Network-Centric Interventions to Contain the Syphilis Epidemic in San Francisco

Caterina Scoglio

Electrical and Computer Engineering
Kansas State University

Workshop Advances on Epidemics, Delft, August 31 – September 1

Co-authors

- **David Juher** and **Joan Saldana**, Departament d'Informàtica, Matemàtica Aplicada i Estadística, Universitat de Girona 17071 - Girona, Catalunya (Spain)

- **Robert Kohn**, San Francisco Department of Public Health, USA

- **Kyle Bernstein**, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

D Juher, J Saldaña, R Kohn, K Bernstein, C Scoglio
Network-Centric Interventions to Contain the Syphilis Epidemic in San Francisco
Scientific Reports 7, 2017

The problem: syphilis

- The number of reported *early syphilis* cases in USA has increased steadily during the past several years.
- It is not yet clear what factors are responsible for such an increase.
- The members of the men-who-have-sex-with-men (MSM) community have a significantly higher chance of syphilis and HIV compared to non-member individuals.

Current mitigation strategy

- *Partner notification* definition: all partners of an infected individual (neighbors in the sexual network) are notified of the infectious state

- Partner notification goals:
    1) treating exposed partners,
    2) preventing infection of healthy partners,
    3) preventing reinfection of the index cases.

- Partner notification performed by Public Health Departments allows the creation of an approximated temporal contact networks (neighbors + duration of the relationship)

Data from SFDPH

Client ID	Partner ID	First Experience	Frequency	Last Experience	Open Date	Critical Period Start	Critical Period End
478214	482624	6/1/2009	STEADY	7/5/2014	7/8/2014	6/26/2013	6/26/2014

↓

Client ID	Partner ID	Critical Period Start	Critical Period End
478214	482624	6/26/2013	6/26/2014

Contact sequence $S_0=\{(i, j, t_i, t_f)\}$

Goal of this study

- To define and evaluate new mitigation interventions for the syphilis epidemic exploiting the network data collected by the SFDPH
- New interventions must be simultaneously effective and require limited additional costs
  - Definition based on the network structure
  - Evaluation based on extensive simulation
- The SAIS model seemed to fit perfectly

Static weighted sexual network

Aggregated static contact network by assuming that the probability of a contact between two nodes per unit of time is proportional to the duration of the partnership (length of the critical contact period in our data).

$$\begin{bmatrix} i & j & t_{ij}^{init} & t_{ij}^{fin} \end{bmatrix}$$

$$\omega_{ij} = \frac{t_{ij}^{fin} - t_{ij}^{init}}{\Delta t} \quad \text{with} \quad \Delta t = \max_{ij}(t_{ij}^{fin}) - \min_{ij}(t_{ij}^{init})$$

For the largest connected component, *A* is a 953x953 symmetric matrix with 2022 non-zero elements, Δt = 1099 days, and the resulting mean link weight $\overline{\omega} = 0.1766$, with a maximum weight of 0.6142 and a minimum weight of 0.0237.

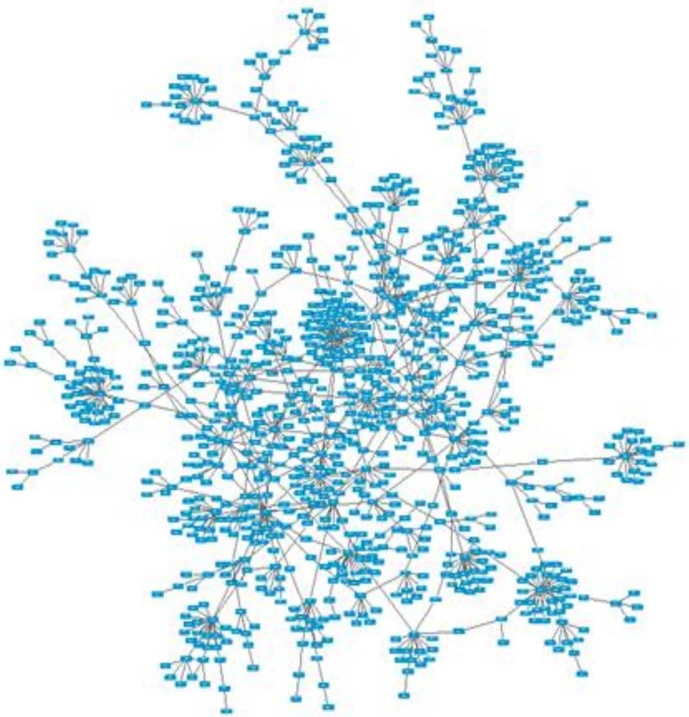
Starnini, M., Baronchelli, A., Barrat, A. & Pastor-Satorras, R. Random walks on temporal networks. *Physical Review E* **85,** 056115 (2012).

The sexual network

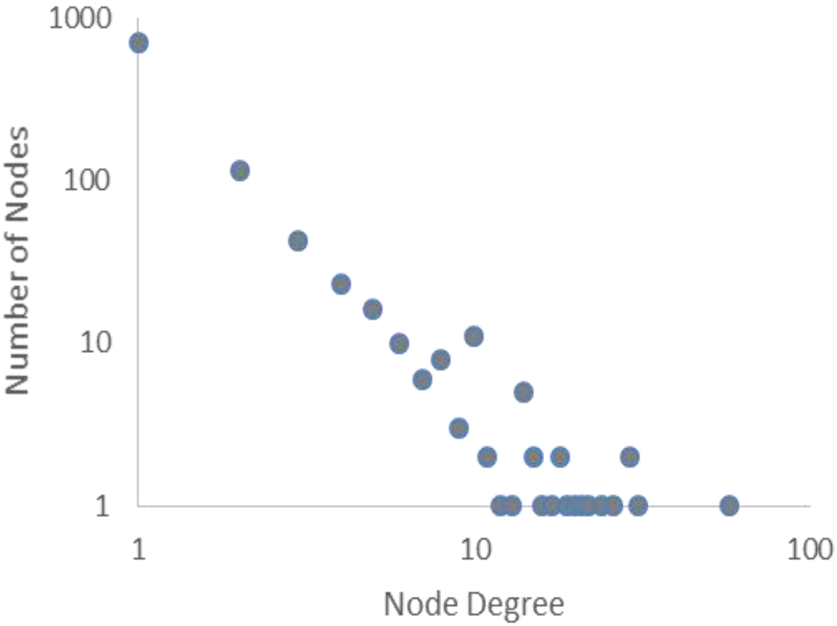
	Complete network	Sexual network largest component
Number of nodes	2428	953
Number of edges	2046	1011
Density	6.94E-4	0.002
Average Node Degree	1.68	2.12
Node Degree Variance	6.43	13.27
Clustering coefficient	0.014	0.017
Connected components	457	1

The complete network includes 457 disconnected components, where 280 components have only two nodes (61.3%), 75 components have 3 nodes, 56 components have a number of nodes included in the range 5 to 9. Only eleven components include more than 9 nodes, and one largest component includes 953 nodes.

Sexual network–connected component

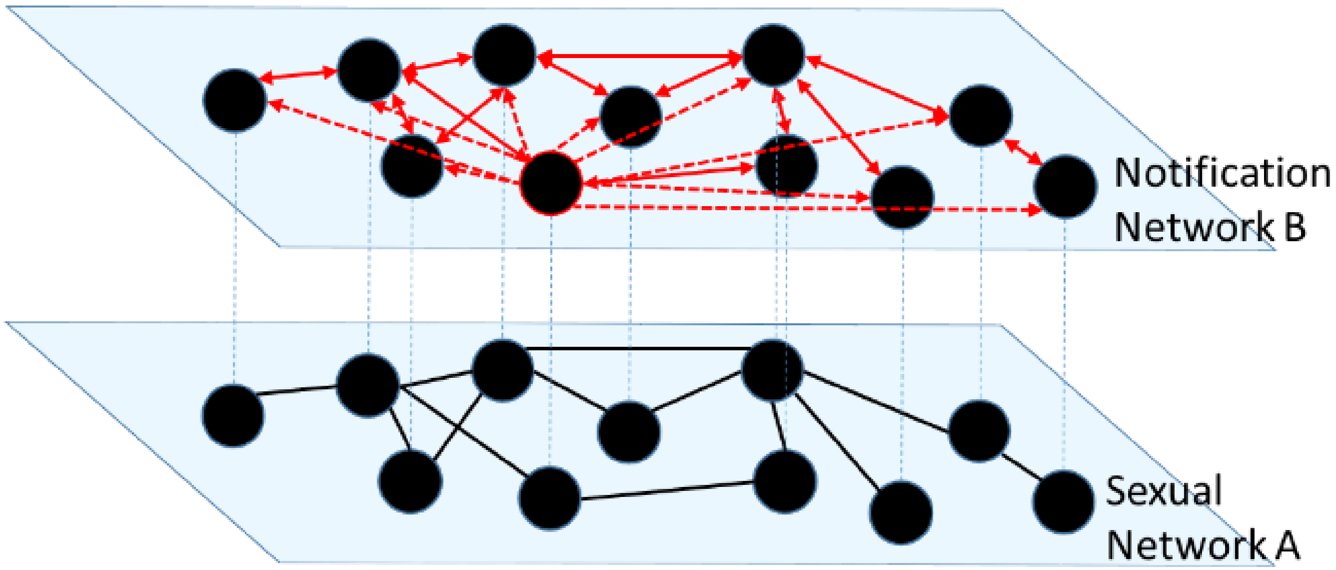


One node has degree 58, four nodes have a degree in the range 11 to 13, and eleven nodes have degree 10 . On the other side, 808 nodes have degree one and two.



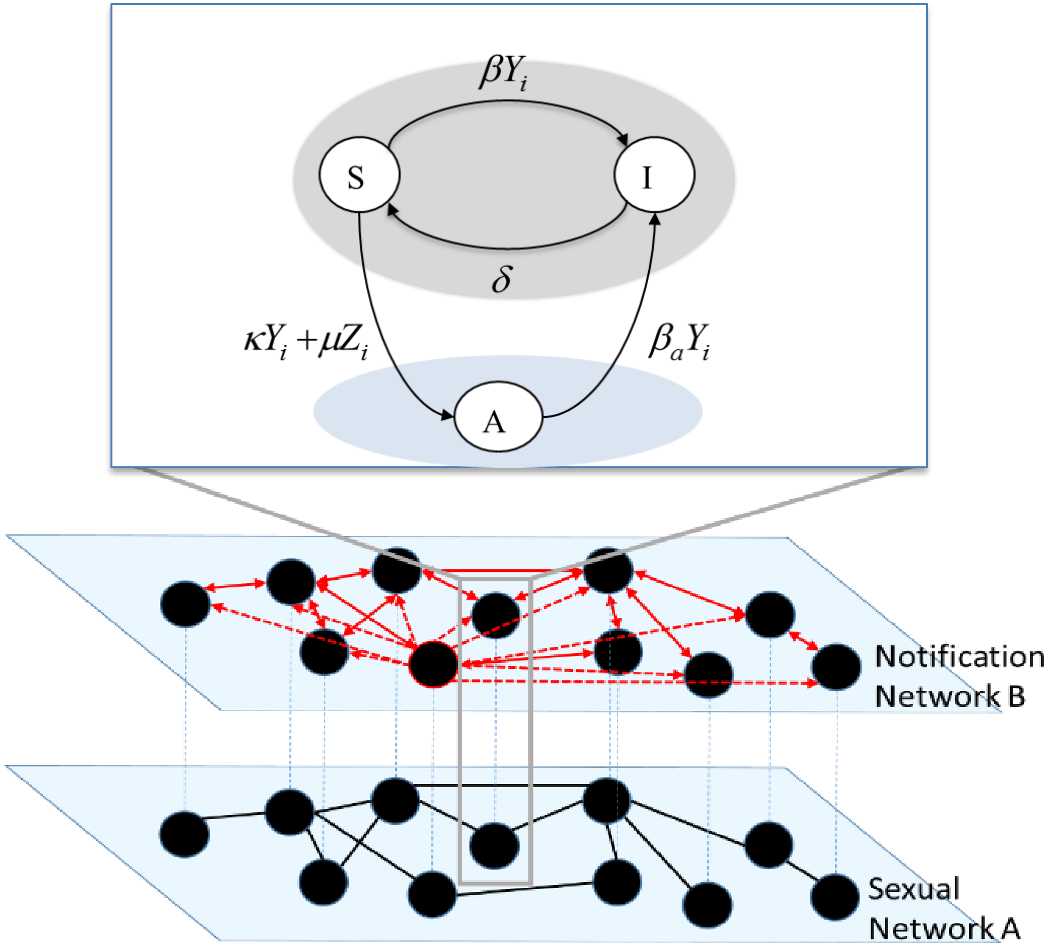
Largest connected component of the sexual contact network (left) and its node degree distribution (right). Network weighted by the duration of partnership.

The notification network



- Through the notification network, the infection status of nodes can be notified to neighbors.
- In the standard partner notification strategy, sexual network = notification network.
- What about different structures of the notification network?

The multilayer SAIS model



Transition diagram for each node in the network among the three states: susceptible (S), alert (A), and infected (I), when the transition to the alert state can be induced by infected neighbors in the notification network.

F. D. Sahneh and C. M. Scoglio, "Optimal information dissemination in epidemic networks," 2012 *IEEE 51st IEEE Conference on Decision and Control (CDC)*, Maui, HI, 2012, pp. 1657-1662.

MF Differential Equations

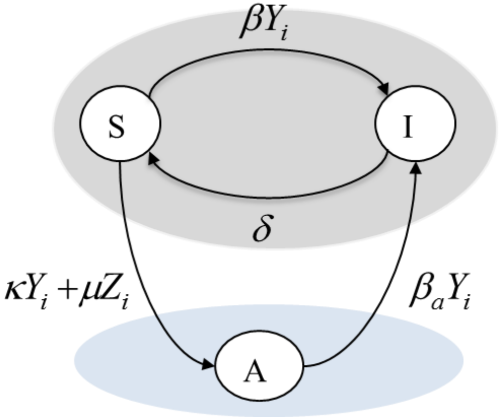
- Infection probability

$$\dot{p}_i = \beta(1 - p_i - q_i)\sum_{j \in N_i} a_{ij} p_j + \beta_a q_i \sum_{j \in N_i} a_{ij} p_j - \delta p_i , \quad i \in \{1, \ldots, N\}$$

- Alert probability

$$\dot{q}_i = (1 - p_i - q_i)\left\{\kappa \sum_{j \in N_i} a_{ij} p_j + \mu \sum_{j \in N_i} b_{ij} p_j\right\}$$

$$-\beta_a q_i \sum_{j \in N_i} a_{ij} p_j , \quad i \in \{1, \ldots, N\}$$



The Two Thresholds

- Thresholds in the SAIS model

$$\tau_a = \frac{\beta_a}{\delta}$$

$$\tau_{c1} = \frac{1}{\lambda_1(A)} \qquad \tau_{c2} = \tau_{c1} + \frac{\kappa}{\beta_a}(\tau_{c1} - \tau_a)$$

- Thresholds in the SAIS model with information propagation

$$\tau_{c1} = \frac{1}{\lambda_1(A)} \qquad \tau_{c2} = \tau_{c1} + \left(\frac{\kappa}{\beta_a} + \boxed{\varphi(A,B)}\frac{\mu}{\beta_a}\right)(\tau_{c1} - \tau_a)$$

Information Dissemination Index

- $\varphi(A, B)$ index to measure the efficacy of the information dissemination network *B* with respect to the contact network *A*

$$\varphi(A, B) = \frac{v_1^T B v_1}{\lambda_1(A)}$$

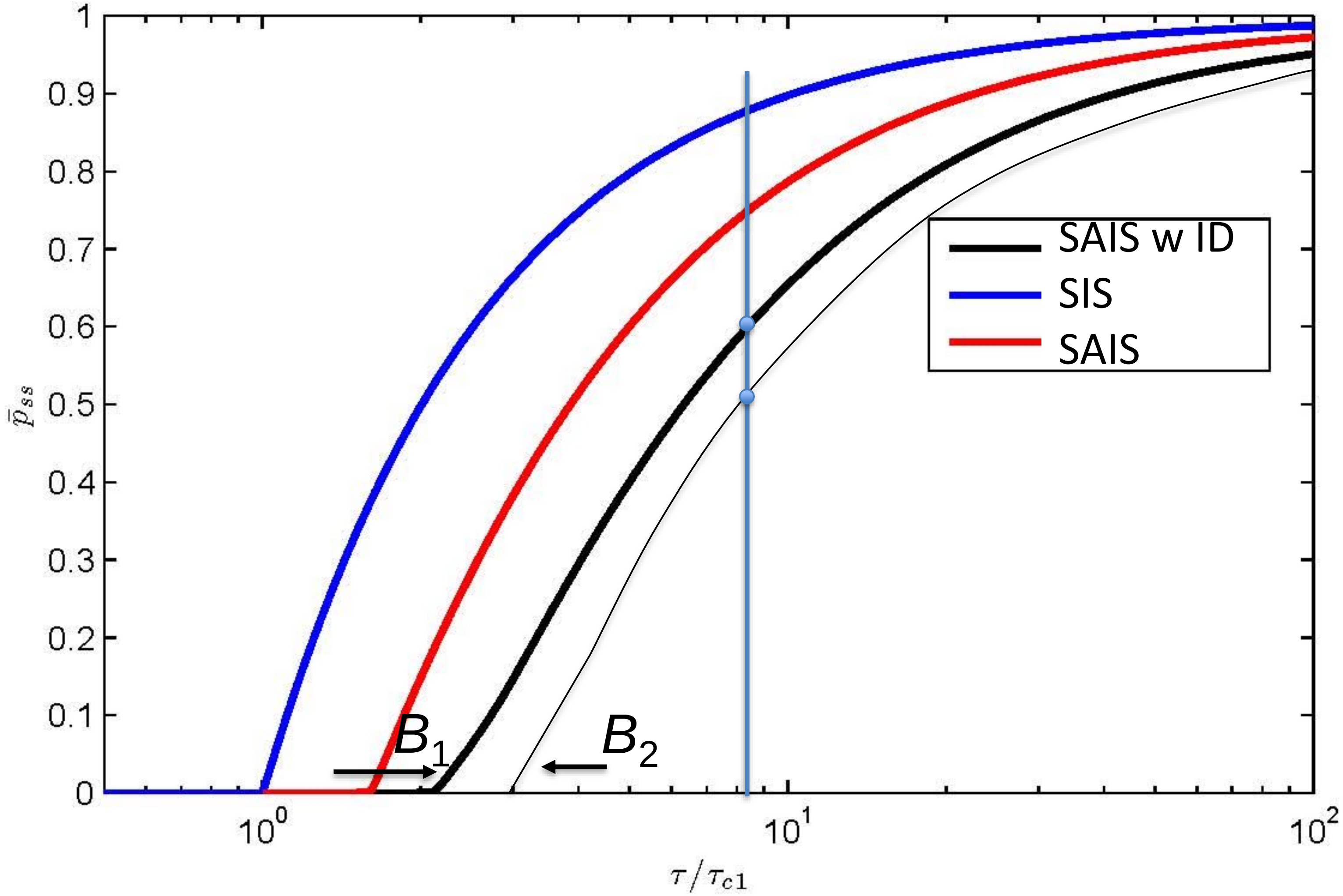
- where v_1 is the eigenvector of network A corresponding to the largest eigenvalue $\lambda_1(A)$.
- $\varphi(A, B)$ *can be maximized designing properly B!*

Maximizing the Second Threshold

- If there are no constraints on B, the second threshold is maximized when B is the fully connected graph.
- Given the number C of nodes which can broadcast their infection state, the optimal solution is to select the C nodes that have the largest eigenvalue centrality in the contact network *A*.

$$\text{Maximizing} \quad \varphi(A,B) = \frac{v_1^T B v_1}{\lambda_1(A)}$$

Epidemic Size Curves



Community notification strategies

- *Community notification* has the goal of alerting all nodes in the network when the index case belongs to a small subset of *core* nodes in the network
- Only a generic alert message will be sent to the community with the purpose of inviting everyone to adopt effective preventing measures and to look for testing and treatment if needed

Simulations

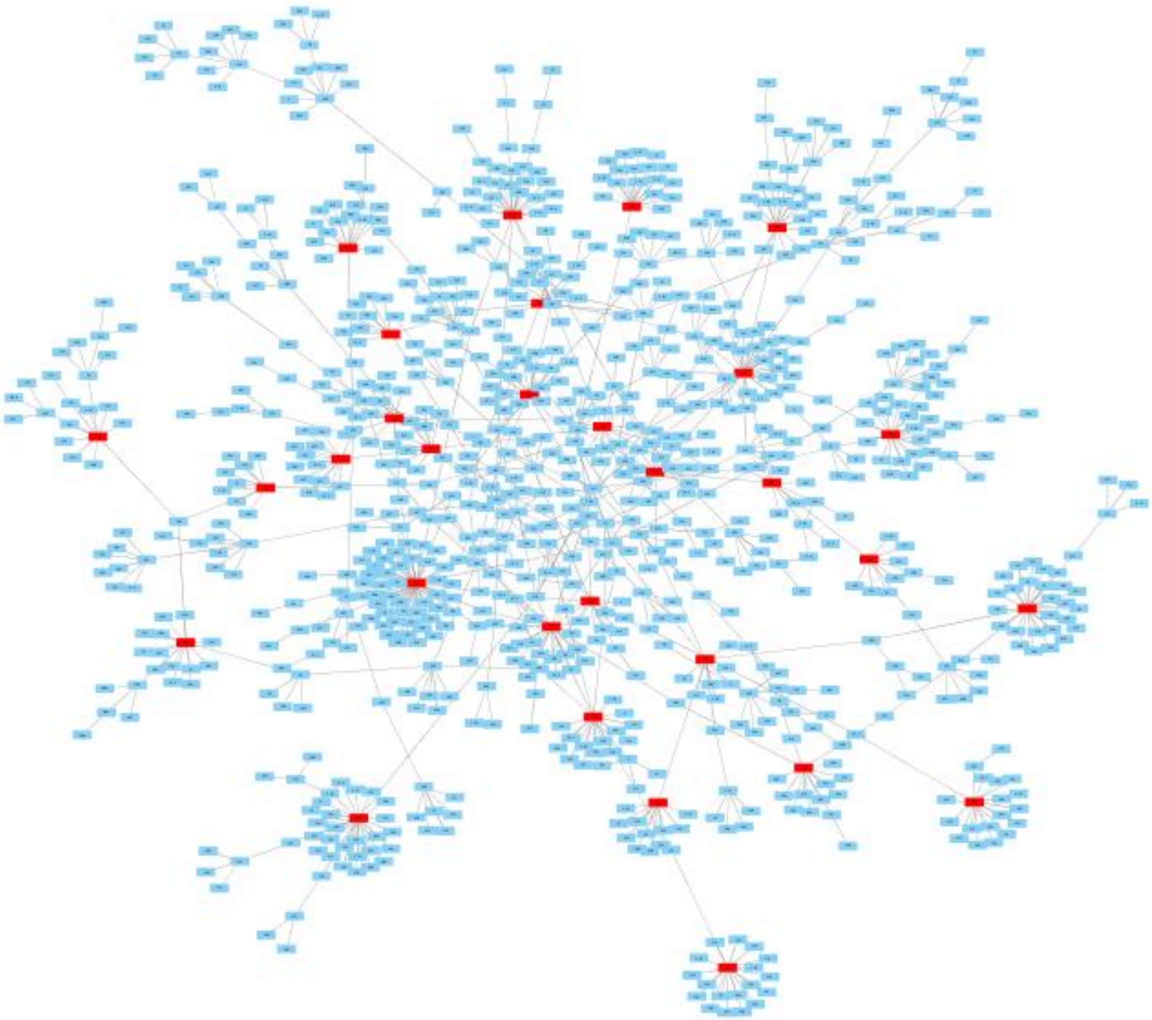
- Gillespie-based algorithm
- Each simulation has been repeated 350 times selecting different sets of initially infected nodes (5%)
- With the selected parameter values, the simulations converge in the long term to a non-zero number of infected individual

Parameter	β	β_a	κ	δ
(1/day)	0.94	0.47	0.02	0.30

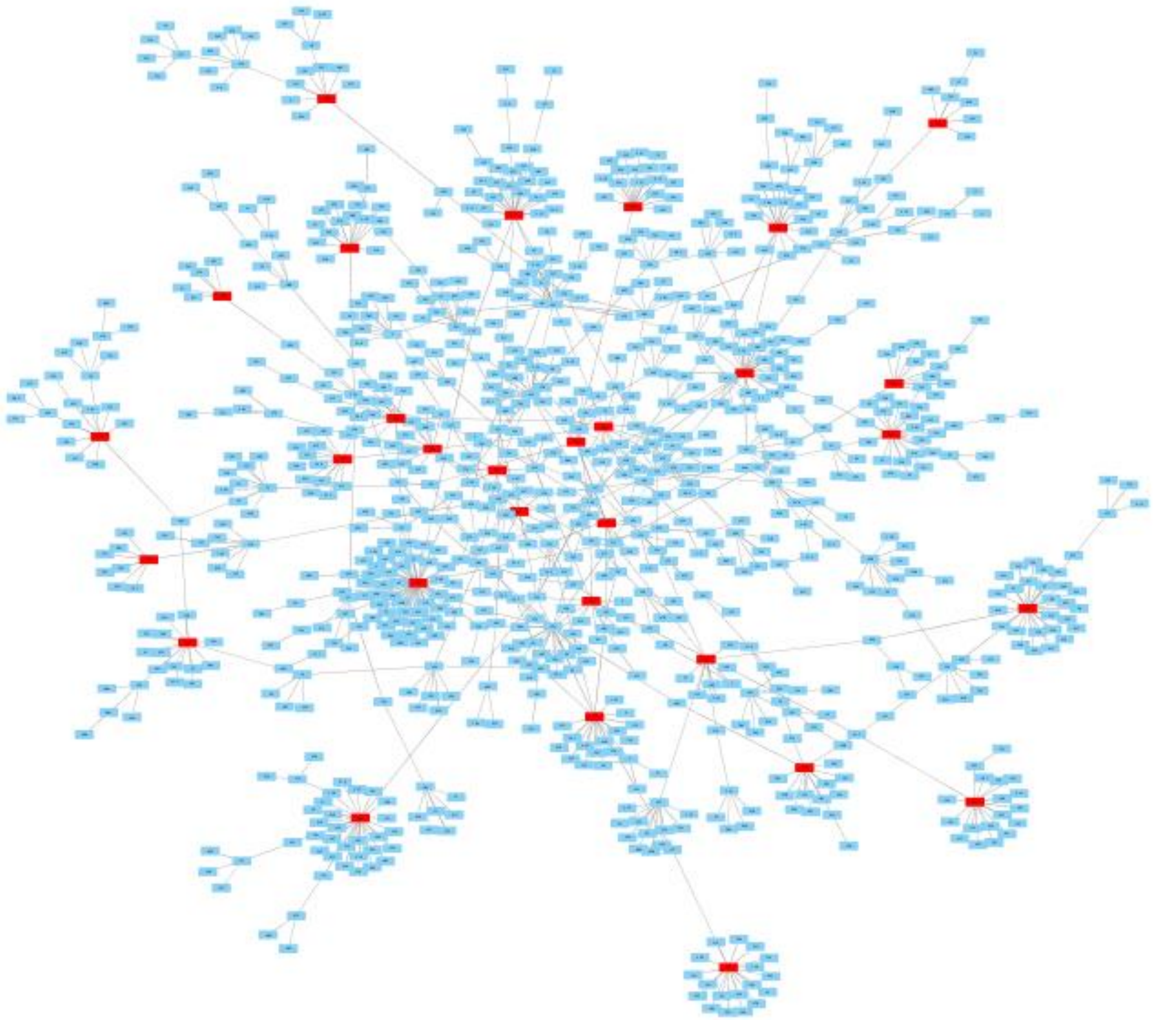
Network centric node selections

- Amount
  - (5, 14, 30); more than 30 nodes → marginal improvement
- Centrality measure
  - Degree centrality (DC), Weighted degree centrality (WDC), Eigenvector centrality (EC).
- Node selections define the notification network

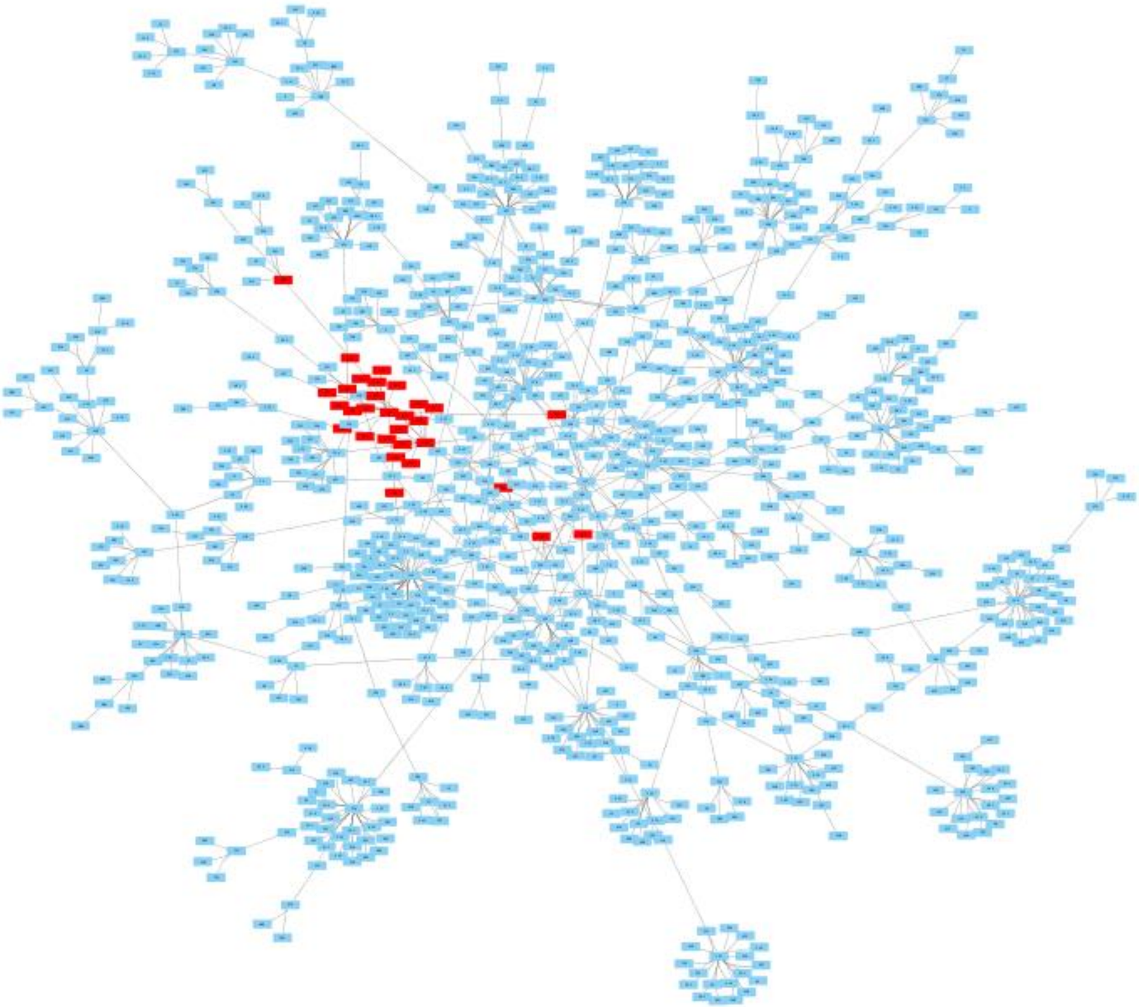
Degree centrality: top 30



Weighted Degree centrality: top 30

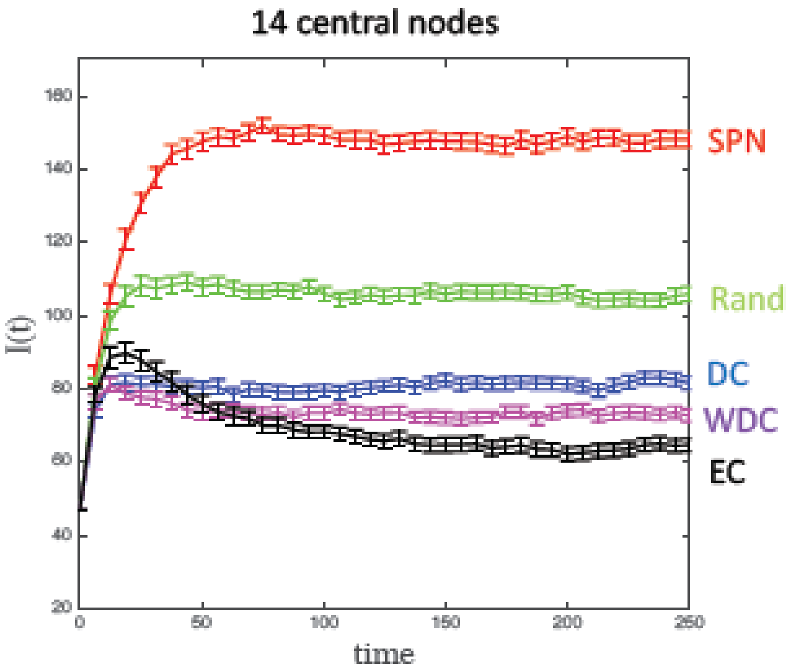


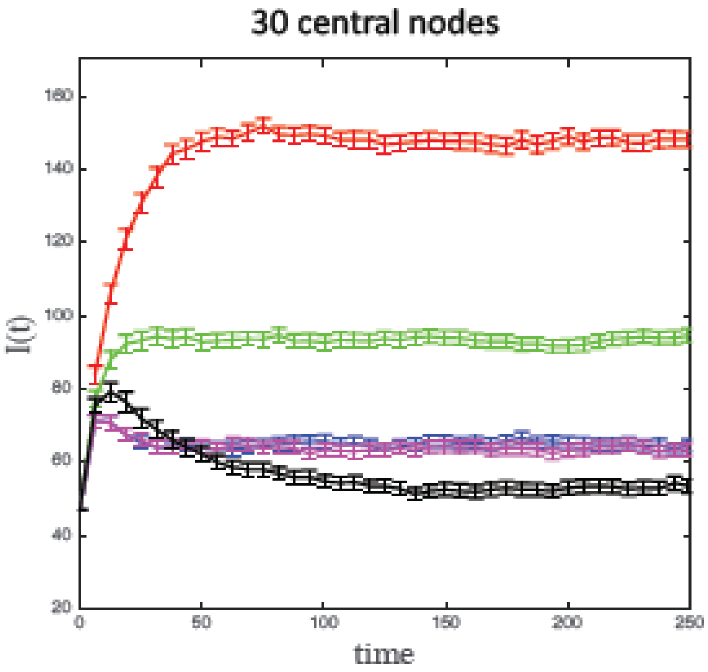
Eigenvector centrality: top 30



Epidemic Dynamics

Epidemic dynamics under 4 selection strategies


14 central nodes


30 central nodes

Comparison

TABLE V	*C*=0	*C*=5				*C*=14				*C*=30			
	SPN	Rand	DC	WDC	EC	Rand	DC	WDC	EC	Rand	DC	WDC	EC
Additional links	0	4754	4587	4622	4739	13312	12980	13013	13285	28491	28015	28056	28483
Average Number of infected individuals	148	123	100	91	87	105	81	72	64	94	64	63	53

Sensitivity to link weights

- We perform two types of numerical experiments:

1. *Homogeneously weighted network*:

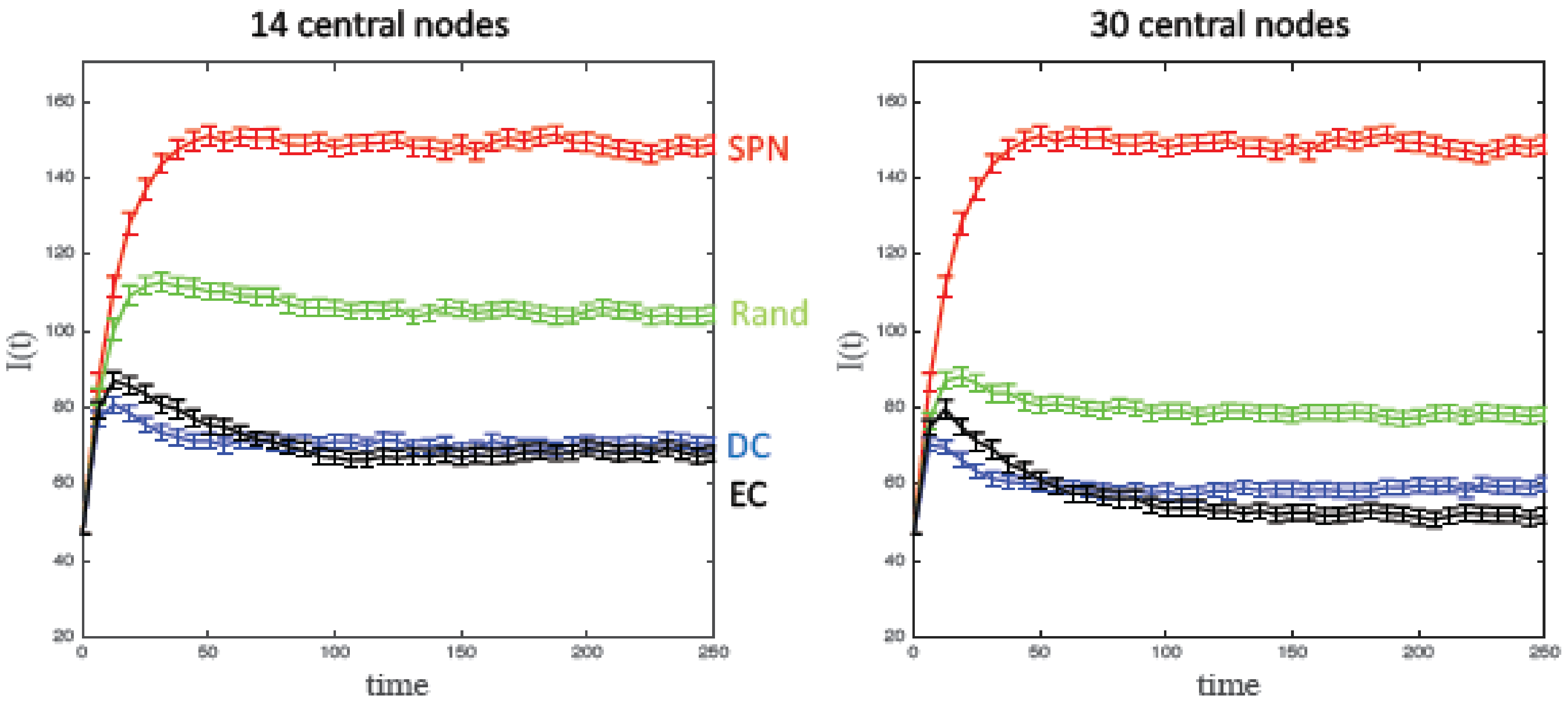
   link weight = mean link weight of the original network

   = $\bar{\omega} = 0.1766$

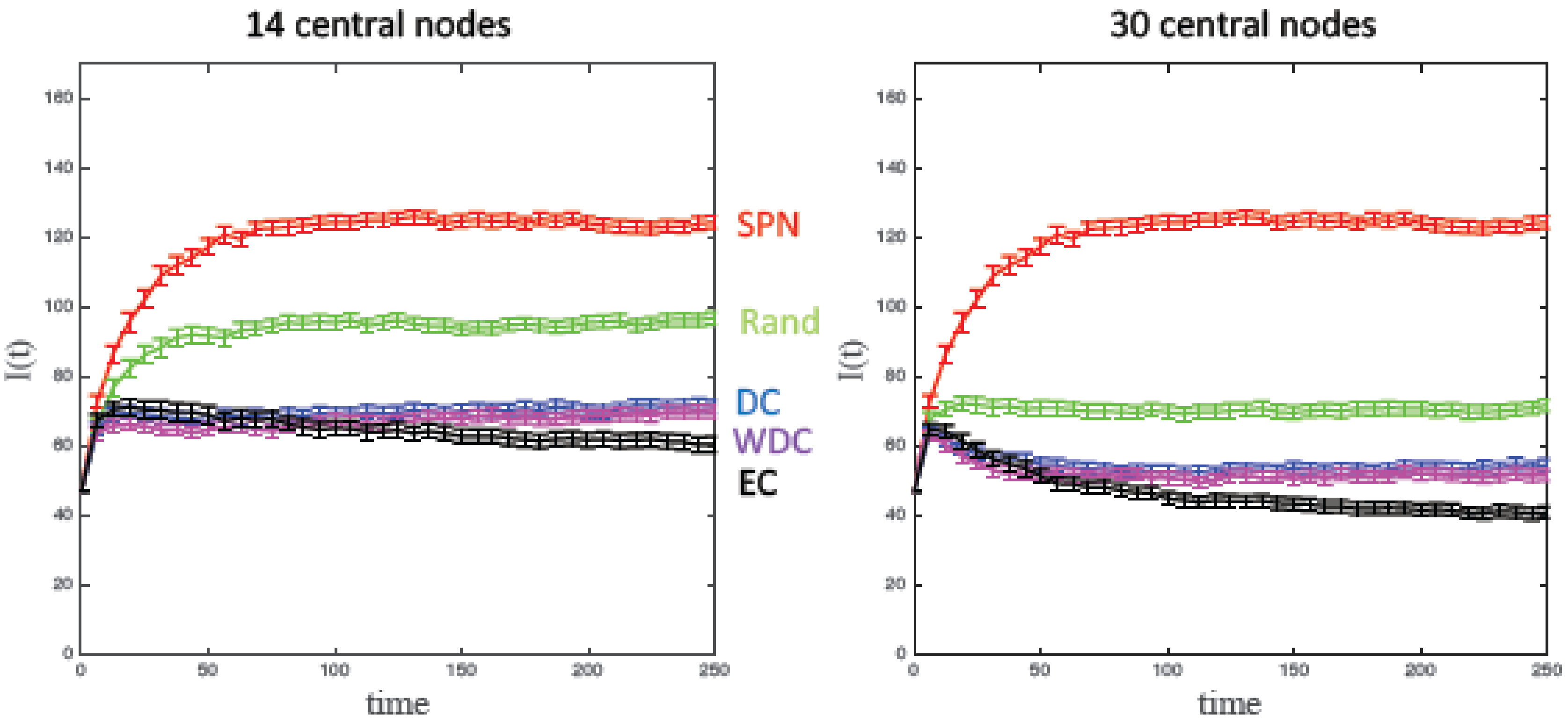
2. Random assignment of links weights ~ *exponential distribution* with expected value = $\bar{\omega}$

   A weight = 1 is assigned when the generated weight is > 1 (to preserve the interpretation of ω_{ij} as a probability)

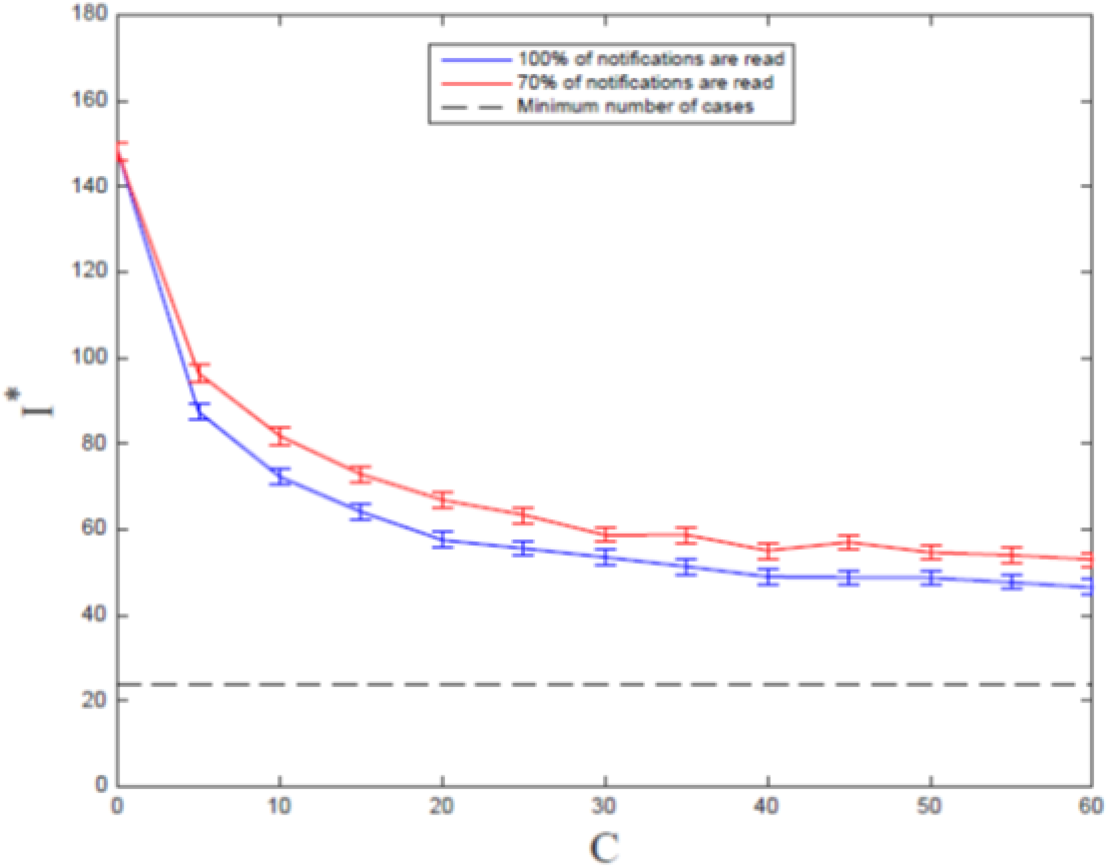
Homogeneously weighted network



Random weighted network



Sensitivity to Response



THE MEAN NUMBER OF INFECTED INDIVIDUALS AT EQUILIBRIUM AS A FUNCTION OF THE NUMBER OF SELECTED NODES BASED ON EIGENVECTOR CENTRALITY (EC) WITH 100% RESPONSE (BLUE LINE) AND 70% RESPONSE (RED LINE).

Sensitivity to Missing Links

- Random addition of 200, 300, 500, and 1000 links

- Two methods of attachment of these new links:
  - Both ends of a link are attached to nodes chosen **uniformly at random**
  - One of the two ends of the new links is **preferentially attached to nodes of high degree**: attaching prob. ~ $k^{0.7}$ (the other end is attached at random)

In the long run:

- EC remains the best selection strategy under all scenarios except when 1000 links are added under semi-preferential attachment where WDC performs better than EC

Current and future work

- How to improve the quality of noisy data and infer the network topology
  - Identification of missing links using Susceptible-Infected-Susceptible traces
  - The goal is to find the probability for the existence of uncertain links in a partially known network using the trace of SIS epidemic process
  - $Likelihood\ (trace\ |\ transmission\ rates)$, perform Bayesian inference
  - A. Vajdi, C. Scoglio
    Identification of Missing Links Using SIS Epidemic Traces
    arXiv preprint; arXiv:1708.01631 [physics.soc-ph]

Concluding remarks

- Preliminary study: many additional aspects of the modeling and implementation still require further assessment and development.

- If the notification is performed with a *mobile phone app*, minimizing the number of links is not a critical goal, and strategies like “broadcast the alert message if any of the 30 nodes with highest eigenvector centrality is infected” seems to be promising.

Acknowledgments

- This work was supported by the National Science Foundation Award CIF-1423411 and Award DMS 1515810.