Graph Interventions with Spectral Embeddings

Christian van der Loo[1,2], Rajmonda Caceres[2], Lirong Xia[1]
[1] Rensselaer Polytechnic Institute, Troy, NY
[2] MIT Lincoln Laboratory, Lexington, MA

Motivation

The COVID-19 pandemic highlighted the need for improving modeling of pandemic forecasting and control. We model spread of COVID-19 as a diffusive process over a proximity-based **interaction network**.

Nodes in the network represent “synthetic individuals” belonging to communities (such as towns), and edges or interactions in the network capture the amount of potential exposure individuals have because of proximity at various points of interest (e.g. shops) [3].

By applying interventions on this network, policies can be created that stem the spread of disease through the simulated network.

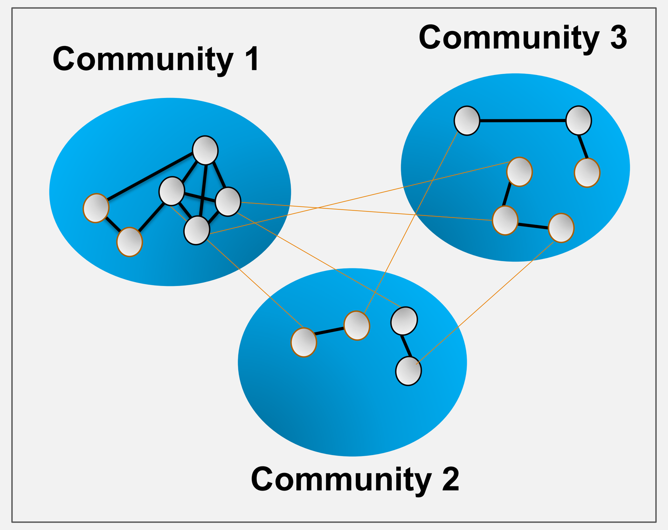

Figure 1: *Stochastic Block Model network structure*

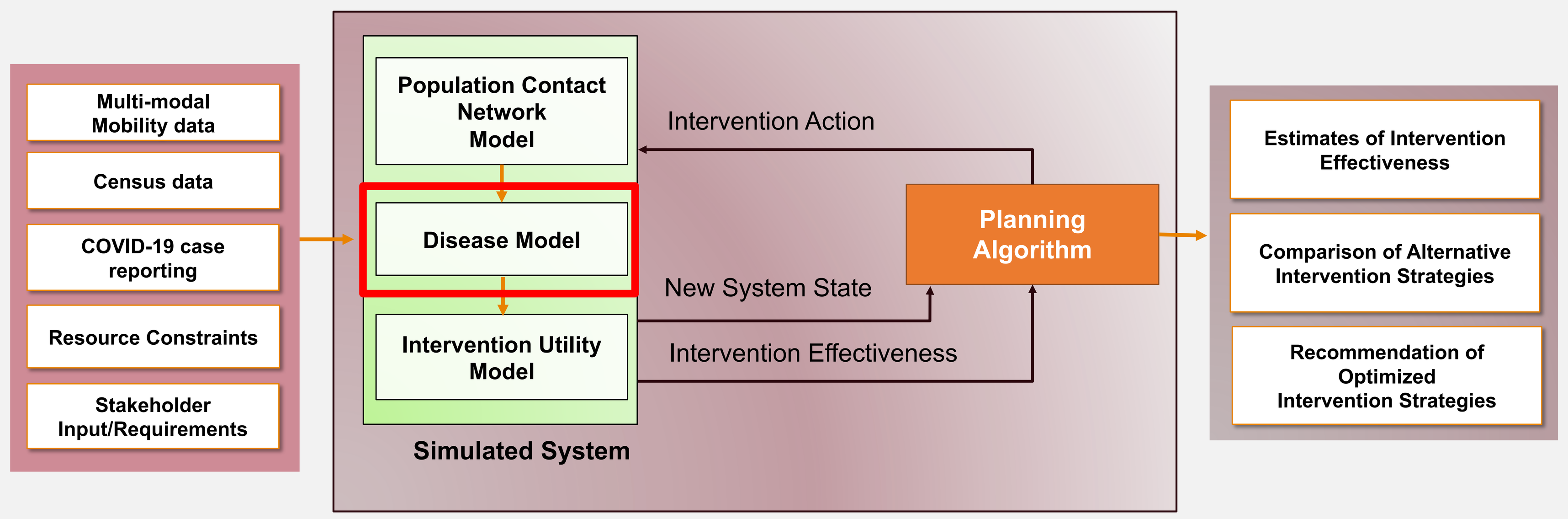

Figure 2: *End-to-End Decision System for Disease Intervention*

The most expensive part of planning is simulating the disease repeatedly, such as with an MDP approach[1]. If we could fit a model or some surrogate process to the simulation, we could gain efficiency when planning for pandemic control.

Approach

Graph spectral properties are known to correlate and reflect characteristics of graph spreading processes. These properties are derived from the adjacency and Laplacian matrices of a given graph.

> *“…for a large family of dissemination processes, the largest […] eigenvalue λ of the adjacency matrix A or an appropriately defined system matrix is the only graph parameter that determines the tipping point of the dissemination process”* [2]

The Idea: *Use the spectral properties of a graph G, and a disease state with S, E, I, and R counts, to predict disease progression (the SEIR curve) over a number of timesteps. This can act as a surrogate for an actual disease simulation.*

- Deep neural networks can capture complex relationships
- An LSTM can capture time-varying progressions

Experimental Setup

Graph Parameters:

- Graph size of 5000
- Node degree sampled from Poisson $P(X = 6)$
- 2 communities

We create stochastic block model networks in this regime, and sweep the between-community and within-community connectivity. When these two are equal, the graph is effectively an Erdős–Rényi.

Spectral Properties:

From graph G we can get adjacency matrix A and Laplacian matrix L.

- The adjacency “eigengap” is $\lambda_1 - \lambda_2$ of A. This value increases as the graph becomes more ER-like.
- The maximum adjacency eigenvalue is the largest eigenvalue of A.
- The smallest Laplacian eigenvalue is the smallest, non-zero eigenvalue of L.

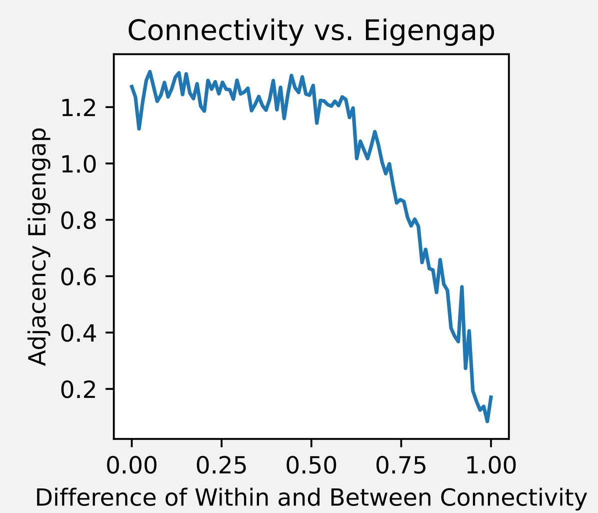

Figure 3: *Spectral properties have measurable sensitivity to community structure.*

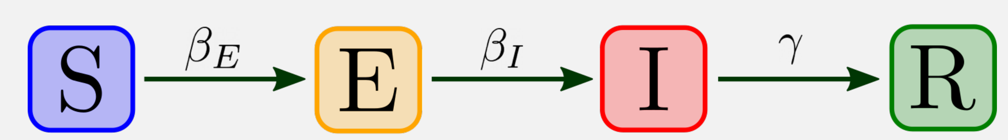

Figure 4: *The SEIR process, where the probability of state transition is represented by the Greeks above.*

SEIR was simulated over the graphs, with interventions ranging from 0% to 70%. The intervention (and spectral properties) were calculated by applying the intervention to all edge weights.

SEIR Parameters: $\beta_I = 0.11$, $\beta_E = 0.24$, $\gamma = 0.13$

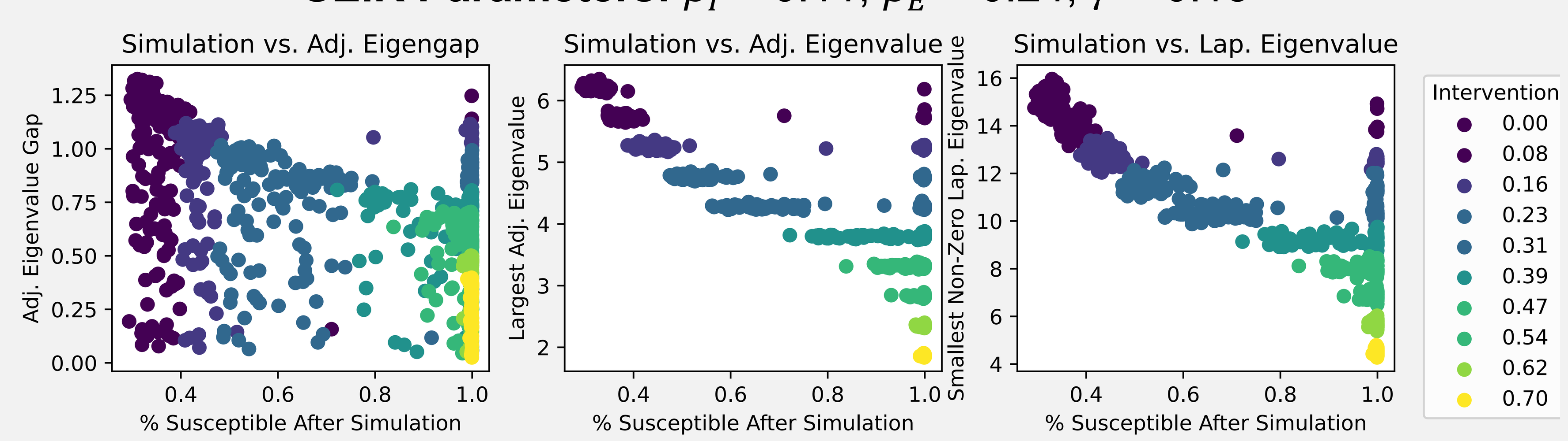

Figure 5: *Spectral properties track the magnitude of disease spread during the SEIR process.*

A deep network was created to learn the relationship between the spectral properties and the SEIR curve. A model input consists of 7 elements to capture network data:

$$V_0 = [\ S_0,\ I_0 + E_0,\ R_0,\ \texttt{n\_nodes},\ \texttt{max\_adj},\ \texttt{sm\_lap},\ \texttt{adj\_gap}\ ]$$

$$f(V_i) = [V_{i+1}, V_{i+2}, \dots, V_{i+150}]$$

LSTM 16 → **Dense** 16 → **Dense** 32 → **Repeat** T=150 → **LSTM** 16, T=150 → **Dense** 32, T=150 → **Dense** 3, T=150

Results

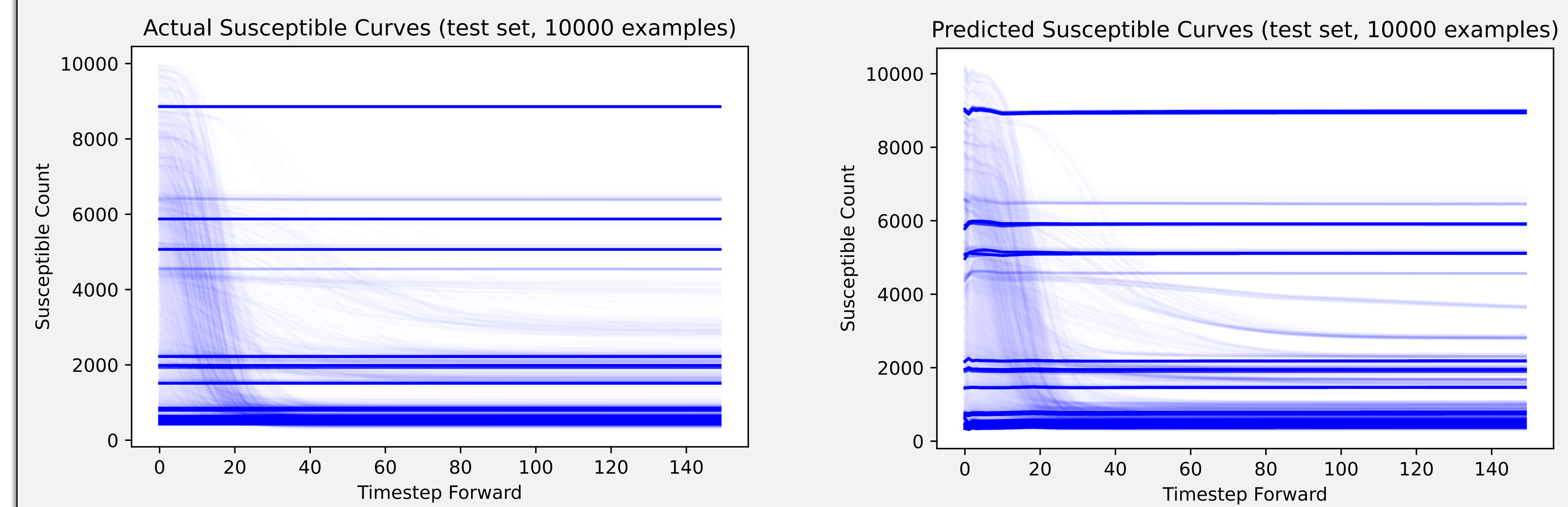

Figure 6: *An ML model can fit SEIR curves with a 7-variable encoding*

Accuracy: The model had a mean squared error of 1.417×10^{-4} on our test set. The predicted curves in this regime were off by approximately 67 people in a given timestep, or roughly 0.5-1% on average.

Performance: SEIR curves were created via simulation and prediction over 400 random graphs in the same regime, and performance was measured.

- **Simulated:** 2.80sec* per simulation
- **Model:** 0.06sec* per model forward pass

Impact: It’s possible to use spectral properties (in this regime) to predict curves that are “good enough” for planning purposes, at much higher performance than traditional simulation could allow for.

** Experiments ran on an Intel(R) Xeon(R) Platinum 8260 CPU @ 2.40GHz, averaged over 100 runs. SEIR was performed with graph_tool’s SIRState simulation tool. The model was done in Tensorflow.*

Conclusion and Future Work

Conclusion

- The deep LSTM model is able to predict the SEIR curve of a simulation over an interaction network using high-level network properties.
- We showed that spectral properties can be used as a surrogate to measuring disease spread over interaction networks with community structure.

Future work

- Future research directions could indicate other structures beyond SBMs.
- Proposal in process to fund future development of this work in the area of dynamic graphs and more robust state embeddings.

References

[1] Ho, Christopher et al. “Control of epidemics on graphs”. 2015 54th IEEE Conference on Decision and Control (CDC). N.p., 2015. 4202–4207. Web.

[2] Tong, Hanghang, B. Aditya Prakash, Tina Eliassi-Rad, Michalis Faloutsos, and Christos Faloutsos. "Gelling, and Melting, Large Graphs by Edge Manipulation." Proceedings of the 21st ACM International Conference on Information and Knowledge Management - CIKM '12 (2012). Print.

[3] Smedemark-Margulies, Niklas et al. “Inference in Network-based Epidemiological Simulations with Probabilistic Programming”. AI for Public Health Workshop, ICLR. N.p., 2021. Web.

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Rensselaer

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

This material is based upon work supported by the Under Secretary of Defense for Research and Engineering under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Under Secretary of Defense for Research and Engineering.

© 2021 Massachusetts Institute of Technology.

Delivered to the U.S. Government with Unlimited Rights, as defined in DFARS Part 252.227-7013 or 7014 (Feb 2014). Notwithstanding any copyright notice, U.S. Government rights in this work are defined by DFARS 252.227-7013 or DFARS 252.227-7014 as detailed above. Use of this work other than as specifically authorized by the U.S. Government may violate any copyrights that exist in this work.