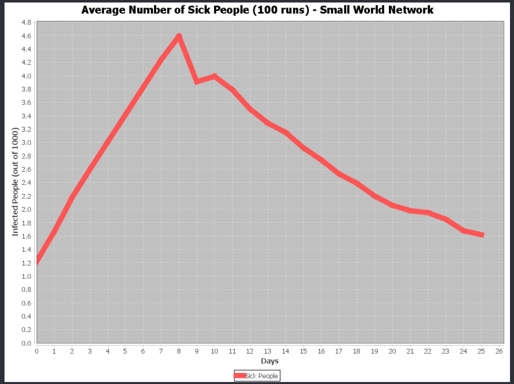
Different Types of Networks - Comparison

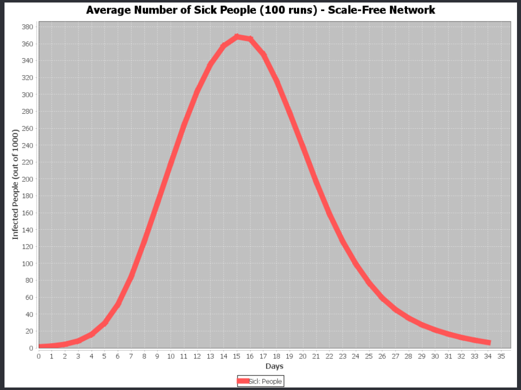
Random
- 140 total
- 50 days

Small-world
- 4.6 total
- 25 days

Scale-free
- 370 total
- 34 days







Census-based simulations

- In order to make our simulations more realistic:
  - We developed a program to automatically download data from American FactFinder
    - User enters town/city and state
  - Run simulations on networks automatically constructed from census data

AMERICAN FactFinder

Geography: Cambridge city, Massachusetts

1 - 40 of 40

Subject	Number	Percent
TENURE		
Occupied housing units	44,032	100.0
Owned with a mortgage or loan	10,878	24.7
Owned free and clear	4,357	9.9
Renter occupied	28,797	65.4
TENURE BY HOUSEHOLD SIZE		
Owner-occupied housing units	15,235	100.0
1-person household	5,600	36.8
2-person household	5,535	36.3
3-person household	2,120	13.9
4-person household	1,357	8.9
5-person household	441	2.9
6-person household	114	0.7
7-or-more-person household	68	0.4
Renter-occupied housing units	28,797	100.0
1-person household	12,333	42.8
2-person household	9,776	33.9
3-person household	3,915	13.6
4-person household	1,825	6.3
5-person household	628	2.2
6-person household	209	0.7
7-or-more-person household	111	0.4

Census Based Network Generation

Automatically generate a network using our network generation tool along with census data:

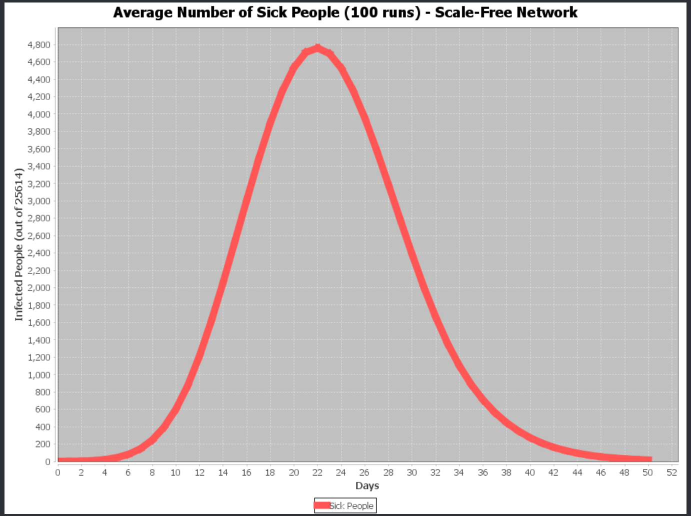
1. Household Information
   a. Number of Households
   b. People per Household
   c. All people within Household are connected
2. Age Distribution Data
   a. Susceptibility and duration of disease change based on age

Automated Census Simulation Results

Lexington, MA

Total population: 25 614

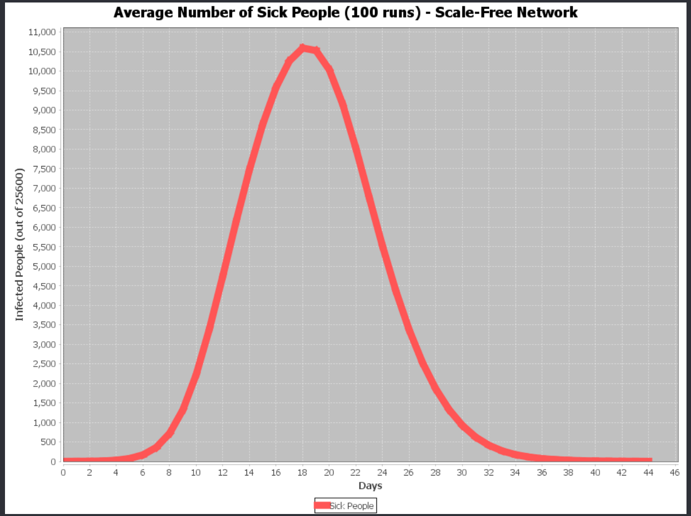
4,800 Sick (20%), 50 Days



Simulation without Census data

Total Population: 25 600

10,750 Sick (44%), 44 Days



Interventions

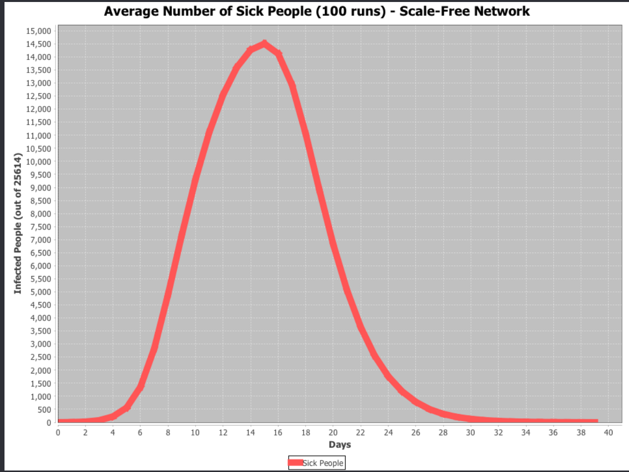
How do vaccinations affect the spread of the epidemic?

We introduce an intervention feature to make people initially immune (state R in SIR model)

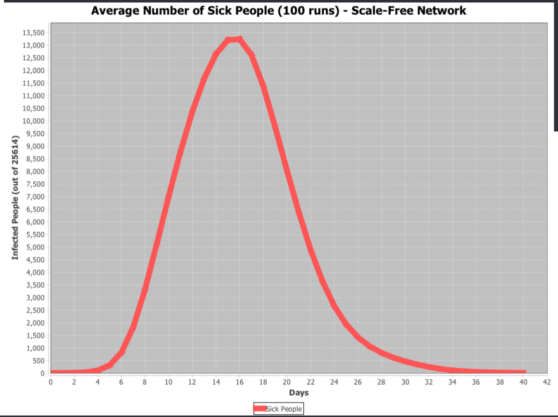
Interventions (Lexington)

Effect of untargeted vaccinations in Lexington

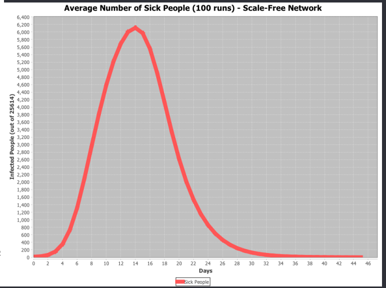
0% - ~14600 Sick



10% - ~13200 Sick



50% - ~6200 Sick



Strii

Scientific and **T**echnological **R**esearch for the **I**nnovation of **I**nfections (STRII)

Collecting disease data in real time

Crowdsourcing Application

Can we collect Individual’s health statuses and locations?

a. Develop a crowdsourcing application
    i. Allow users to enter health status daily
    ii. Detect geolocation of users
    iii. Save results in a cloud
    iv. Model each day with Google Earth
b. Use collected data to predict disease spread

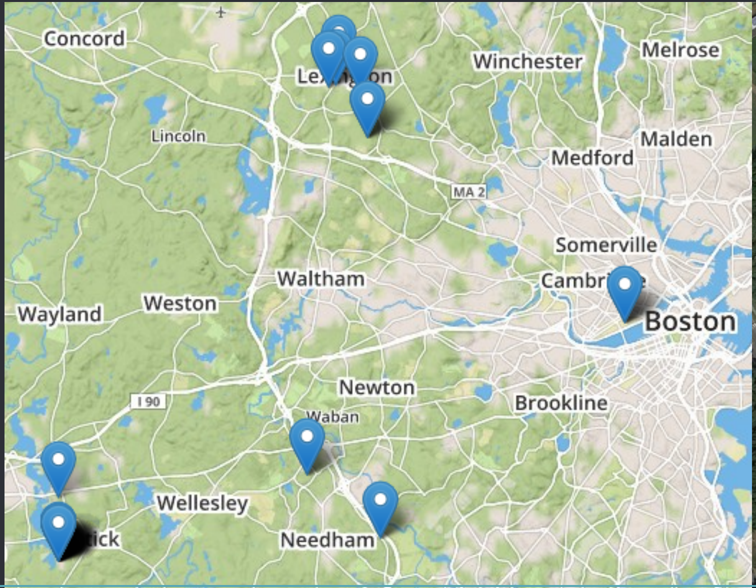
Strii

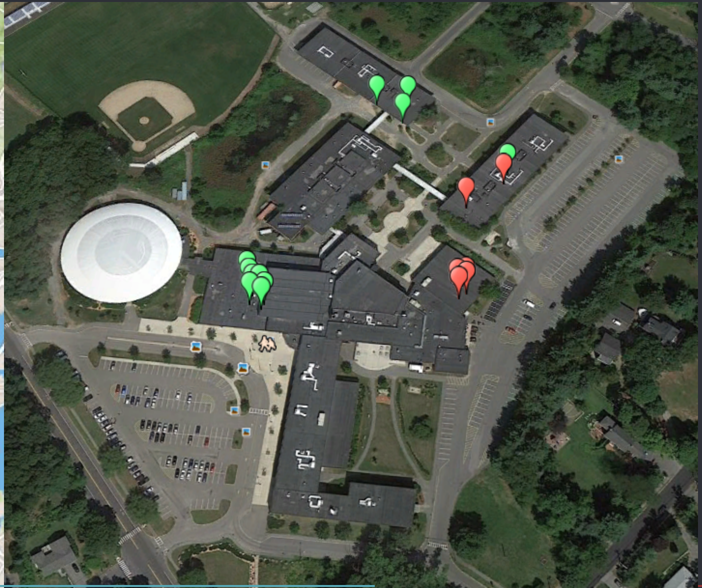
- User connects to the cloud application and updates his/her daily status in the database
- We use the database of information to analyze any potential epidemics
  - View clusters with increased incidence
  - Trace the epidemic to the origin
  - Calculate optimal nodes for vaccination

Strii website

Google Earth Example

- An example of our software and its accuracy on large and small scales





Strii website

- One day snapshot of our database
- Stored in the cloud
- Automatically updates with new info

FIRST_NAME	LATITUDE	LONGITUDE	D417
Alik	42.359548	-71.090982	1
Natasha	42.28347977378744	-71.20805166016626	1
Andrew	42.425036	-71.214023	0
Strii	42.4249805	-71.2140268	1
Grishka	42.305546	-71.2427909	1
Lelya	42.305631299999995	-71.24281069999999	1
Valya	37.4638072	-122.1670806	1
Chris	37.463783299999996	-122.1671384000001	1
Yvette	0	0	1

Strii website - Instructions for joining

- Go to e1em.com/join
- Enter name and join!

Yes, many of you can probably break our demonstration with a moderately simple SQL injection.

Please don’t.

Strii website - Map

Live map demonstration

- The online map has live location updating, unlike our application
- e1em.com/map

Conclusion

Conclusions

- Developed an automated tool for simulating disease spread
  - Analyzed different types of networks
  - Tried out different vaccination strategies
- Developed an automated system to interpret Census data
  - Ran simulations on a more realistic model
- Developed crowdsourcing application (strii)
  - Collected very specific and localized data about disease spread

Future improvements

- Incorporate Strii data in the developed predictive models and simulation
- Make Strii application widely available
  - Develop Android applications
- Gain more Strii data
  - Allow millions of users to access application

Special Thanks to:

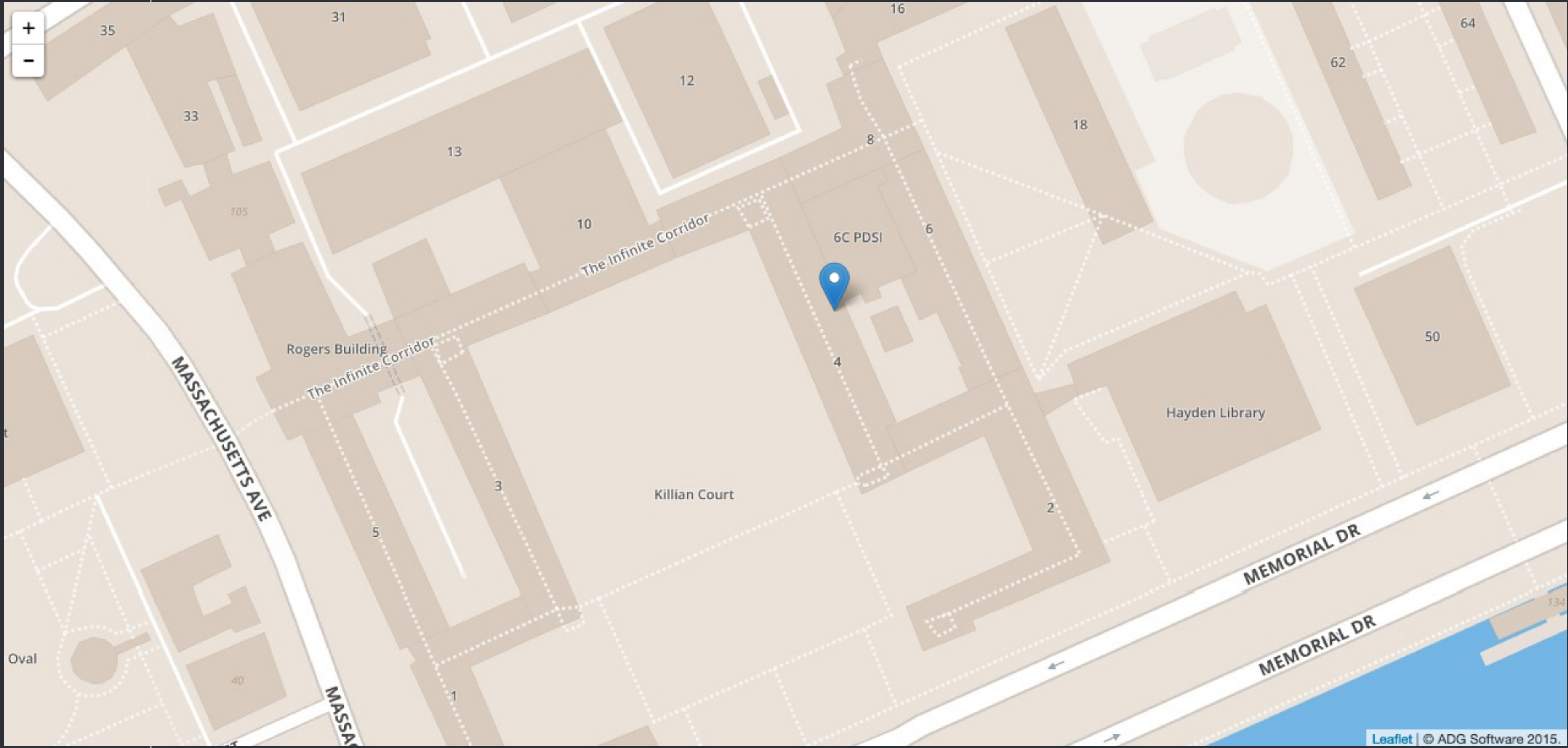
- Dr. Natasha Markuzon, our mentor, for suggesting the project and guiding us through the process
- MIT PRIMES program, and in particular Dr. Slava Gerovitch and Dr. Pavel Etingof, for making this wonderful experience possible and being a wonderful dad
- Everybody who helped us collect data by joining Strii
- Parents, for supporting us and driving us to meetings
- Skype for allowing our communications
- Bim for being our executive canine specialist in moral support
- Google and leafletjs for their beautiful APIs
- Electron Neutrino for hosting our strii cloud database
- Edward Boatman for creating the lock icons
- Our audience, for listening (or at least pretending)

No Thanks To:

- Apple, or any affiliated programs

Thank you!

Questions?



35
31
16
64
62
12
33
13
8
18
105
10
The Infinite Corridor
6C PDSI
6
50
Rogers Building
The Infinite Corridor
4
MASSACHUSETTS AVE
Hayden Library
3
Killian Court
2
5
MEMORIAL DR
MEMORIAL DR
Oval
40
1
MASSA
Leaflet | © ADG Software 2015.