Network Reconstruction from Viral State Observations and Prediction of the Epidemic Nodal State

Bastian Prasse & Piet Van Mieghem

Mathematical Modeling and Statistical Analysis of Infectious Disease Outbreaks Marseile (CIRM), 17-21 February, 2020



1

Expected Cum. Fraction of Covid-19 infected in



B. Prasse, M. A. Achterberg, L. Ma and P. Van Mieghem, "Network-Based Prediction of the 2019-nCoV Epidemic Outbreak in the Chinese Province Hubei", arXiv: 2002.04482 (12 Febr. 2020)



Expected Cum. Fraction of Covid-19 infected in Hubei



Computed yesterday; not in arXiv: 2002.04482



Outline



Epidemic model

Network reconstruction

Reconstruction and prediction accuracy



Compartmental Epidemic Models

- Simplest epidemic model: every individual is either *Infected (I)* curing or *Susceptible (S)* SIS epidemic S model: infection
- More compartments may be more accurate: •

E.g., *Exposed (E)* but not contagious and *Recovered (R)*
SEIR epidemic
model:
$$S$$
 infection E incubation I curing R

- Almost all compartmental epidemic models are special cases of **GEMF** (Generalized Epidemic Mean-Field)
- Here: focus on SIS model, but results apply to all GEMF models

Sahneh, F. D., C. Scoglio and P. Van Mieghem, "Generalized Epidemic Mean-Field Model for Spreading Processes over Multi-Layer Complex Networks", IEEE/ACM Transactions on Networking, Vol. 21, No. 5, pp. 1609-1620, 2013. TUDelft



Group-Based SIS Epidemic Model

Epidemic spread between *N* groups of individuals



At any *discrete* time k = 1, 2, ..., every group i = 1, ..., N has a viral state vector $v_i[k] = (S_i[k], I_i[k])^T$

- *S_i*[*k*]: fraction of Susceptible individuals in group *i*
- $I_i[k]$: fraction of Infected individuals in group *i*

 $\succ I_i[k] + S_i[k] = 1$ for every group *i* at every time *k*

B. Prasse and P. Van Mieghem, "Network Reconstruction and Prediction of Epidemic Outbreaks for NIMFA Processes", arXiv:1811.06741, November 2018.



SIS Governing Mean-field Equations

Nonlinear difference equation of infection probability I_i[k] of node *i*:

$$I_{i}[k+1] = \underbrace{(1-\delta_{i})I_{i}[k]}_{\text{curing}} + S_{i}[k] \sum_{j=1}^{N} \beta_{ij}I_{j}[k]$$

infections

- $\circ \delta_i$: Curing probability of group *i*
- o β_{ij} : Infection probability from group *j* to group *i*
- Group interactions are specified by the *infection probability* $(\beta_{11}, \dots, \beta_{1N})$

$$B = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1N} \\ \vdots & \ddots & \vdots \\ \beta_{N1} & \cdots & \beta_{NN} \end{pmatrix}$$

- *Problem* : matrix *B* unknown for real epidemics
- *Solution attempt* : estimate *B* from observing the epidemic

Prasse, B. and P. Van Mieghem, 2019, "The Viral State Dynamics of the Discrete-Time NIMFA Epidemic Model", IEEE Transactions on Network Science and Engineering, to appear.



Outline



Epidemic model

Network reconstruction

Reconstruction and prediction accuracy



Network Reconstruction from Epidemics (1)



 $\{x_i[1], x_i[2], \dots, x_i[n]\}$

 $x_i[k] \in \{0,1\}$ is the viral SIS state of node *i* at discrete time k

 $x_i[k] = 0$: node i is healthy $x_i[k] = 1$: node i is infected

Objective: Infer the N×N adjacency matrix A of the network from viral SIS state observations of all nodes

Solution: Infeasible (NP-hard)

Prasse, B. and P. Van Mieghem, 2018, "Exact Network Reconstruction from Complete SIS Nodal State Infection Information Seems Infeasible", IEEE Transactions on Network Science and Engineering, Vol. 6, No. 4, October-December, pp. 748-759.



Network Reconstruction from Epidemics (2)



Objective: Infer the N×N adjacency matrix A of the graph from a set of viral probabilities over time of all nodes

Solution: possible when using a mean-field model

Key observation: Governing mean-field equations are linear in the (weighted) adjacency matrix *B*

B. Prasse and P. Van Mieghem, "Maximum-Likelihood Network Reconstruction for SIS Processes is NP-Hard `, arXiv:1807.08630, July 2018.



Network Inference based Prediction algorithm (NIPA)

Observations $I_i[1], ..., I_i[n]$ Network reconstruction $N \times N$ infectionfor every group i = 1, ..., Nprobability matrix B = ?

n: number of observations

Network reconstruction is *equivalent* to a linear system for every group *i*: (β_{i1})

$$v_i = F_i \begin{pmatrix} \rho_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix}$$

Vector v_i and matrix F_i follow from the viral state observations $I_i[1], ..., I_i[n]$

Obtain matrix estimate \hat{B} by solving LASSO:

$$\min_{\beta_{i1},\dots,\beta_{iN}} \left\| v_i - F_i \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix} \right\|^2 + \rho_i \sum_{j=1}^N |\beta_{ij}| \quad \text{for } i = 1, \dots, N$$

R. Tibshirani, "Regression shrinkage and selection via the Lasso," Journal of the Royal Statistical Society: Series B (Methodological), 1996



Network Reconstruction as Linear System

• Network reconstruction results in set of linear equations:

$$Fb = v$$

- Matrix F and vector v: transformations of the viral state observations $v_i[0], v_i[1], \dots$ of every node i
- Problem: Matrix *F* is extremely ill-conditioned
 - Lagrangian optimization with constraints (multipliers)

 A-priori estimates of the matrix B can be taken into account (Bayesian): e.g. mobility pattern between cities
Convexity (efficient algorithms)

basis pursuit (L1-norm regularisation)

$$\min_{B} \|B\|_1$$

s.t.
$$Fb = v$$

 \diamond good without model errors nor a-priori estimate of the graph

Prasse, B. and P. Van Mieghem, 2018, "Network Reconstruction and Prediction of Epidemic Outbreaks for NIMFA Processes", arXiv 1811.06741



Rank problem of Matrix F in Fb=v



Intuition: only the initial state agitates eigenstates; thereafter the epidemics process autonomously tends to its steady-state mainly steered by the principal eigenvector (corresponding to the largest eigenvalue)

Solution of the rank problem of matrix F:

- Multiple outbreaks/realizations of epidemic on same network
- Assume that the curing rate changes around a mean value (curing rate control)
- Singular values of F decrease exponentially. Hence, model errors have significant effect

13



Outline



Epidemic model

Network reconstruction

Reconstruction and prediction accuracy



Prediction of Epidemic Outbreaks



- True viral state $I_i[k]$: generated on the true network B
- Predicted viral state $\hat{I}_i[k]$: generated on the reconstructed network \hat{B}



Prediction Accuracy

- Barabasi-Albert network with N = 200 nodes
- Only few observations: n = 50



• Very accurate prediction at every time $k \ge 50$

> Does the prediction accuracy imply $\hat{B} \approx B$?



Network Reconstruction Accuracy

1. AUC score: AUC = 0.51

◦ Tossing a coin for reconstructing every link → AUC = 0.5

2. In-degree distribution:





Conclusion



Accurate prediction of epidemic outbreaks without accurate network reconstruction

B. Prasse and P. Van Mieghem, 2020, "Network Reconstruction and Prediction of Epidemic Outbreaks for General Group-Based Compartmental Epidemic Models", IEEE Transactions on Network Science and Engineering, submitted TUDelft



Books

Performance Analysis of Complex Networks and Systems

Piet Van Mieghem



Graph Spectra for Complex Networks

Piet Van Mieghem



Data Communications Networking

Piet Van Mieghem

19

Articles: http://www.nas.ewi.tudelft.nl



Thank You

Piet Van Mieghem NAS, TUDelft P.F.A.VanMieghem@tudelft.nl



20