Inference of Disease Dispersion Rates of a Branching Model with Spatial and Temporal Information using GNN

Qing Yao[1,2], and Sen Pei[2*]

Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032 1 and Complexity and Networks, Imperial College London, London SW7 2BX, United Kingdom 2

(Dated: June 28, 2023)

Timely control of infectious disease outbreaks requires accurate modeling and prediction of disease spread. The complexity increases when accounting for superspreading events, where a small number of individuals cause a large proportion of infections [1]. This project presents a novel approach to this problem by leveraging Graph Neural Networks (GNN) to infer the dispersion rate of a branching model using spatial and temporal disease spread patterns.

We constructed a graph-based branching model [2] to simulate the spatial spread of an infectious disease across U.S. counties, factoring in inter-county commuting and superspreading events [3]. The branching model used a negative binomial distribution to generate the number of secondary infections each individual causes.

We trained a GNN on this model to learn the dispersion rate of superspreading events. The GNN was able to capture the highly stochastic transmission dynamics and resulted in more accurate predictions of disease spread patterns than traditional models. Our GNN-based approach allows for the quick and robust estimation of disease dispersion rates, even in the face of significant uncertainty and variability. This study highlights the potential of GNNs in advancing disease modeling and in facilitating a more effective response to disease outbreaks.

---

[1] Lloyd-Smith, James Oet al. Nature, 438(7066):355–359, 2005.

[2] Jacob, Christine. International journal of environmental research and public health, 7(3):1186–1204, 2010.

[3] Pei, Sen et al. Proceedings of the National Academy of Sciences, 115(11):2752–2757, 2018.

* sp3449@cumc.columbia.edu 2