Equal Graph Partitioning of an Estimated Infection Network

CHEONG Siew Ann

cheongsa@ntu.edu.sg

http://www1.spms.ntu.edu.sg/~cheongsa/

Motivation

- Social impact of severe respiratory diseases
  - Spanish Flu (1918)
    - 50-100 million deaths
  - SARS (2003)
    - 8273 cases worldwide
    - 775 deaths
  - H5N1 (2005)
    - Pandemic warned
  - H1N1 (2009)
    - 1.6 million cases worldwide
    - 14,378 deaths

Motivation

- Economic impact of severe respiratory diseases
  - SARS (2003)
    - Singapore Airlines
      - S$4.13B revenue in 2002
      - S$2.87B revenue in 2003
    - Singapore economy
      - −3% GDP in 2003
  - H1N1 (2009)
    - Economic cost being estimated

Epidemiology

- Biology (10%?)
  - Viral pathogen
  - Target cells lining respiratory tract
  - Viral particles re-emitted in bodily fluid droplets
  - Latent phase (~3 days)
    - Already contagious
  - Infected phase (~7 days)
    - Symptomatic vs asymptomatic
  - Self-limiting through immunological response
    - Immunity period (~1 year) determined by evolution rate of virus

Epidemiology

- Sociology (90%?)
  - Social proximity necessary
    - Inhalation of viral particles
  - High population density
    - Low mixing rates of subpopulations
    - SIR framework does not apply
  - Complex social interactions
    - Co-workers, classmates, social activities
    - Buses, trains, planes, ships
    - Shopping, dining, leisure gatherings

Network Approach

- Contact tracing
  - SARS (2003)
  - H1N1 (2009)
- Infection network estimation
  - Surveys (Xiao, NTU)
    - Manual, partial
    - No anonymous social interactions
  - Cellular phone collocation (Marathe, Virginia Tech)
    - Automatic, total?
    - Inadequate spatial resolution
    - Biological barriers to infection

Our Idea

- Common cold vs rare severe respiratory diseases
  - Biologically very similar
  - Sociologically nearly identical
  - Identical or very similar infection network
- Use common cold incidences to estimate network
  - No-intervention benchmark
  - One epidemic every 2-3 months

Proof of Principle

- Database of incidences
  - Does not exist for common cold
  - Generate from artificial social network
  - Censor data
- Network estimation
  - Tentative links weaken with time
  - Real links reinforced periodically
- Network intervention
  - Equal graph partitioning

Artificial Social Network

- Jin, Girvan & Newman, 2001
  - Saturation
    - $k(i) \leq k_{max}$ for all nodes i
  - Link formation
    - Randomly select i and j
    - If $k(i)$, $k(j) < k_{max}$, form link with probability P_0, if i and j have no mutual neighbors
    - Else form link with probability $P_1 > P_0$
  - Link deletion
    - Randomly select i
    - Delete one random neighbor of i with probability Q

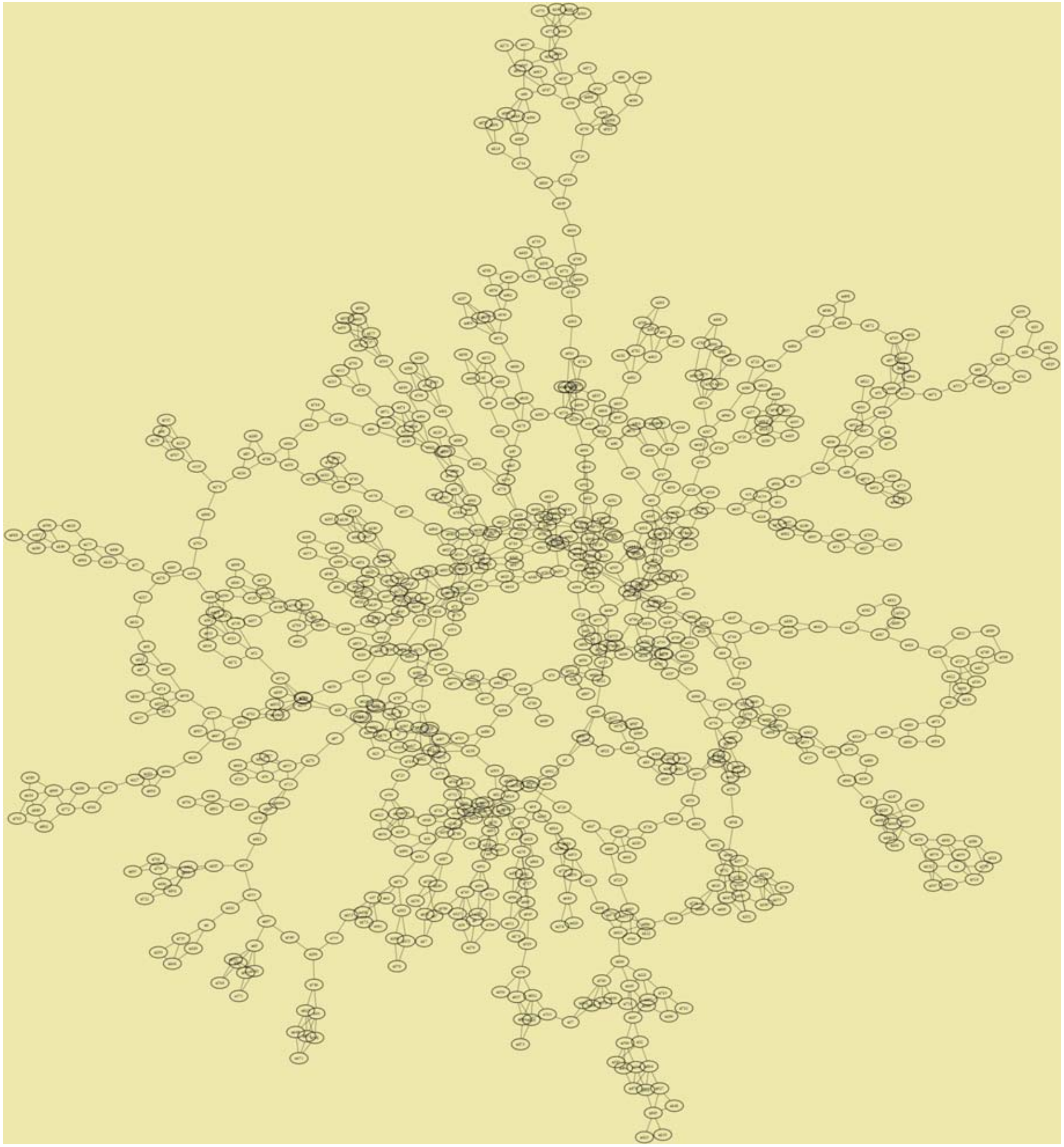
$N = 1000$

$\langle k \rangle = 3$

$P_0 = 0.10$

$P_1 = 0.40$

$Q = 0.552$



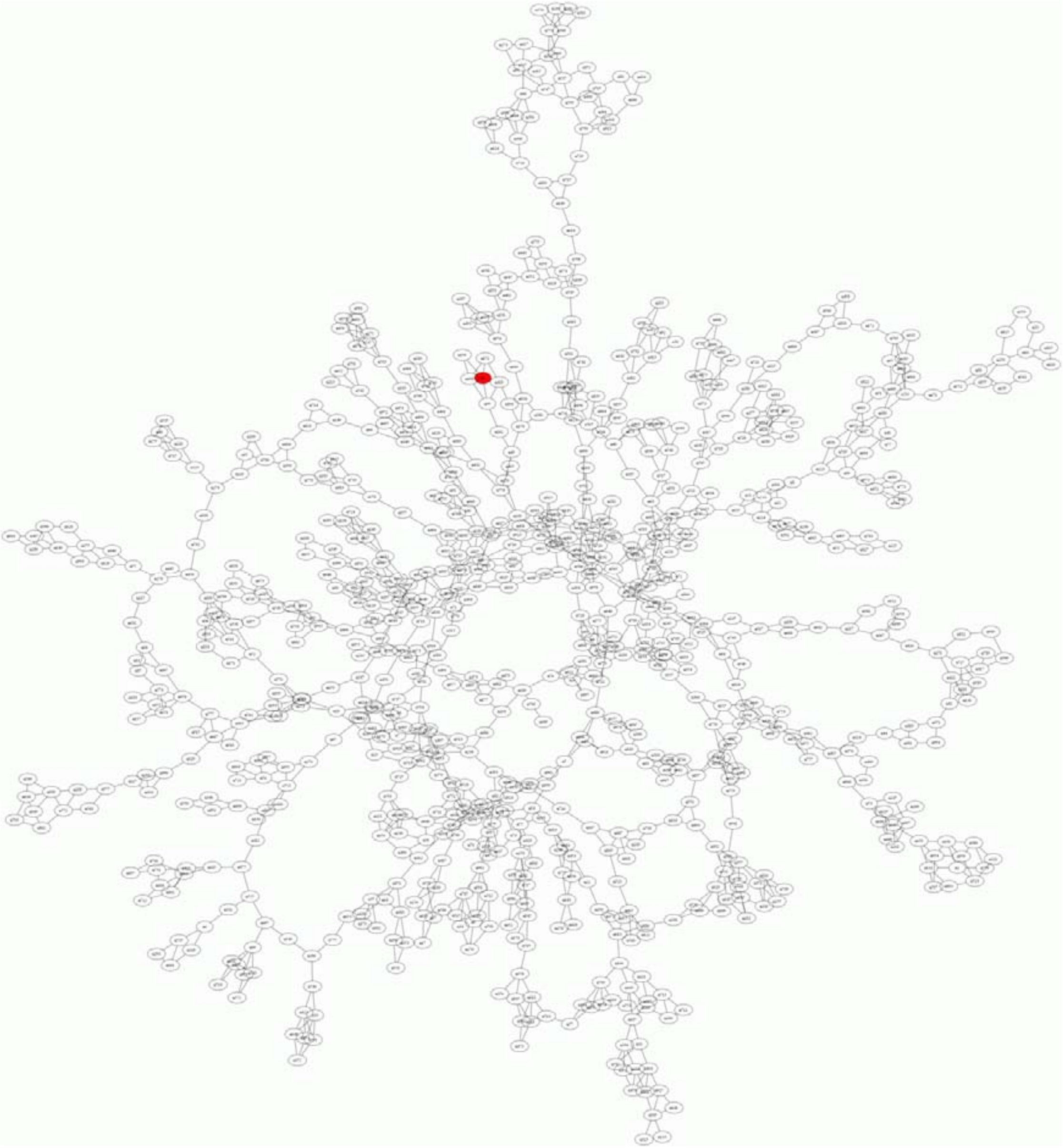
SIR Epidemics

- Epidemic
  - At $t = 0$, all nodes susceptible (S)
  - At $t = 1$, one random node infected (I)
  - At each $t > 1$, susceptible neighbors of infected nodes infected with probability p
  - Each infected node recovers (R) after $t_R = 1$ time step
- Censorship
  - List of infected nodes at each t
  - Only fraction $f < 1$ used for estimation
    - Simulate low level of reporting

N = 1000

<k> = 3

p = 0.8



Network Estimation

- Link Formation
  - Tentative link
    - between ALL reported nodes at time step t and ALL reported nodes at time step $t - 1$
    - Weight $w = 1$
- Link Reinforcement
  - Over S epidemics
    - False links formed once or twice
    - Real links formed $O(S)$ times
- Link Elimination
  - Weights of ALL links decay at constant rate
    - Help keep background of false links low

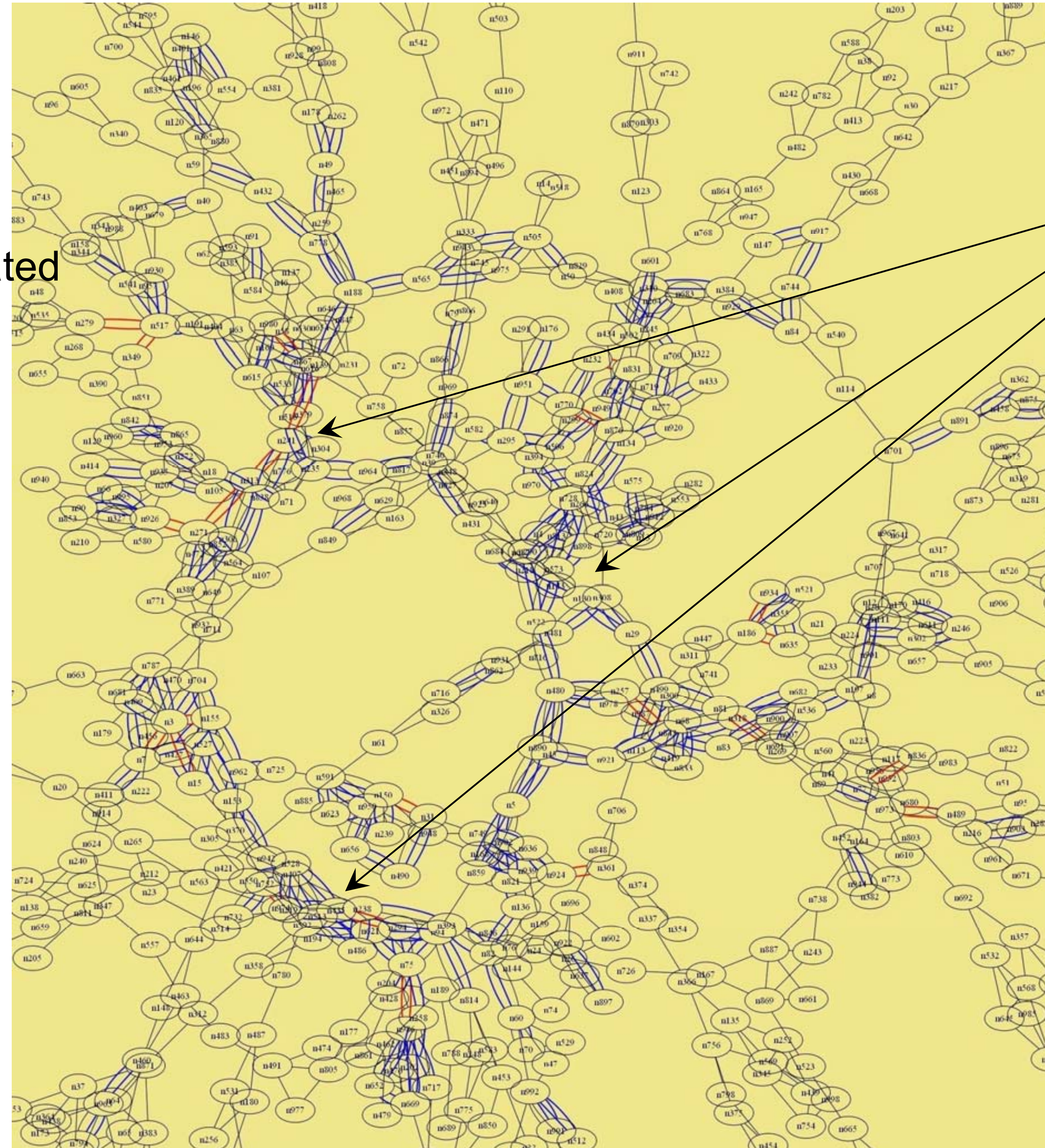
S = 100

f = 0.3

Not estimated

Correct

Wrong

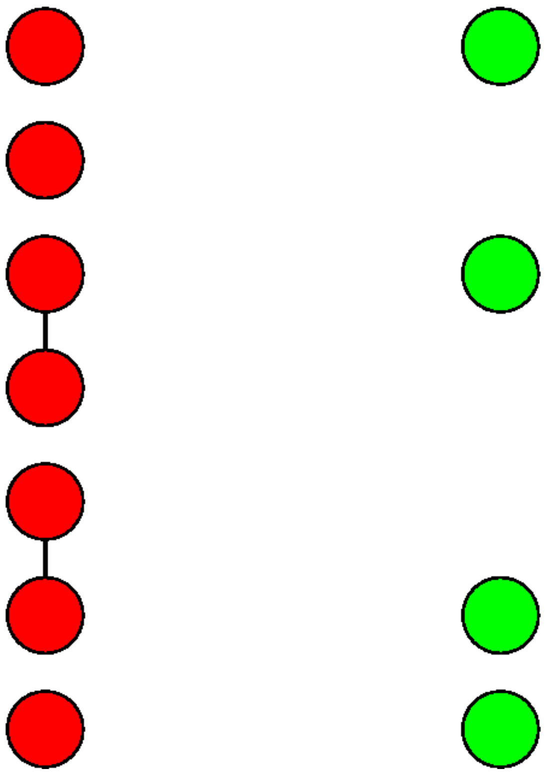


backbone of infection network

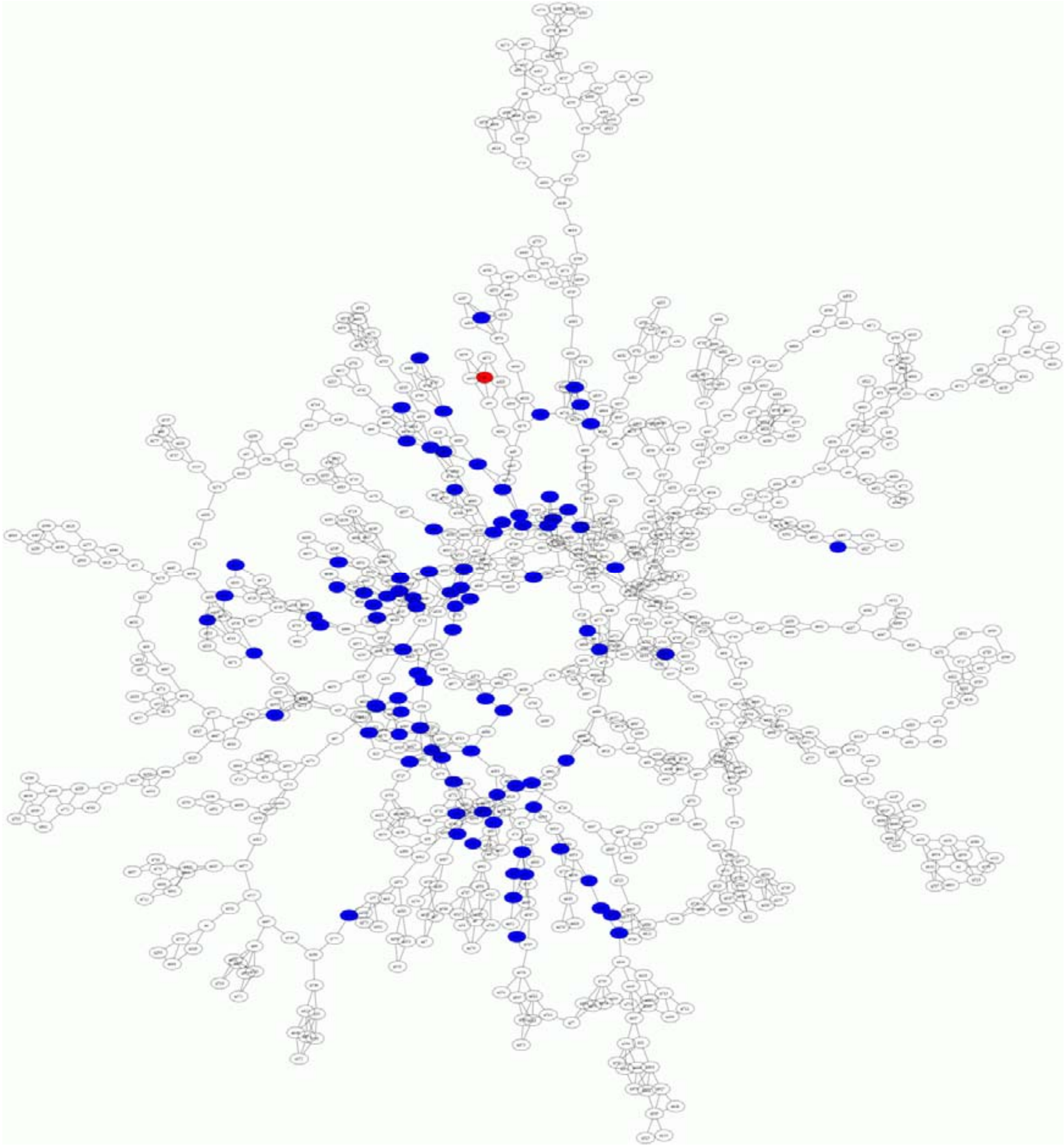
Network Intervention

- Nested Dissection (Lipton, 1979)
  - Efficient solution of sparse linear systems
- Equal Graph Partitioning (Chen, 2008)
  - Efficient immunization of completely mapped complex network
- Our question
  - Will EGP be effective:
    - Partially mapped network?
    - Errors in mapped network?
    - Epidemic in progress?

Equal Graph Partitioning



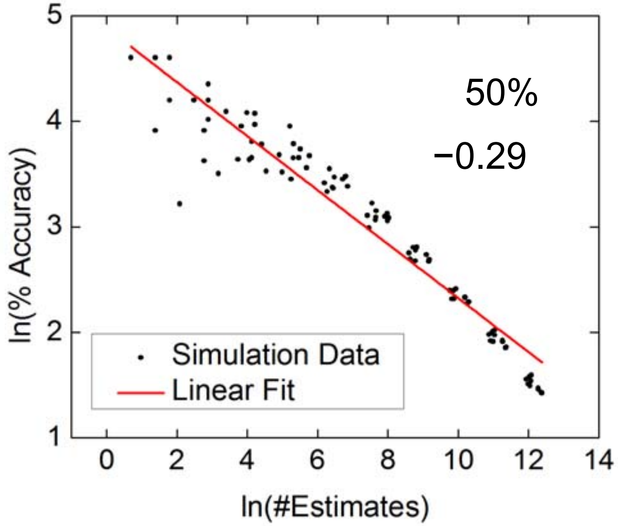
quarantine
healthy
individuals!



Systematic Study

- Estimation accuracy vs
  - Number of estimated links
    - N = 10k; $<k>$ = 10; c = 0.05
  - Censor rate $(1 - f)$
    - N = 10k; $<k>$ = 10; c = 0.05
  - Number of epidemics S
    - N = 10k; $<k>$ = 10; c = 0.05
  - Network sizes
    - N = 1k, 10k, 100k; $<k>$ = 10; c = 0.05
  - Average degree $<k>$
    - N = 10k; $<k>$ = 10, 20; c = 0.05

Number of Estimated Links

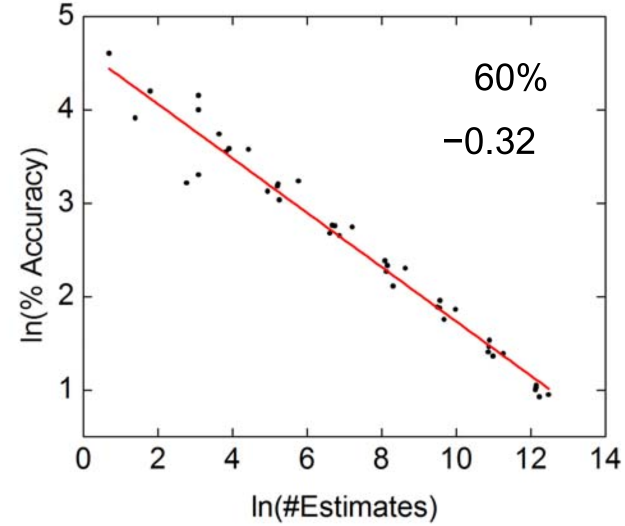


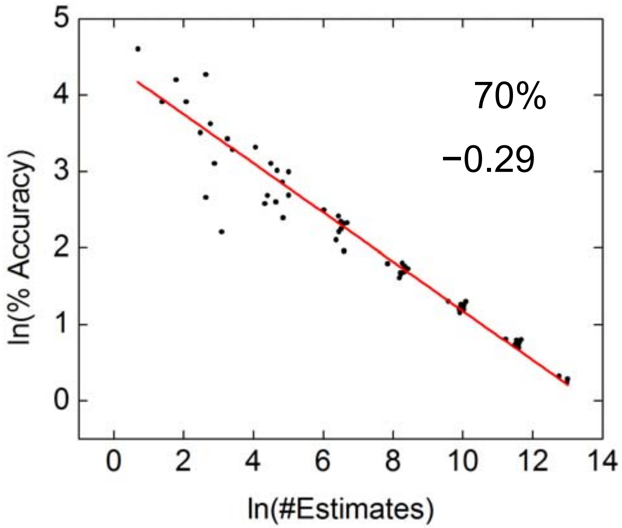
N = 10k

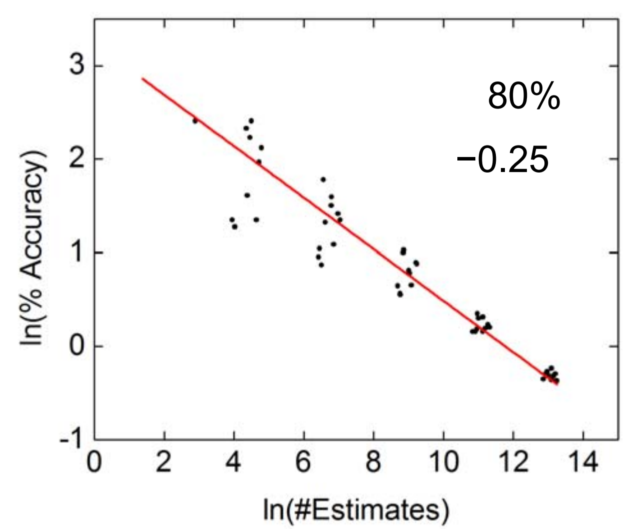
$\langle k \rangle$ = 10

c = 0.05

S = 100







Censor Rate

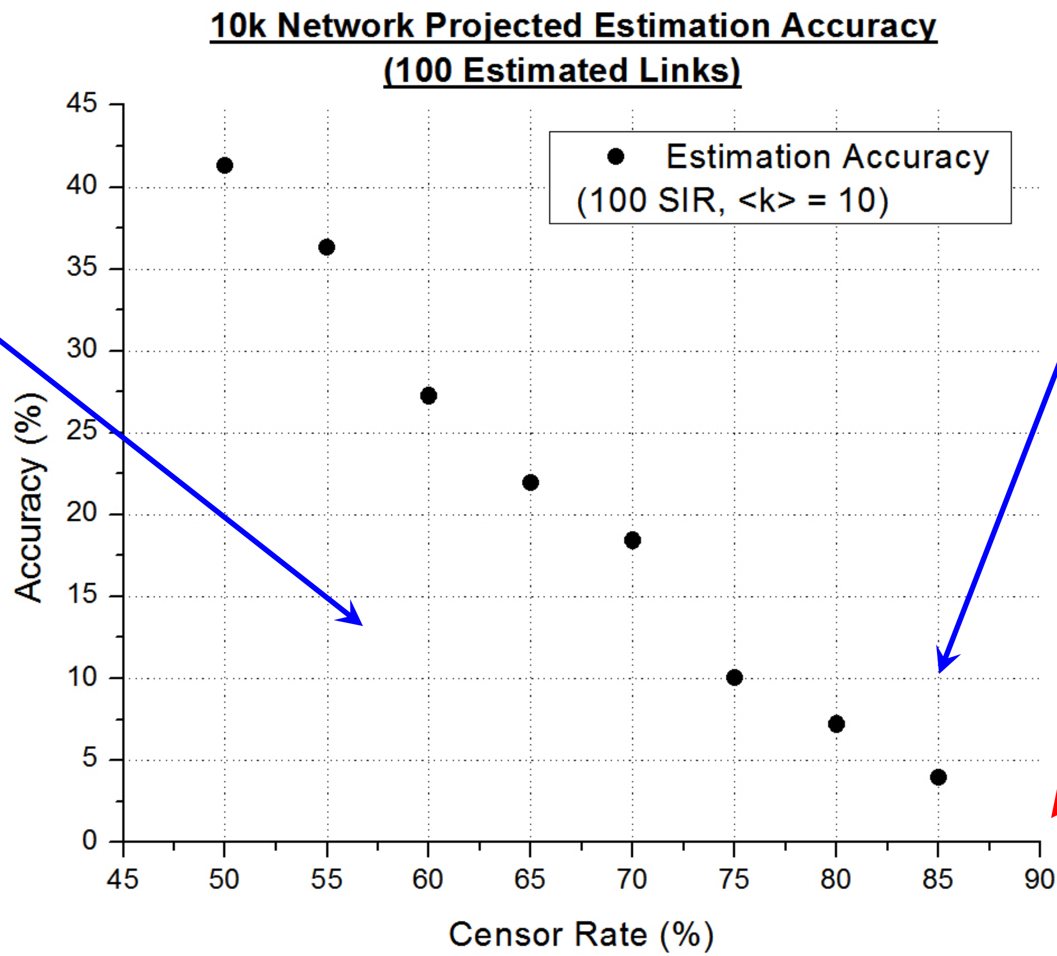
Visit to GP cheap in Singapore = low censor rate

N = 10k

$<k>$ = 10

c = 0.05

S = 100



Visit to GP expensive in US = high censor rate

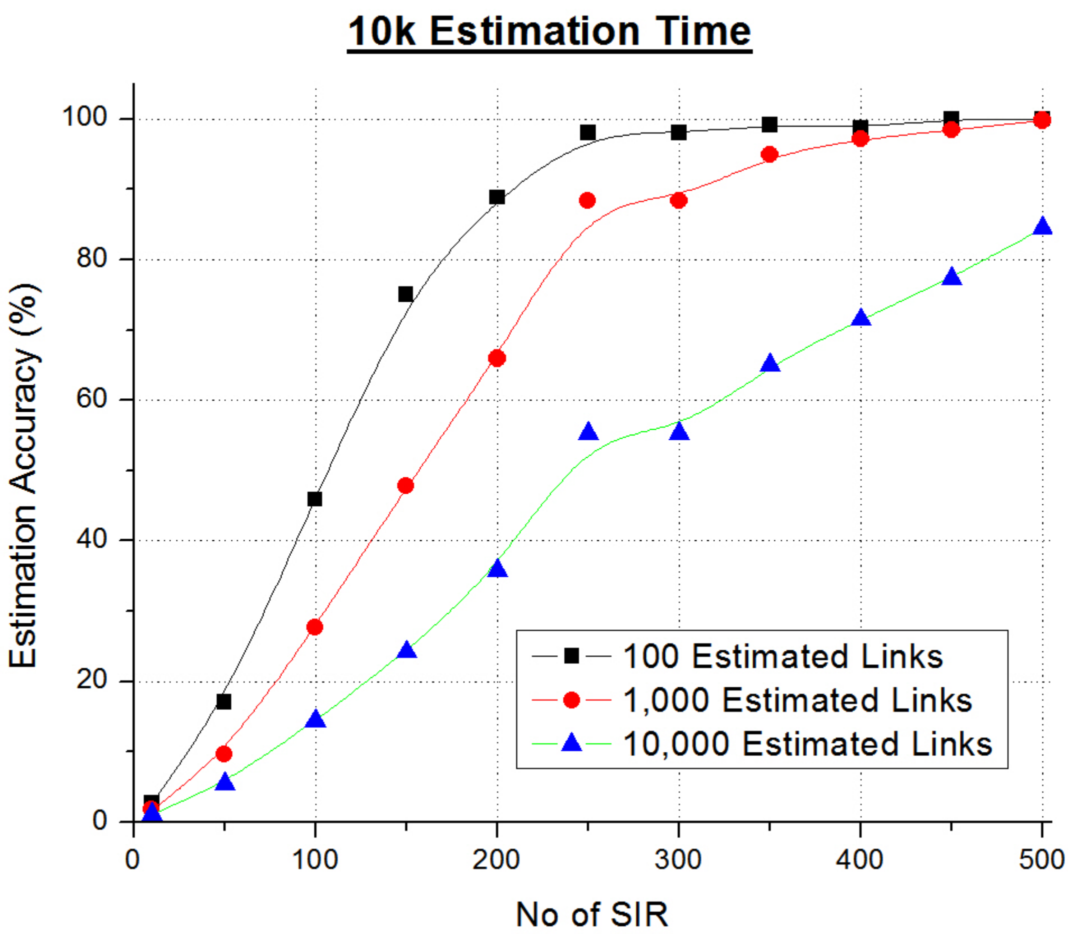
> 90%, all false links?

Number of Epidemics

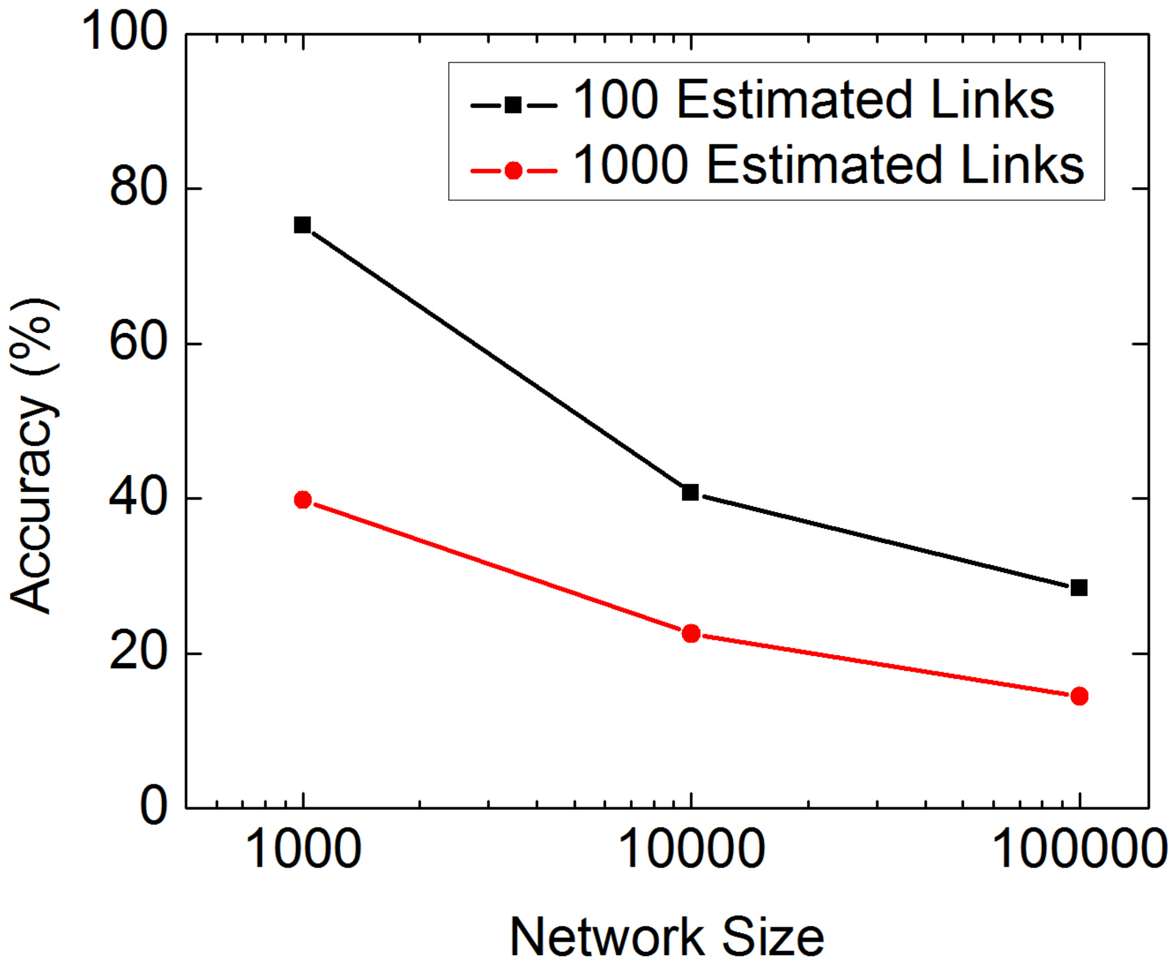
N = 10k

<k> = 10

c = 0.05



Network Sizes



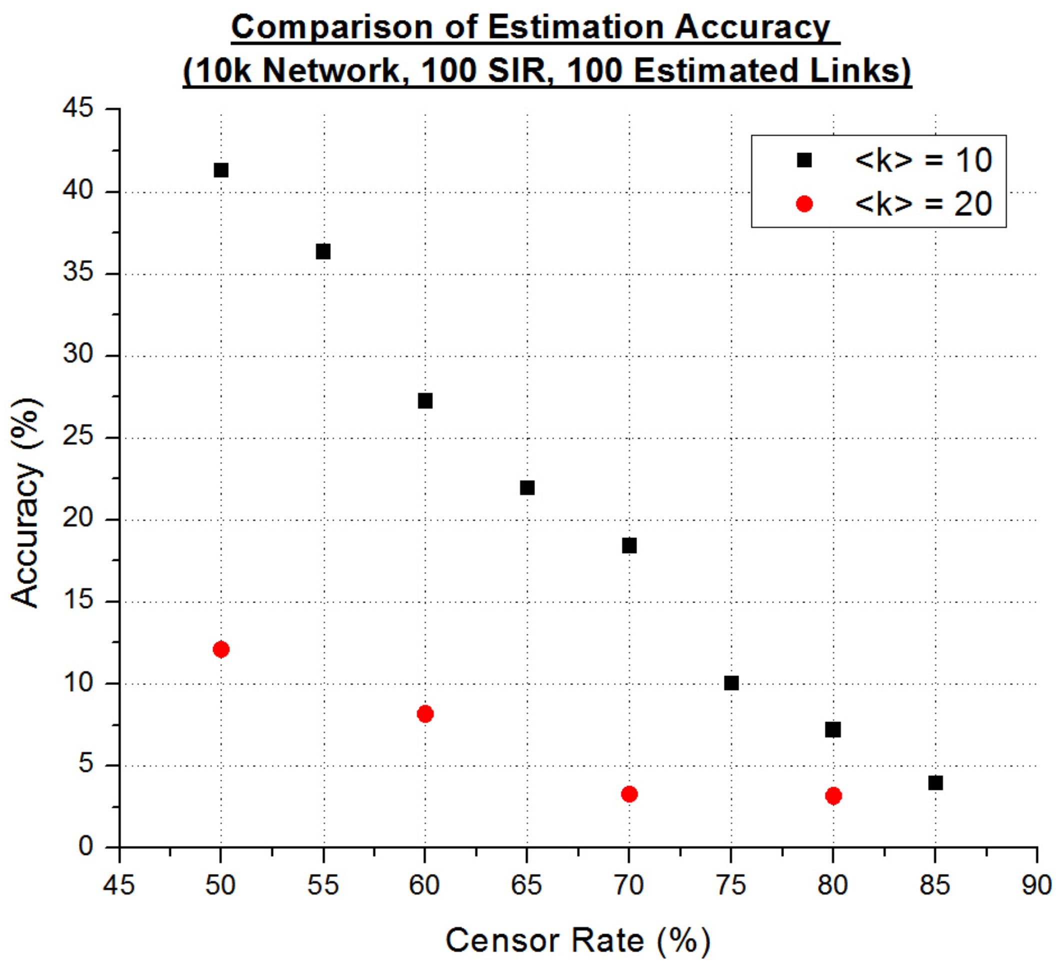
Average Degree

N = 10k

$<k>$ = 10, 20

c = 0.05

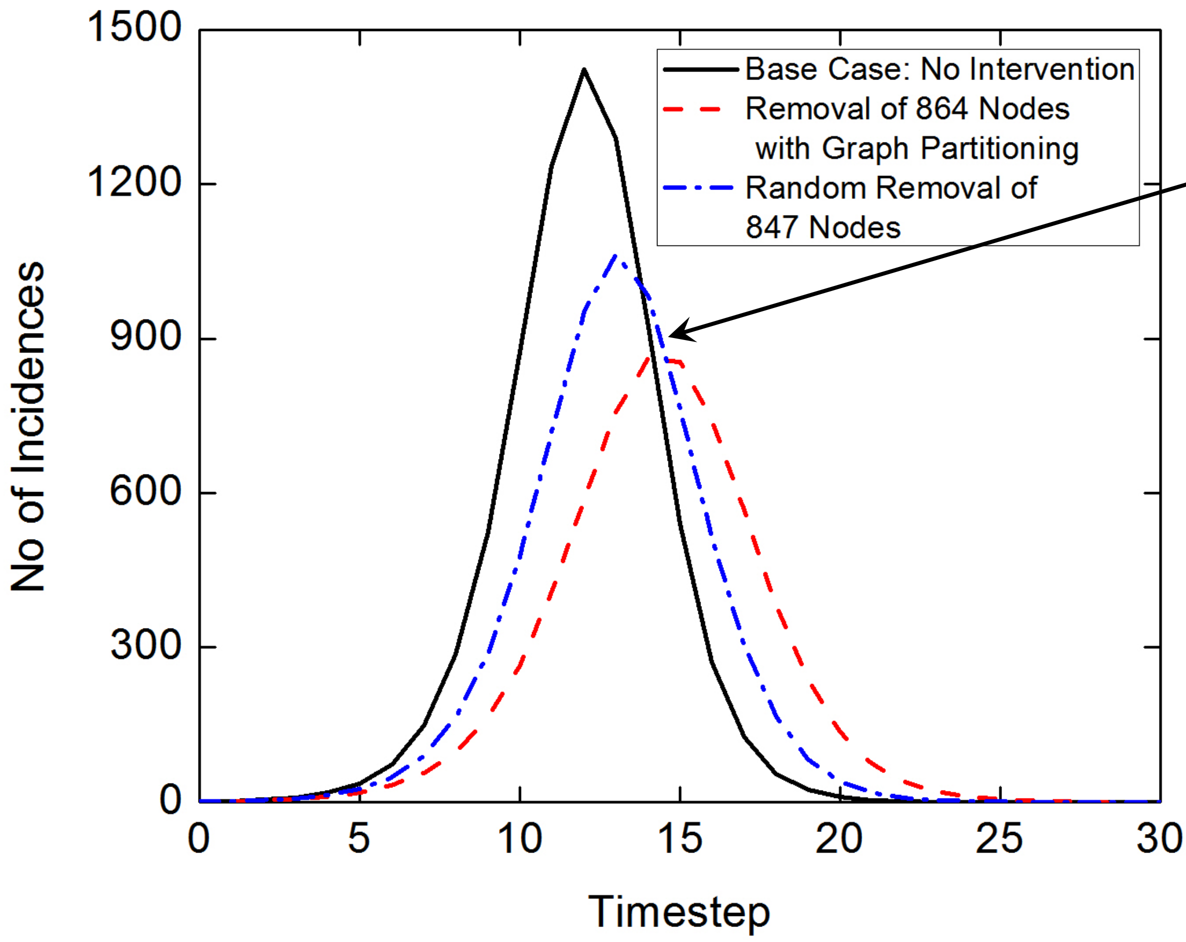
S = 100



Depressing?

- Low accuracy for high censor rate
- Every order of magnitude increase in *N*
  - Accuracy halves
- Doubling of <*k*>
  - Accuracy quarters
- How to reliably estimate $N = 10^6$ network?
  - Encourage self reporting through online portal
  - <*k*> finite even in large cities
  - Combine information from cellular phone collocation

Pre-Epidemic EGP Intervention

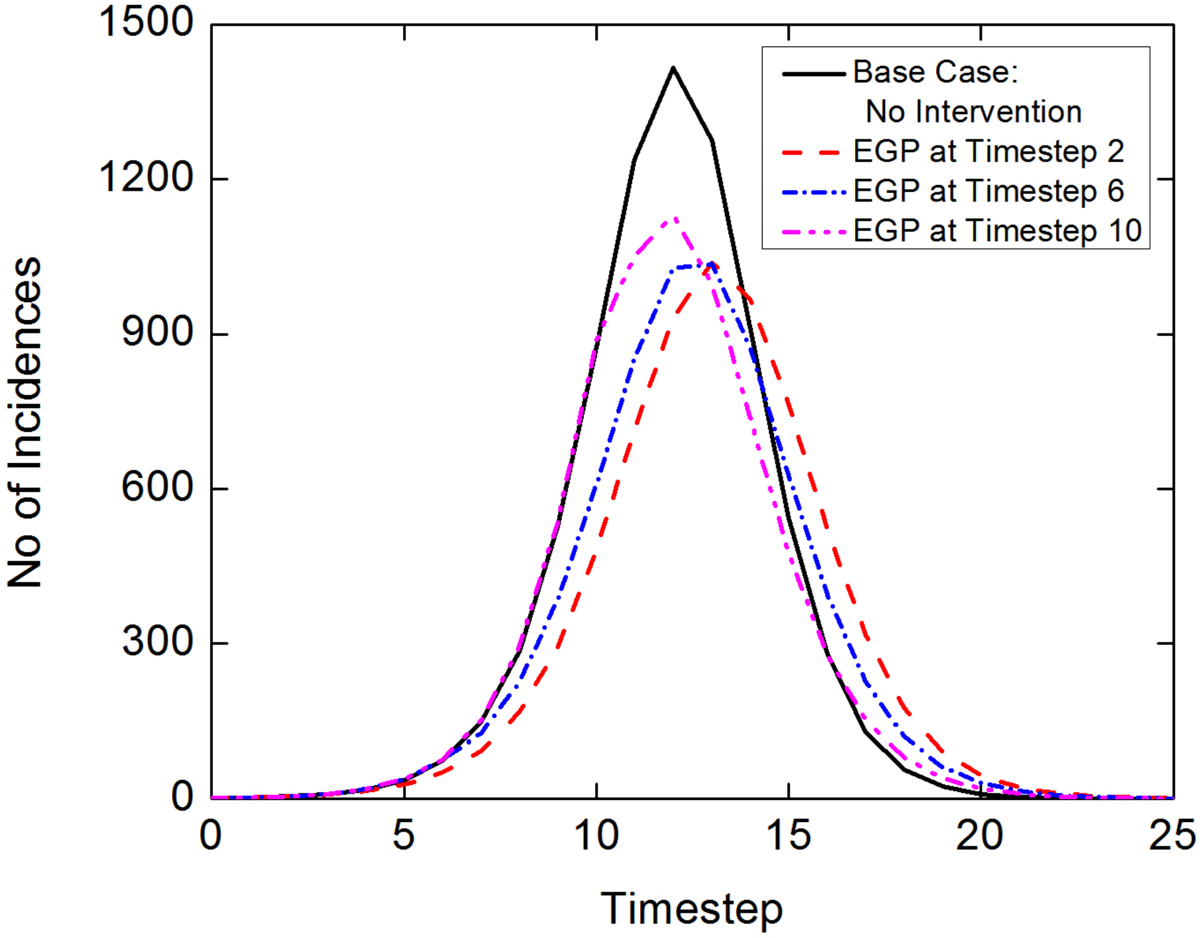


surprisingly effective!

16.6% of 10,000 estimated links correct

estimated links concentrated along backbones of infection network

In-Progress EGP Intervention

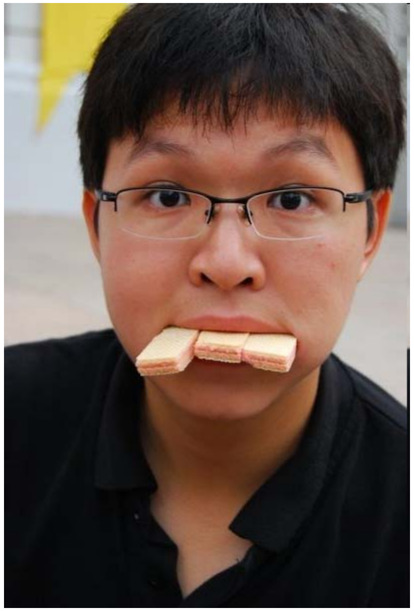


Conclusions

- Estimation of infection network through censored incidence data alone
  - SIR epidemics on artificial social network
  - Decay-reinforcement link estimation
- Infection network partially estimated
  - Asymptotic perfect accuracy possible
  - Accuracy decreases with increasing *N* and <*k*>
- EGP intervention surprisingly effective
  - Concentration of estimated links along infection backbone
  - Effective even when applied late into epidemic

Acknowledgments

- Jeremy HADIDJOJO
- Undergraduate Research Experience on Campus (URECA) Programme



Thank You!