eScience '22 DEMOCRATIZING SCIENCE

UNIVERSITY of VIRGINIA
BIOCOMPLEXITY INSTITUTE

UNIVERSITY of VIRGINIA

A Web-Based System for Contagion Simulations on Networked Populations

Tanvir Ferdousi, Aparna Kishore, Lucas Machi, Dustin Machi, Chris J. Kuhlman, S. S. Ravi

Biocomplexity Institute and Initiative
University of Virginia

IEEE eScience 2022, October 11-14, 2022
Salt Lake City, Utah, United States

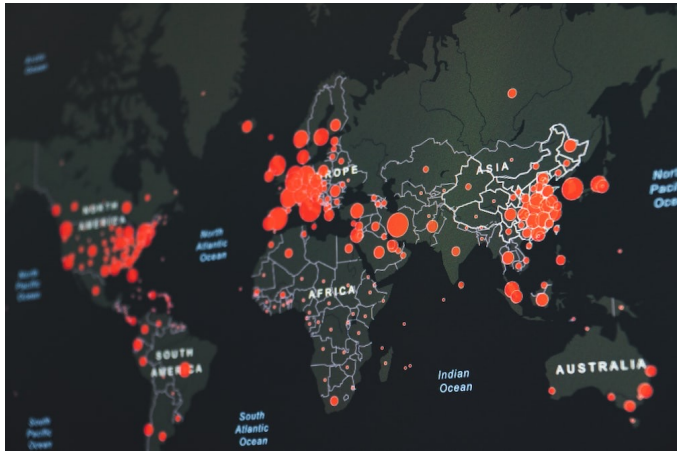
Acknowledgments: This research is supported by University of Virginia Strategic Investment Fund award number SIF160, VDH grant VDH-21-501-0135-1, and NSF Grants OAC-1916805 (CINES), CCF-1918656 (Expeditions) and CMMI-1916670 (CRISP 2.0).

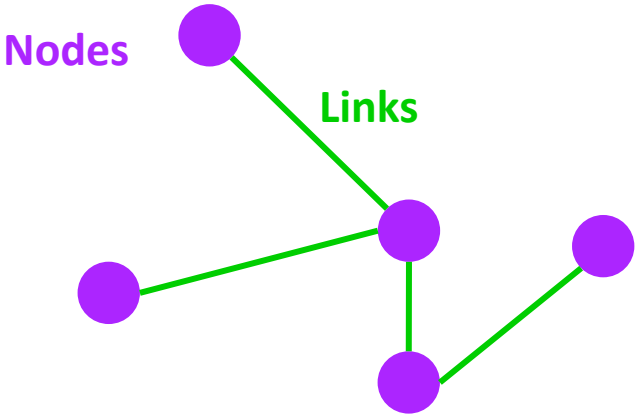
Outline

- Motivation
- System Architecture
- Performance Evaluation
- Case Studies
- Related Work
- Conclusions
- Limitations & Future Directions

Introducing Networks

- Networks are everywhere.
- Many situations can be represented by entities (nodes) & relationships (links).
- Network models and simulations are widely used.





Motivation

- Many **social processes** can be formulated as **contagion dynamics on networks**.
- **Agent-based simulation tools** are commonly used for networked systems.
- People need **software development experience** to **build and run** such tools.
- Some simulations require **high performance computing** resources.
- **Modeling and Simulation as a Service (MSaaS)** can alleviate many issues.

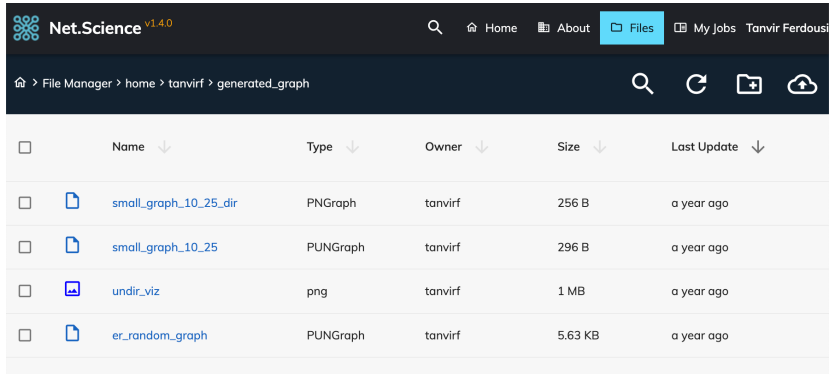
Highlights of Our Solution

- Network Simulation as a Service (NetSimS).
- Web-based open access system.
- Simulation module for contagion dynamics on networks.
- Graph seeding module for network state initialization.

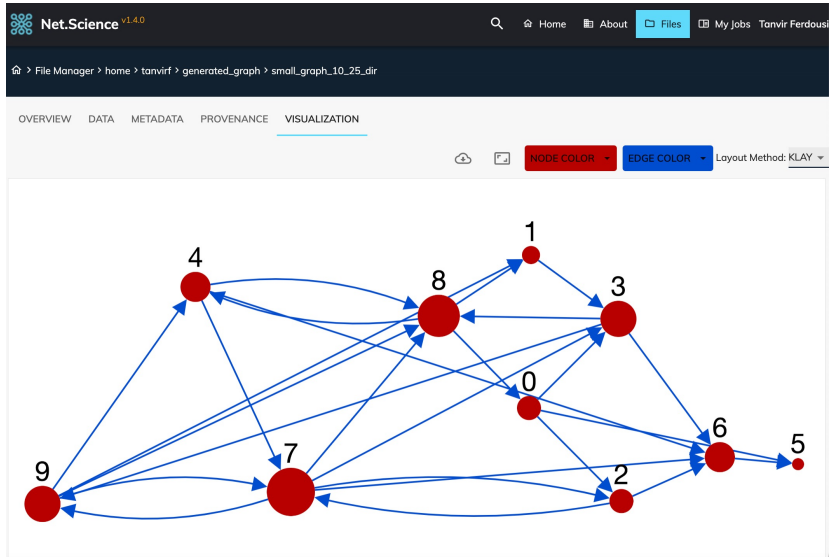
A Cyberinfrastructure for Network Science

- Open to public and accessible at: http://net.science
- Features
  - Modeling & Simulation as a Service (MSaaS).
  - Intuitive GUI & accessible APIs.
  - On demand high performance computing (HPC).
- The results in paper were generated using net.science.

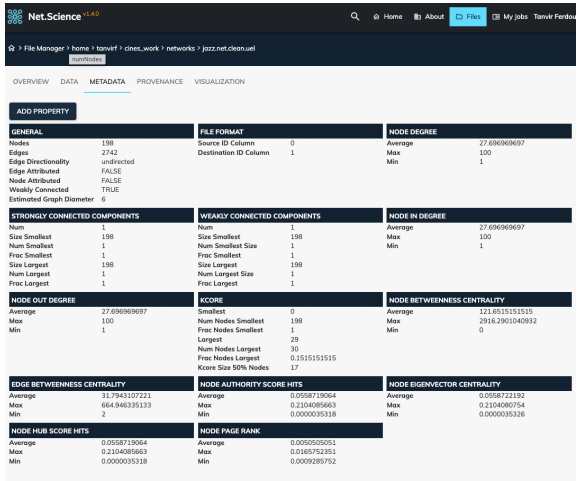
File System Browser



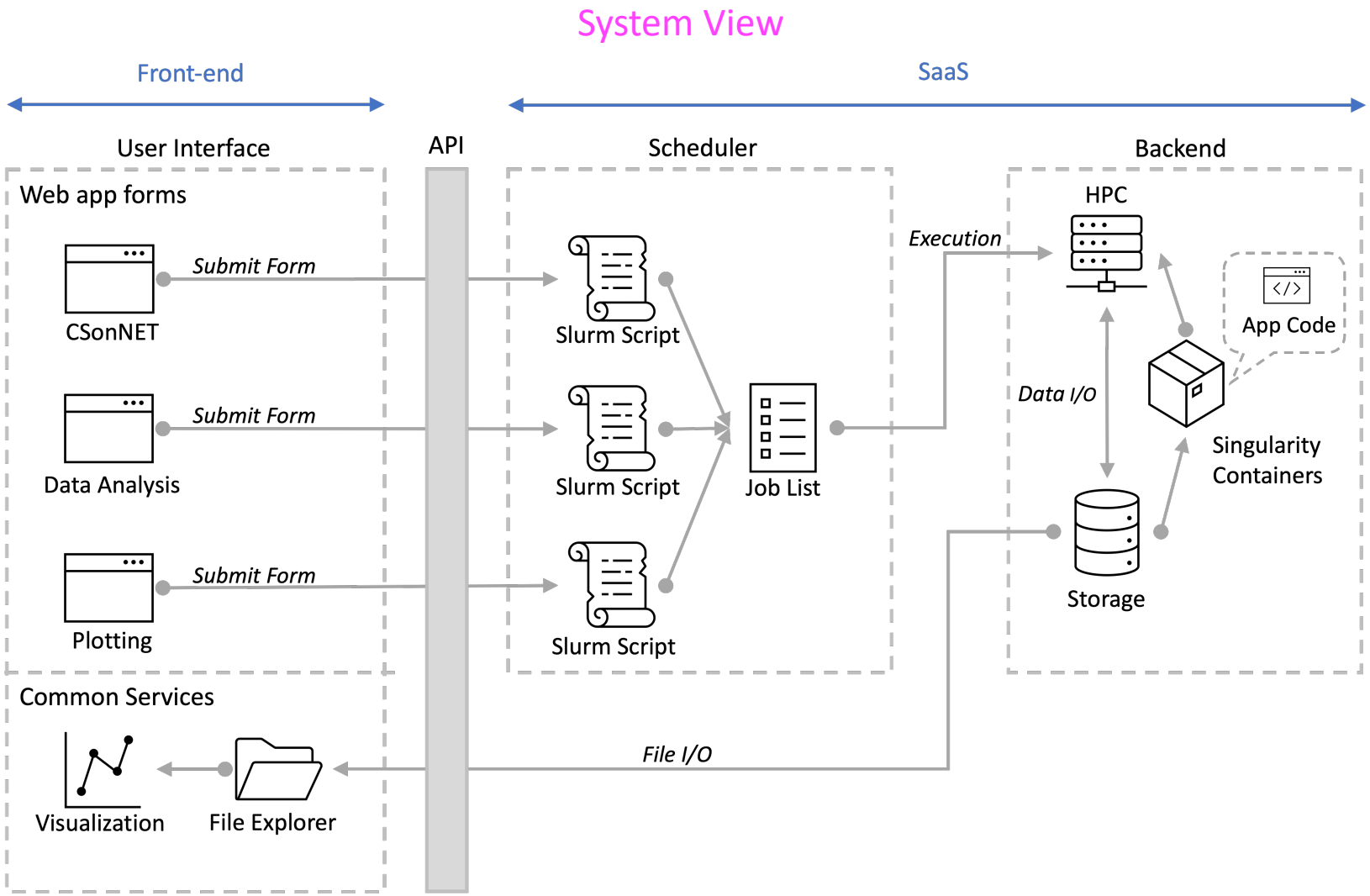
Interactive Graph Visualization



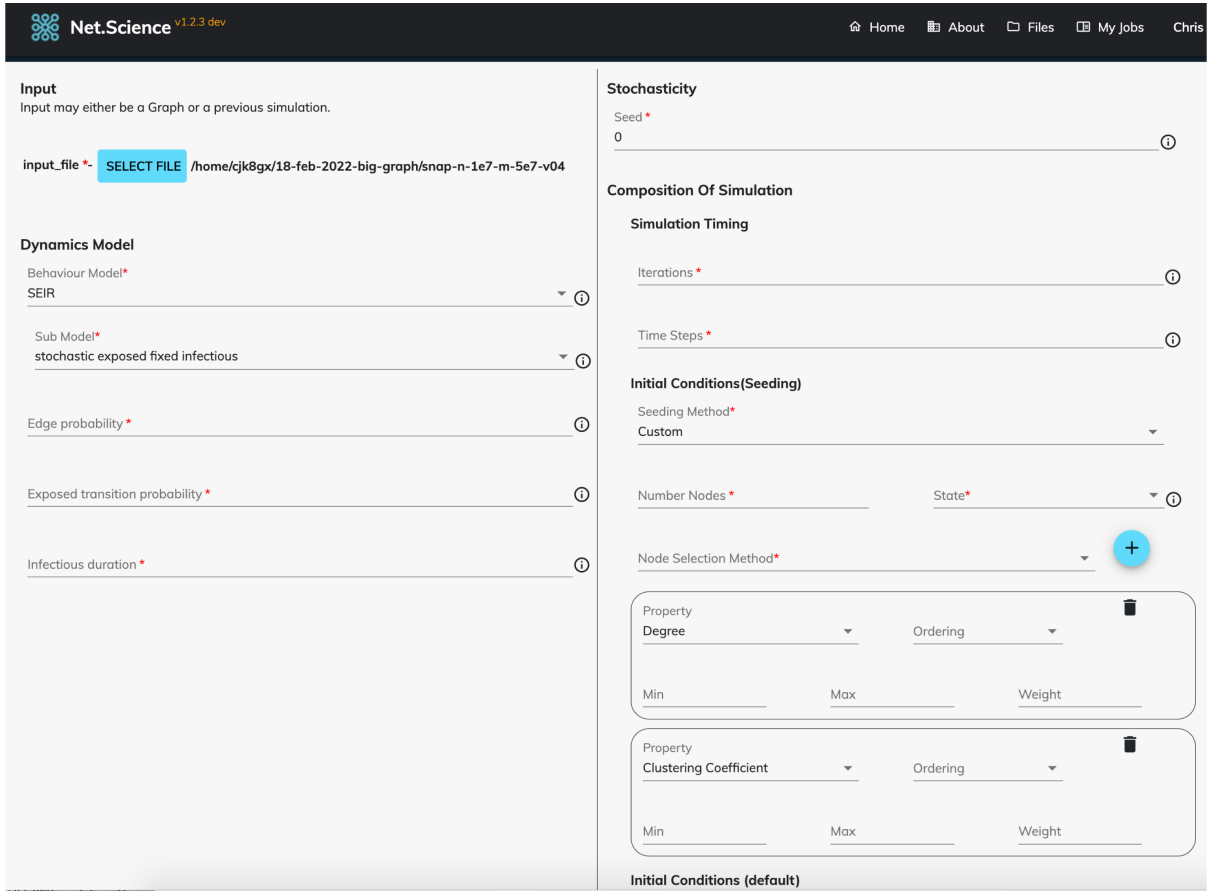
Graph Metadata



Network Simulation as a Service (NetSimS)

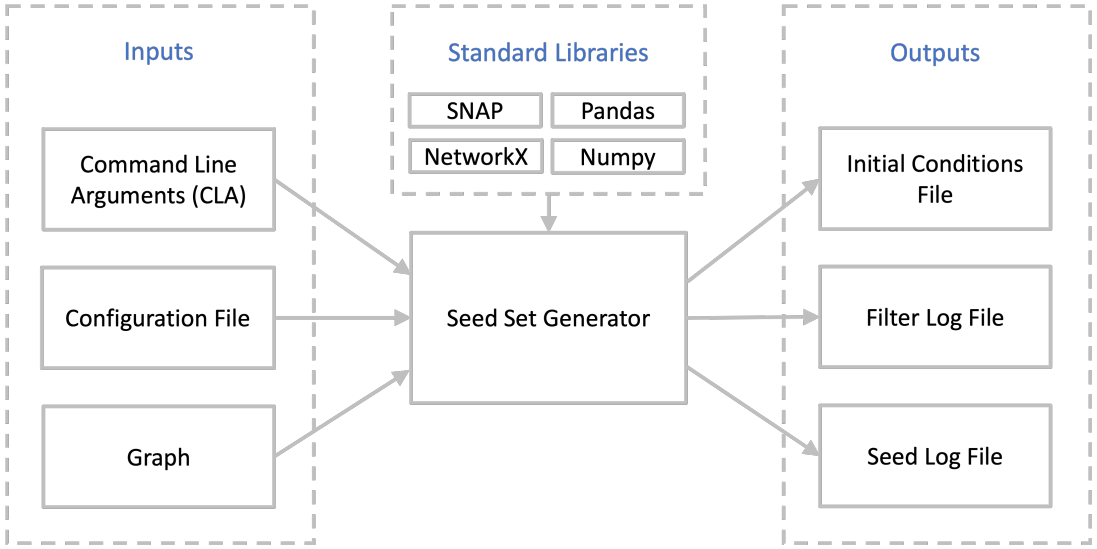


Web Interface



- Representative inputs for simulation
  - Agent model and parameters
  - Network
  - Number of iterations, time per iteration
  - Initial conditions
- Many types of web forms
  - Contagion simulation
  - Post-processing simulation data
  - Plotting
- Web app uses same API as other 3rd party tools.

Initial Conditions Module



seed.config.01

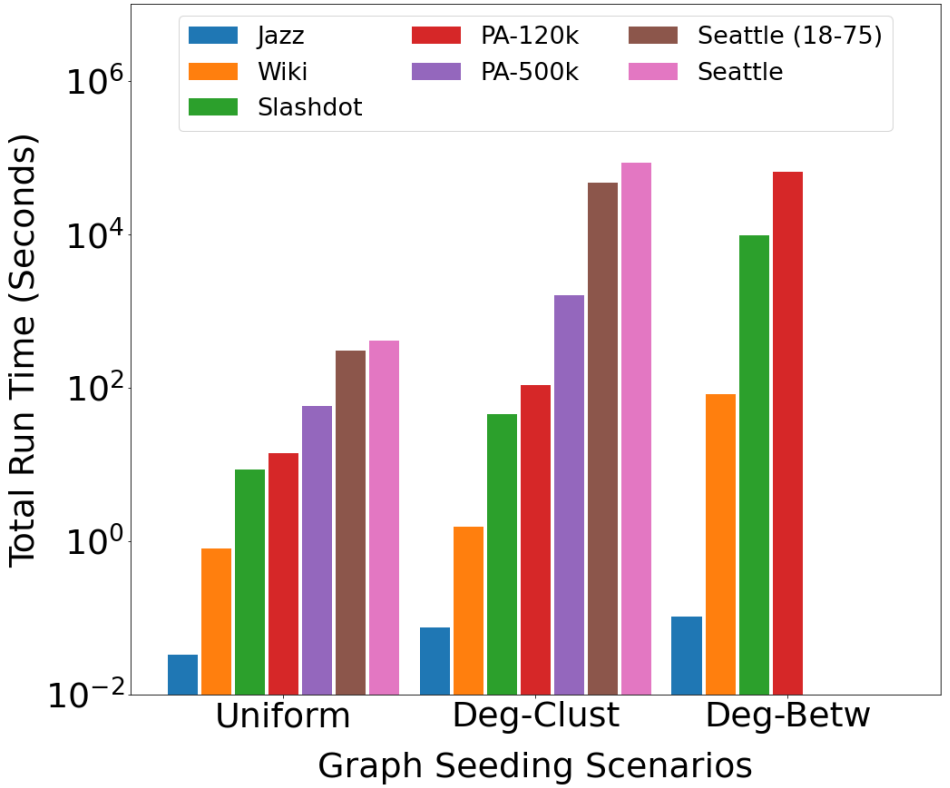
```
#graph_metric,sort_method,metric_min,metric_max,metric_weight
outdeg,high,3,4,1
betwcnt,high,0.2,6,1
clustcf,high,0.4,0.8,1
```

- Initialize graph node states.
  - Compute graph metrics.
  - Filter, sort nodes using metrics.
- Combine multiple metrics.
- Generate multiple seed sets
  - Deterministic
  - Stochastic
- 11 graph metrics are supported
  - Including degree, betweenness, clustering coefficient, pagerank, and hub/authority scores.

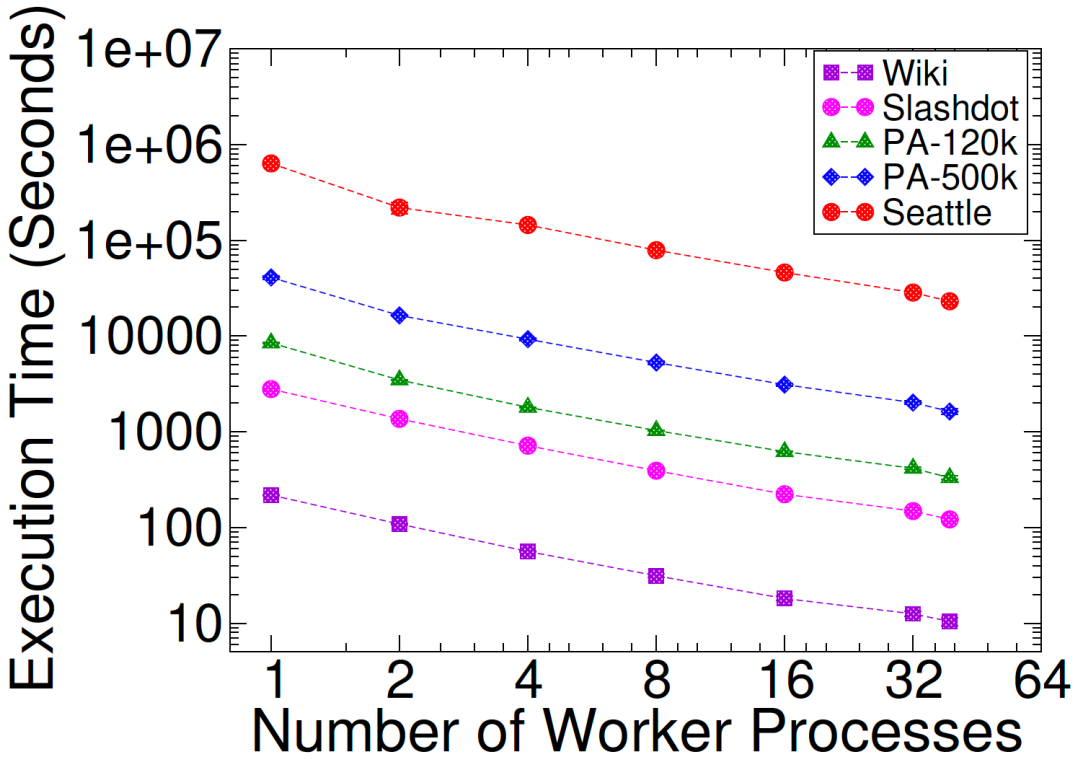
Performance Evaluation – Network Initialization

Network	Structural Properties				
	Num. Nodes	*Num. Edges*	*Avg. Degree*	*Max. Degree*	*Largest k-core*
Jazz, M	198	2,742	27.7	100	29
Wiki, M	7,066	100,736	28.51	1,065	53
Slashdot, M	77,360	469,180	12.13	2,539	54
PA-120k, S	120,000	2.4 M	39.99	2,686	20
PA-500k, S	500,000	9.99 M	39.99	5,405	20
Seattle (18-75), C	2.56 M	40.34 M	31.49	664	42
Seattle, C	3.52 M	66.51 M	37.82	879	43

M: Mined, C: Constructed, S: Synthetic

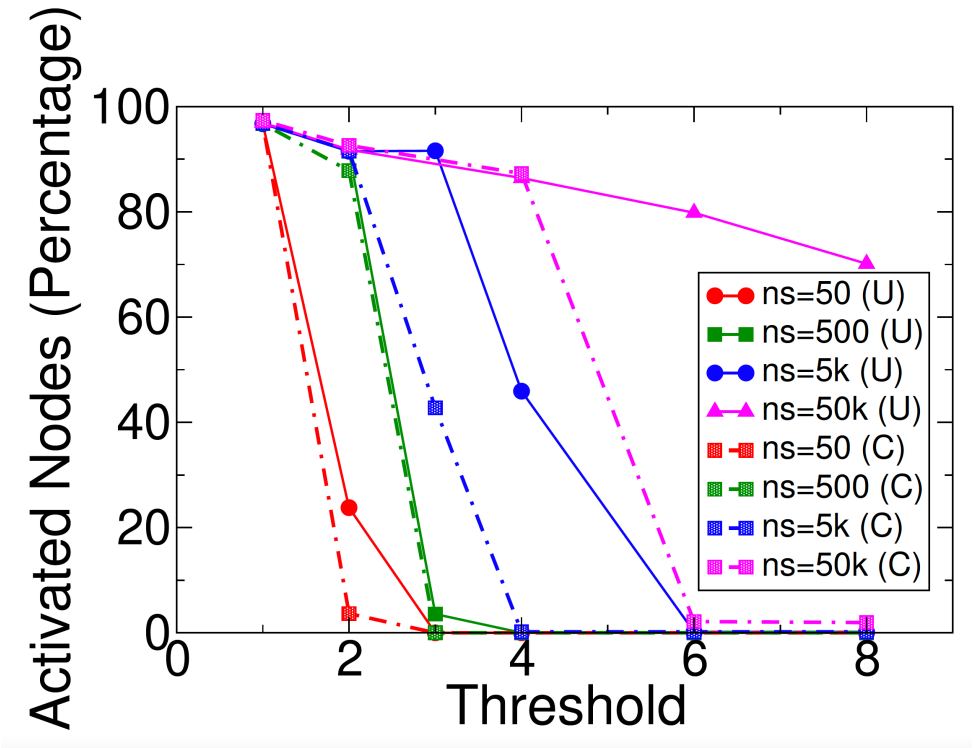


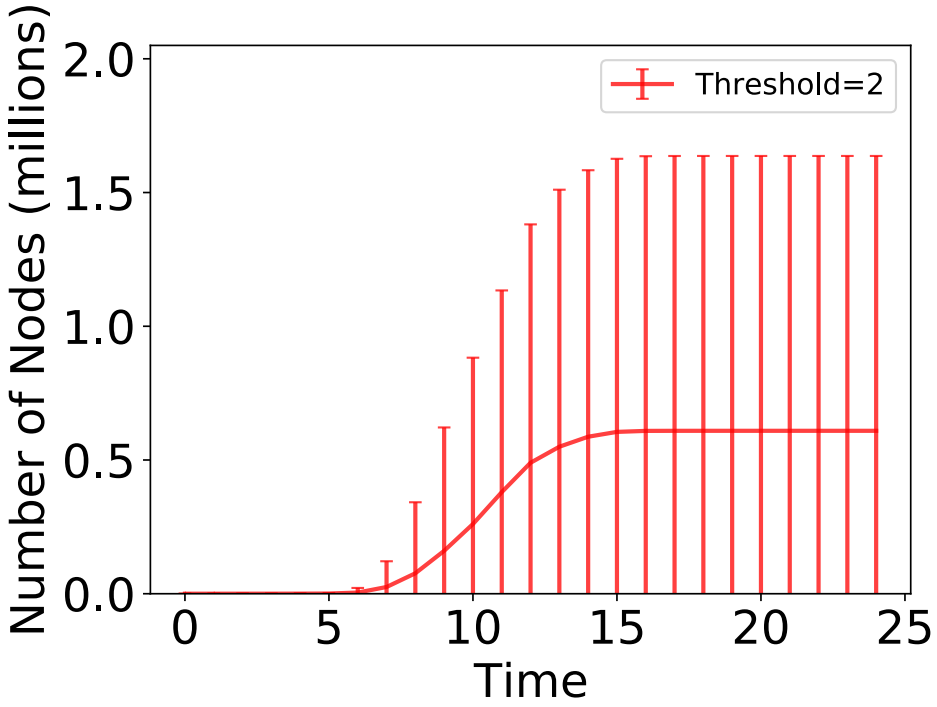
Strong Scaling Study – Agent-Based Simulation



- Uses SIR model
  - Transmission probability: 0.004
  - Infectious duration: 4 days
- Y-Axis: total time to run 100 iterations of 100 timesteps
- 300 nodes initialized as infected (I).
- The system exhibits strong scaling.

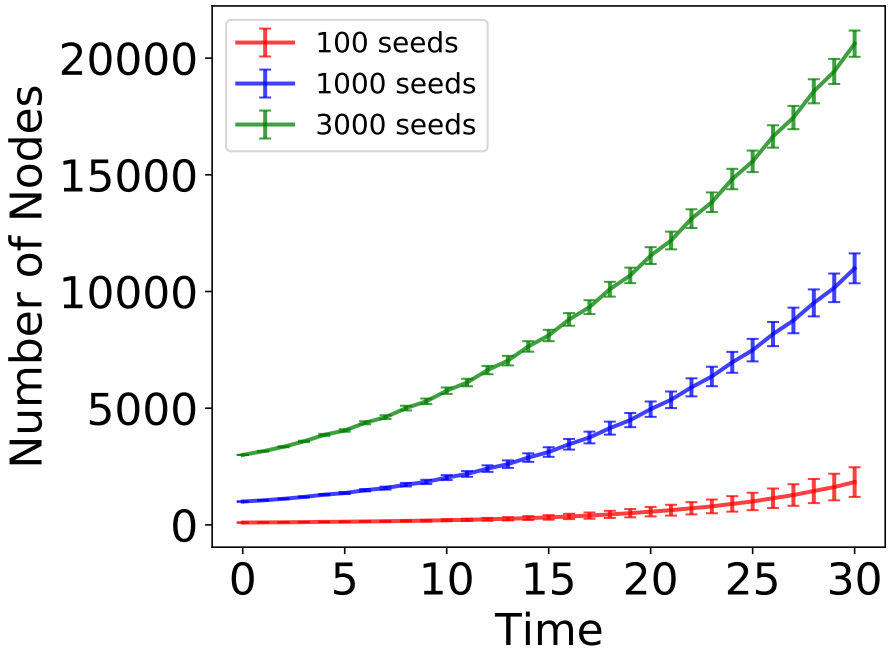
Case Study I – Simulation of Social Protests

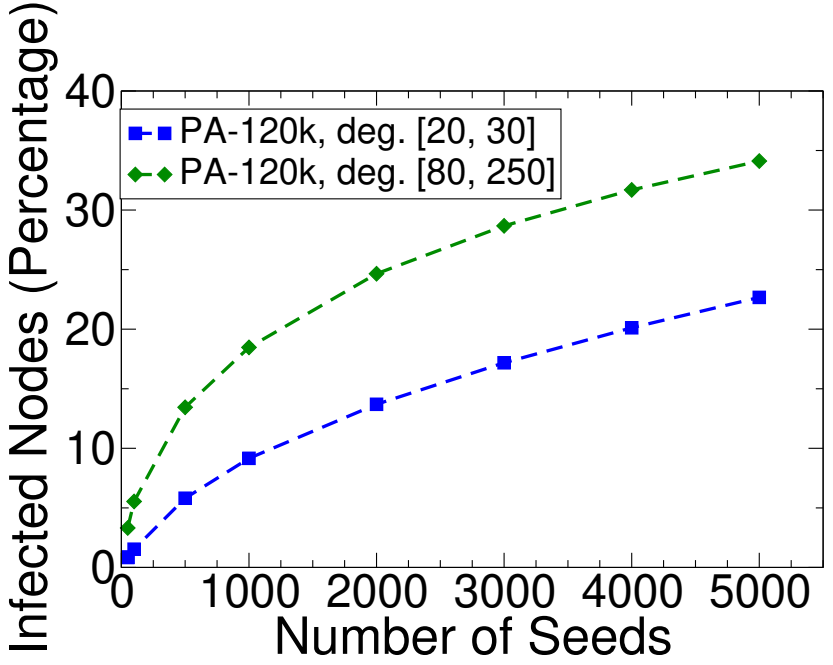




- Simulated using the Granovetter (1978) threshold model.
- Seattle (18-75) network was used.
- Two different seed node sampling methods were used (U, C).

Case Study II – Simulation of Virus Transmission





- (Left) Cumulative infection for SIR on PA-120k with three seed configurations with degree in [20,30].
- (Right) Final fraction of infected nodes after 30 days using two seeding configurations.

Related Work

- With IDE/GUI
  - Repast Simphony [North et al., 2013]
  - NetLogo [Wilensky, 2011]
- Geared towards parallel and high-performance computing
  - Repast HPC [Collier et al., 2013]
  - µsik [Perumalla, 2005]
- Modeling and Simulation as a Service (MSaaS)
  - Anylogic [http://anylogic.com]
  - CloudSME [Taylor et al., 2018]

Conclusions

- A system that generates initial conditions and runs contagion simulations on networks.
- On demand high performance computing (HPC).
- Open access, versatile simulation framework within net.science.
- Can benefit researchers, educators, and students.

Limitations & Future Directions

- Addition of new models requires working on both front and backends.
- Input network sizes are limited by available memory and processing power.
- Parallelization of graph analysis algorithms in the initial conditions module to improve performance.

Questions ?

Tanvir Ferdousi
Email: tanvir@virginia.edu
Project Site: https://net.science