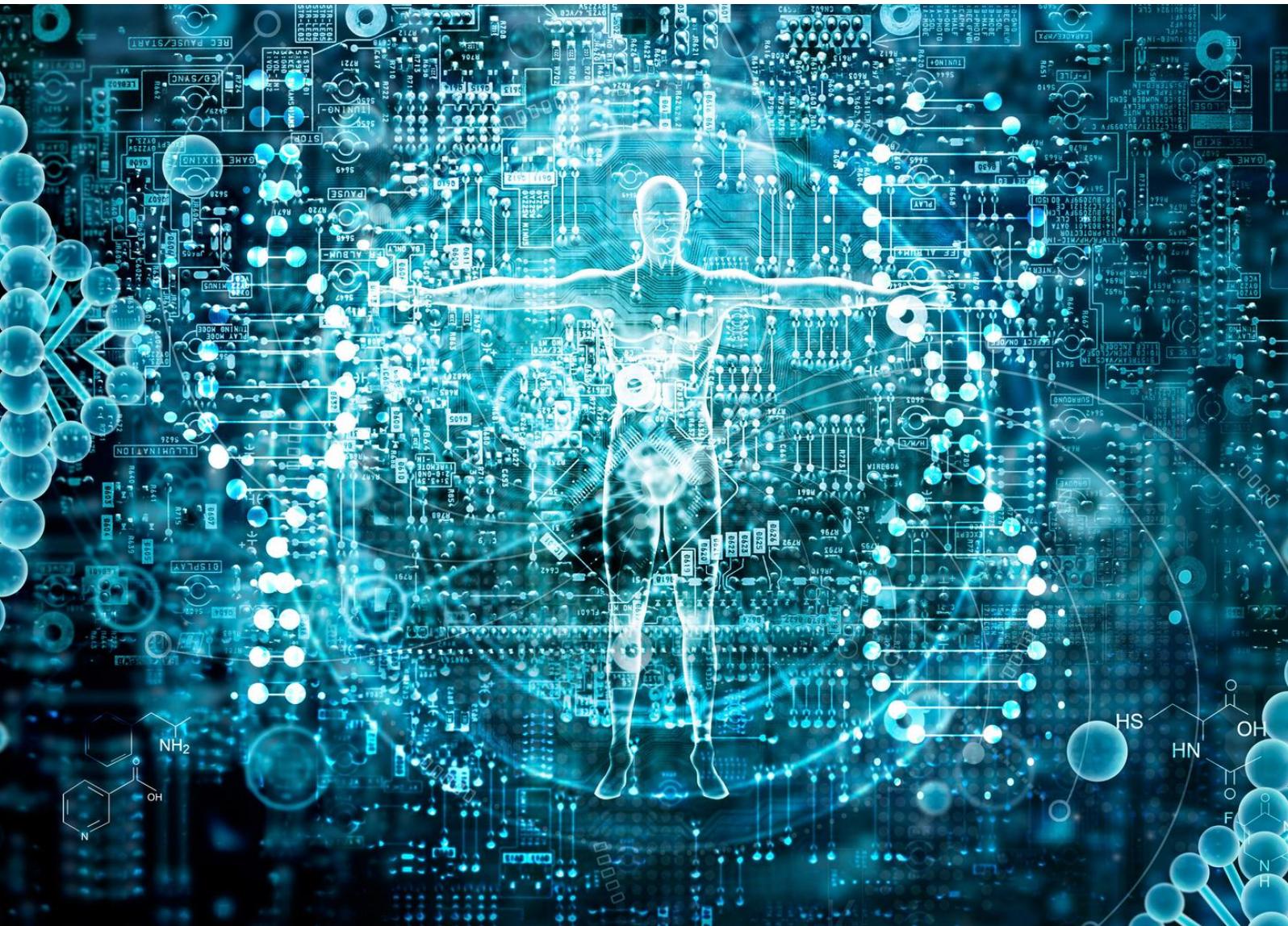


November 2022

CompEpi Quarterly

Newsletter of the Global Pervasive Computational Epidemiology project



Georgia Institute of Technology
Lawrence Livermore National Laboratory

JSU JACKSON STATE UNIVERSITY
ONEHEALTH TRUST



UNIVERSITY OF VIRGINIA
PRINCETON UNIVERSITY



Yale
VIRGINIA TECH



computational-epidemiology.org

NSF Midpoint Review

It's been a busy fall, and thanks to the team's hard work we can proudly say that our NSF midpoint review was a great success!

This was an important time for the team, as most of our research has pivoted away from COVID-19 and we work to determine our team's direction for the remainder of this project. We've got some exciting initiatives ahead to help develop more collaborative opportunities, improve our presence in the BPC space, and engage our students and researchers in innovative ways.

You can read about our revised team management plan on page 8, but one thing we really want to highlight is our new GPCE Dashboard. This is a one-stop resource for the team, complete with quick forms to add things to our internal repository, information about upcoming events, reminders for important dates, and links to our social media sites. This is YOUR resource, so if there are additional items you think would be useful, please let us know!



Golda & Erin

Expeditions: GPCE Dashboard ☆

GPCE Team Dashboard

Important Dates

Annual report due April 1, 2023
Next All-Team meeting: May 2023 Hybrid meeting - date TBD

Helpful Links

GPCE Website

YouTube

Twitter

Slack

Internal Repo

Teams

Scroll down for specific team links and updates

Admin: Madhav Marathe, Anil Vullikanti, Golda Barrow, Erin Raymond

EAC Liaisons: Madhav Marathe, Anil Vullikanti, Simon Levin, Rita Colwell, TBD

Research Integration: Samarth Swarup, Jeffrey Townsend, Michael Bernstein, Dave Higdon, Stephen Eubank

IGES: Geoffrey Fox, Naren Ramakrishnan, Kristy Collins, Anuj Mubayi, Baltazar Espinoza, Natarajan Meghanathan

Education & Outreach: Aravind Srinivasan, B. Aditya Prakash, Lenny Heath, Natarajan Meghanathan, Anil Vullikanti, SS Ravi, Baltazar Espinoza

Infrastructure: Anil Vullikanti, Rok Sosic, Judy Fox

Knowledge Transfer: Bryan Lewis, Stephen Eubank, Ramanan Laxminarayan, Rok Sosic

Upcoming Events

Seminar series
2nd & 4th Thursdays, [Zoom link](#)

- 23 Feb Mami Taniuchi, UVA
- 9 Mar Srini Venkatramanan, UVA
- 23 Mar Sriram Pemmaraju, Univ of Iowa
- 13 Apr Michela Meister, Cornell
- 27 Apr Colin Carlson, Georgetown
- 11 May Linsey Marr, VT
- 8 Jun Baltazar Espinoza, UVA

Workshops

External Advisory Committee

Sir Roy Anderson, Imperial College London

Jack Dongarra, Univ. of Tennessee

Dylan George, CDC

Lise Getoor, UC Santa Cruz

Linda Hayden, ECSU

Sir Richard Roberts, NE Biolabs

Publications and Presentations Input Form

Internal Repository Dataset Input Form

Internal Repository Source Code Input Form

Internal Repository Tools & Websites Input Form

COMING SOON
CECURE Teaching Materials Input Form

Teams

BPC Team	Education & Outreach Team	Infrastructure Team
BPC Google folder BPC Standalone Project Plan (Dec. 2022) BPC Team meeting notes	Education & Outreach Google folder Education & Outreach Plan (Dec. 2022) Education & Outreach Team mtg notes	Infrastructure Google folder Meeting schedule: TBD

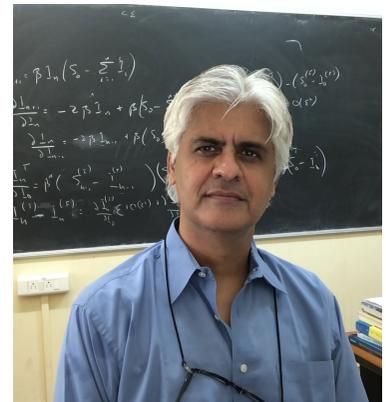
.....

Seminar Series - 14 July 22



Shift, Scale and Restart Smaller Models to Estimate Larger Ones: Agent-based Simulators for Covid Modelling

Agent-based simulators are a popular epidemiological modelling tool to study the impact of various non-pharmaceutical interventions in managing an evolving pandemic. They provide the flexibility to accurately model a heterogeneous population with time and location varying, person specific interactions. To accurately model detailed behaviour, typically each person is separately modelled. This however, may make computational time prohibitive when the region population is large and when time horizons involved are large. We observe that simply considering a smaller aggregate model and scaling up the output leads to inaccuracies. In this talk we primarily focus on the COVID-19 pandemic and dig deeper into the underlying probabilistic structure of an associated agent based simulator (ABS) to arrive at modifications that allow smaller models to give accurate statistics for larger models. We exploit the observations that in the initial disease spread phase, the starting infections behave like a branching process. Further, later once enough people have been infected, the infected population closely follows its mean field approximation. We build upon these insights to develop a shifted, scaled and restart version of the simulator that accurately evaluates the ABS's performance using a much smaller model while essentially eliminating the bias that otherwise arises from smaller models.



Dr. Sandeep Juneja



Sandeep is a senior professor at the School of Technology and Computer Science in Tata Institute of Fundamental Research in Mumbai. His research interests lie in applied probability including in sequential learning, financial mathematics, Monte Carlo methods, and game theoretic analysis of queues. Lately, he has been involved in modelling Covid-19 spread in Mumbai, and in mathematics of agent-based simulation models. Sandeep received his B. Tech. from IIT Delhi and Masters in Statistics and PhD in Operations Research from Stanford University. He is currently on the editorial board of Stochastic Systems. Earlier he has been on editorial boards of Mathematics of Operations Research, Management Science and ACM TOMACS.

*Would you like to present at a future Seminar?
Email Golda and Erin with your availability, title, and abstract.*

Seminar Series - 25 Aug 22

Click here
to watch on
YouTube

Uncertainty and the management of outbreaks: harnessing the power of multiple models

During outbreaks of weeds, pests and infectious diseases, uncertainty hinders our ability to forecast dynamics, and to make critical decisions about management. In particular, disparate epidemiological projections from different modeling groups, arising from different scientific descriptions of the underlying biological and management processes, may hamper intervention planning and response by policy makers. Drawing on methods from expert elicitation and judgment, we can harness the expertise of multiple modeling groups in a decision theoretic framework. Case studies of several diseases illustrate how to gain a better idea of expected outcomes, and the risk of particularly problematic consequences. This approach allows policy makers to focus on what matters most as they make critical decisions.



Dr. Katriona Shea



PennState
Eberly College of Science

.....

Katriona Shea is Professor of Ecology and the Alumni Professor in the Biological Sciences at the Pennsylvania State University. She received her BA (Hons) in Physics from Oxford in 1990 and her PhD in theoretical population ecology from Silwood Park, Imperial College, London University in 1994. Following postdoctoral positions in California and Australia, she joined the faculty at Penn State in 2001.

Professor Shea uses a wide range of empirical and quantitative methods to study the ecology and management of invasive and outbreaking species in perturbed environments, with a particular focus on the role of uncertainty.

Professor Shea is an elected Fellow of the Ecological Society of America (ESA) and of the American Association for the Advancement of Science (AAAS).

More information can be found at <https://kshealab.wordpress.com/>.

Seminar Series - 20 Oct 22

Click here
to watch on
YouTube

Bayesian model calibration for constraining ABMs

Agent-based models (ABMs) use rules at the individual (agent) level to simulate a social, ecologic, or social-technical system, producing structured behavior when viewed at an aggregated level. ABMs allow consideration of interactions, behaviors, outcomes, and interventions at a very fine resolution, giving them a connection to reality that is often missing from more aggregated modeling approaches. Of course this comes with a price. For example, ABMs typically come with high computational cost, random realizations, and difficulty in capturing more aggregated properties of the real system. This talk will use a toy example to highlight the key concepts of Bayesian model calibration for constraining ABMs with observations; it'll also touch on Approximate Bayesian Computation (ABC). I'll then go on to show how such methodology can be used in more involved applications, using an ABM developed during the 2014 Ebola epidemic.



Dr. David Higdon



.....

David M. Higdon is a professor in the Statistics Dept at Virginia Tech. Previously, he spent 10 years as a scientist or group leader of the Statistical Sciences Group at Los Alamos National Laboratory. He is an expert in Bayesian statistical modeling of environmental and physical systems, combining physical observations with computer simulation models for prediction and inference. Dr. Higdon has served on several advisory groups concerned with statistical modeling and uncertainty quantification and co-chaired the NRC Committee on Mathematical Foundations of Validation, Verification, and Uncertainty Quantification. He is a fellow of the American Statistical Association.

Seminar Series - 3 Nov 22



Neighborhood-based Bridge Node Centrality Tuple for Complex Network Analysis

We define a bridge node to be a node whose neighbor nodes are sparsely connected to each other and are likely to be part of different components/disjoint clusters if the node is removed from the network. We propose a computationally light Neighborhood-based Bridge Node Centrality (NBNC) tuple that could be used to identify the bridge nodes of a network as well as rank the nodes in a network on the basis of their topological position to function as bridge nodes. The NBNC tuple for a node is asynchronously computed on the basis of the neighborhood graph of the node that comprises the neighbors of the node as vertices and the links connecting the neighbors as edges. The NBNC tuple for a node has three entries: the number of disjoint components in the neighborhood graph of the node, the algebraic connectivity ratio of the neighborhood graph of the node and the number of neighbors of the node. We analyze a suite of 50 complex real-world networks and evaluate the computational lightness, effectiveness, efficiency/accuracy and uniqueness of the NBNC tuple vis-a-vis the existing bridgeness related centrality metrics and the Louvain community detection algorithm. From a computational epidemiology point of view, the NBNC tuple could be used to effectively and efficiently identify bridge nodes that need to be vaccinated in a social network or organizational networks to provide herd immunity to nodes that are not vaccinated.



Dr. Natarajan Meghanathan



.....

Dr. Natarajan Meghanathan is a tenured Full Professor of Computer Science at Jackson State University, Jackson, MS. He graduated with a Ph.D. in Computer Science from The University of Texas at Dallas in May 2005. Dr. Meghanathan has published more than 150 peer-reviewed articles (more than half of them being journal publications). He has also received federal education and research grants from the U. S. National Science Foundation, Army Research Lab and Air Force Research Lab. Dr. Meghanathan has been serving in the editorial board of several international journals and in the Technical Program Committees and Organization Committees of several international conferences. His research interests are Wireless Ad hoc Networks and Sensor Networks, Systems and Software Security, Graph Theory Algorithms, Machine Learning, Cloud Computing and Computational Biology.



Publications & Presentations



Our team has been busy supporting the challenges of real-time epidemic science, particularly pertaining to the COVID-19 Pandemic.

AI Techniques for Forecasting Epidemic Dynamics: Theory and Practice. [Artificial Intelligence in Covid-19](#). Lidströmer and Eldar (Eds), Springer Cham. Nov 10;193-228. Adiga A, Lewis B, Levin S, Marathe MV, Poor HV, Ravi SS, Rosenkrantz DJ, Stearns RE, Venkatramanan S, Vullikanti A, Wang L (2022)

A Web-based System for Contagion Simulations on Networked Populations. [IEEE eScience](#). October. Ferdousi T, Kishore A, Machi L, Machi D, Kuhlman C, Ravi SS (2022)

Sociocultural determinants of global mask-wearing behavior. [PNAS](#). Oct 3; 119(41): e2213525119. Yang L, Constantino SM, Grenfell BT, Weber EU, Levin SA, Vasconcelos VV (2022)

Techniques for blocking the propagation of two simultaneous contagions over networks using a graph dynamical systems framework. [Network Science](#). September; 10(3): 234-260. Carscadden HL, Kuhlman CJ, Marathe MV, Ravi SS, Rosenkrantz DJ (2022)

Using Active Queries to Infer Symmetric Node Functions of Graph Dynamical Systems. [Journal of Machine Learning Research](#). August; 23:1-43. Adiga A, Kuhlman CJ, Marathe MV, Ravi SS, Rosenkrantz DJ, Stearns RE (2022)

Efficiently Learning the Topology and Behavior of a Networked Dynamical System Via Active Queries. [Proceedings of the Thirty-Ninth International Conference on Machine Learning](#). July. Rosenkrantz DJ, Adiga A, Marathe MV, Qiu Z, Ravi SS, Stearns RE, Vullikanti A (2022)

Understanding the coevolution of mask wearing and epidemics: A network perspective. [PNAS](#). Jun 22; 119(26): e2123355119. Qiu Z, Espinoza B, Vasconcelos VV, Chen C, Constantino SM, Crabtree SA, Yang L, Vullikanti A, Chen J, Weibull J, Basu K, Dixit A, Levin SA, Marathe MV (2022)

Vaccination-hesitancy and global warming: distinct social challenges with similar behavioural solutions. [Royal Society Open Science](#). Jun 15; 9: 211515. Fischer I, Rubenstein DI, Levin SA (2022)

The durability of natural infection and vaccine-induced immunity against future infection by SARS-CoV-2. [PNAS](#). Jul 15;119(31):e2204336119. Townsend J, Hassler H, Sah P, Dornburg A (2022)

Combining Theory and Data-driven Approaches for Epidemic Forecast. [Knowledge-guided Machine Learning: Accelerating Discovery using Scientific Knowledge and Data](#). Karpatne, Kannan, Kumar (Eds), CRC Press. Wang L, Adiga A, Chen J, Lewis B, Sadilek A, Venkatramanan S, Marathe M (2022)

.....
Probabilistic Methods in Computer Science, Data Science, and Public Health. [Boeing Distinguished Colloquium](#). Srinivasan A (Apr 2022)

Climate and Health: Conquering Water-Borne Diseases in the 21st Century. [ONWARD Webinar Series on Water-Related Diseases, Their Links To Environmental Conditions, Water Quality, Monitoring Methods, and Solutions](#), AIR Centre. Colwell R (Mar 2022)

Just Because You Are A Pawn You Don't Have To Lose The Game. [UMD CMNS DEI Lecture](#). Colwell R (Mar 2022)

Rethinking Robustness: From Classification to Contextual Bandits. [USC Department of Mathematics Probability and Statistics Seminar](#). Moitra A (Oct 2021)



GPCE Management Teams

Research Integration: Samarth Swarup, Jeffrey Townsend, Michael Bernstein, Dave Higdon, Stephen Eubank



External Advisory Committee Liaisons: Madhav Marathe, Anil Vullikanti, Rita Colwell, Simon Levin

Broadening Participation in Computing: Geoffrey Fox, Naren Ramakrishnan, Kristy Collins, Anuj Mubayi, Baltazar Espinoza, Natarajan Meghanathan



Knowledge Transfer: Bryan Lewis, Stephen Eubank, Ramanan Laxminarayan, Rok Susic

Education and Outreach: Aravind Srinivasan, B. Aditya Prakash, Lenny Heath, Natarajan Meghanathan, Anil Vullikanti, SS Ravi, Baltazar Espinoza



Infrastructure: Anil Vullikanti, Rok Susic, Judy Fox

Admin: Madhav Marathe, Anil Vullikanti, Golda Barrow, Erin Raymond



Student Poster Session

As always, we are very proud of our students! We had 14 poster presentations during the review meeting - great job everyone!

Phase-Informed Bayesian Ensemble Models

Existing Method

- The current Bayesian ensemble combines forecasts from AR, LSTM, GNN, and SGM.
- The BMA pipeline is computed every week, independently for each country.
- The predictive distribution is assumed to be a mixture of Gaussians, with weights given by the posterior probabilities for each model in the BMA.

Proposed Method and Results

Phase segmentation

- Based on the weekly case counts, we classify the trend of the pandemic in one of the three following categories:
 - **Change phase**
 - **Stable phase**
 - **Decline phase**
- Designs a BMA ensemble that uses the knowledge of phase segmentation to train the model.
- Phase specific posterior performance of individual methods is used for estimating the weights.

Phase Informed BMA (PI-BMA)

- Break points are estimated recursively with each new week.
- PI package used: **segment**

Our Approach

- Propose a new Bayesian ensemble method with updated weighting schemes based on different phases of the pandemic.
- Evaluate proposed methods by fitting BMA on past data retrospectively.
- Identify and segment the observed time period into three main phases (below) in the pandemic as - stable, decline, and steady phase.
- Use phase specific performance of models to obtain the model weights in the BMA ensemble, independently for each country.

Performance evaluation across forecasting weeks

- Estimate the R_t value using the incidence case time series and a simulation model.
- We observe that different countries experience different R_t values and different set of methods perform well in different phases.
- The influence of a method i is defined as: $g_i^T = \frac{1}{N} \sum_{t=1}^N \mathbb{1}(g_i^T(t) = i)$

Results and retrospective evaluation

- PI-BMA outperforms other methods in most cases.
- PI-BMA is more robust to model misspecification.
- PI-BMA is more robust to data noise.

Summary and Conclusion

- PI-BMA models are useful but including every model in the ensemble may reduce forecast performance.
- Phases are identified to be important indicator of model specific performance.
- Departmental models are useful during growth and decline phases but tend to over estimate the future case counts.
- PI-BMA leads to improved performance at a critical phase, when compared to other methods.

Interpreting County Level COVID-19 Infection and Feature Sensitivity using Deep Learning Time Series Models

Abstract

Model Structure

Model Results

Population by Industry

Population by Age

Conclusion and Future Works

The stability of vector-mediated immunity against infection by SARS-CoV-2 and the frequency of mRNA booster vaccination

INTRODUCTION

METHODS

RESULTS

CONCLUSIONS

Loimos: A Large-Scale Epidemic Simulation Framework for Realistic Social Contact Networks

Motivation

Model (continued)

Parallel Implementation

Loimos Algorithm

Experimental Design

Results

Conclusions

Future Work

References

Supporting COVID-19 policy response with mobility networks and machine learning

Serina Chang
PhD Candidate, Computer Science, Stanford University | Advised by Jure Leskovec and Johan Ugander

Mobility network models of COVID-19 explain inequities and inform reopening

S. Chang*, E. Pierson*, P.W. Koh* et al., *Nature* 2021
Project website | NYTimes interactive article | Nature News & Views

Supporting COVID-19 policy response with large-scale mobility-based modeling

S. Chang, M. Wilson, B. Lewis, et al., *KDD* 2021

Estimating geographic spillover effects of COVID-19 policies from mobility networks

S. Chang, D. Vrabac, J. Leskovec, and J. Ugander

- We infer mobility networks for ten of the largest metro areas in the US → 553k POIs, 5.4B hourly edges
- Our networks + SEIR model enable accurate fits of COVID cases + deaths and capture its spread with unprecedented granularity: **who got infected where and when**
- Our model reveals how mobility patterns contribute to disparities and inform reopening strategies

Under review; presented at KDD 2022 workshop and ICOS'22

Many COVID policies defined locally → are populations spilling over from more to less restricted areas?

We develop a causal inference framework to make unconfounded estimates of spillovers from the California mobility network

Controlling Epidemic Spread using Probabilistic Diffusion Models on Networks

Motivation

Model

Methods

Results

Conclusions

Future Work

Multiple Covid-19 Time Series Forecasting

Background and Process

Seasonal Models

ARIMA

VAR(p)

LSTM

Deep Learning

Future Work

Individual Infection Risk Prediction: A Graph Learning Perspective

Problem: Predicting risk of infection for an individual person

Baseline models

This problem (surprisingly) difficult

Global infections in local state

Can we infer anything about the risk of a single person on a large graph of people from aggregate statistics?

Can we find techniques that work in production settings (large graphs) or can we learn representations that predict risk of infection with high accuracy?

SPRINT Student Pandemic Researchers - Interventions Data Analysis

Data Analysis

Network Systems Science and Advanced Computing

SPRINT Student Pandemic Researchers - Interventions Methodology

Methodology

Who We Are

What We Do

480+ MANDATES

38 INDEPENDENT CITIES

95 COUNTIES

Network Systems Science and Advanced Computing

From the admin team



We are excited to be a part of this team and community. Please let us know about your publications and presentations, if you'd like to present at our Seminar Series, or if you want to share the progress your group has made. Building a community on this scale requires an intentional effort and commitment, and we are here to support you in any way we can.

Golda Barrow {ggh5e@virginia.edu} & Erin Raymond {er9ff@virginia.edu}

Social media

We are always looking for content, so please let us know if there is a link, comment, or video we can share on social media.



Collaboration opportunities

Don't forget about these great team resources!



Join our Seminar Series the 2nd & 4th Thursday of every month!

- 12 Jan Nita Madhav, Ginkgo Bioworks
- 9 Feb Geoffrey Fox, UVA
- 23 Feb Mami Taniuchi, UVA
- 9 Mar Srini Venkatramanan, UVA

NSF acknowledgement

Remember to acknowledge NSF in your publications. Check with your grant manager for the appropriate number.

- In any publication (including web pages) of any material based on or developed under this project: “This material is based upon work supported by the National Science Foundation under Grant No. [CCF-xxxxxxx].”
- All publications except scientific articles or papers appearing in scientific, technical or professional journals: “Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.”
- NSF support must be orally acknowledged during all news media interviews, including popular media such as radio, television and news magazines.

