

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).

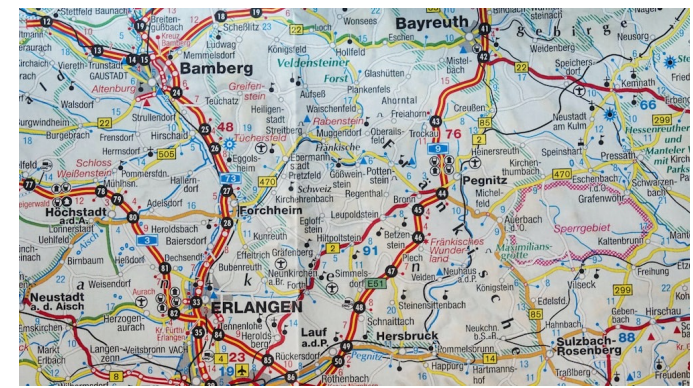
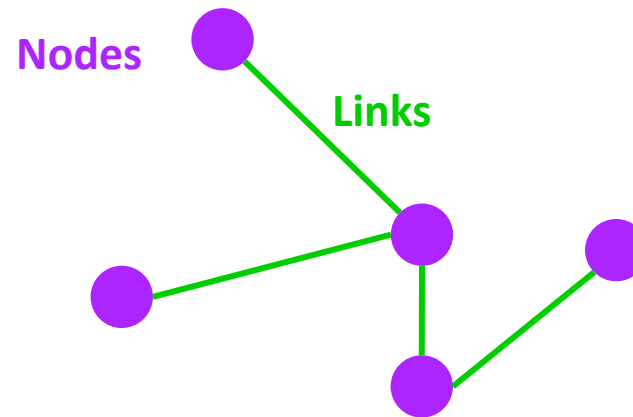
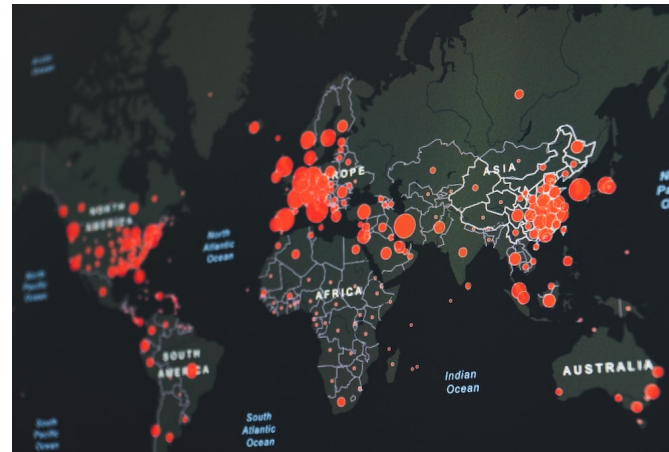


Outline

- 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.

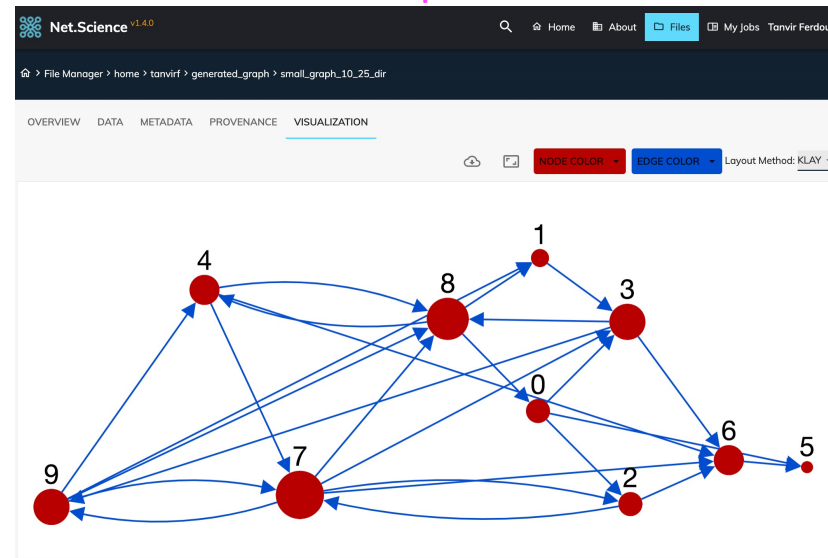
A Cyberinfrastructure for Network Science

- 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

Name	Type	Owner	Size	Last Update
small_graph_10_25_dir	PNGraph	tanvirf	256 B	a year ago
small_graph_10_25	PUNGraph	tanvirf	296 B	a year ago
undir_viz	png	tanvirf	1 MB	a year ago
er_random_graph	PUNGraph	tanvirf	5.63 KB	a year ago

Interactive Graph Visualization



Graph Metadata

GENERAL	FILE FORMAT	NODE DEGREE
Nodes: 198	Source ID Column: 0	Average: 27.69096967
Edges: 2742	Destination ID Column: 1	Max: 100
Edge Directionality: undirected		Min: 1
Edge Attributed: FALSE		
Node Attributed: FALSE		
Weakly Connected: TRUE		
Estimated Graph Diameter: 6		

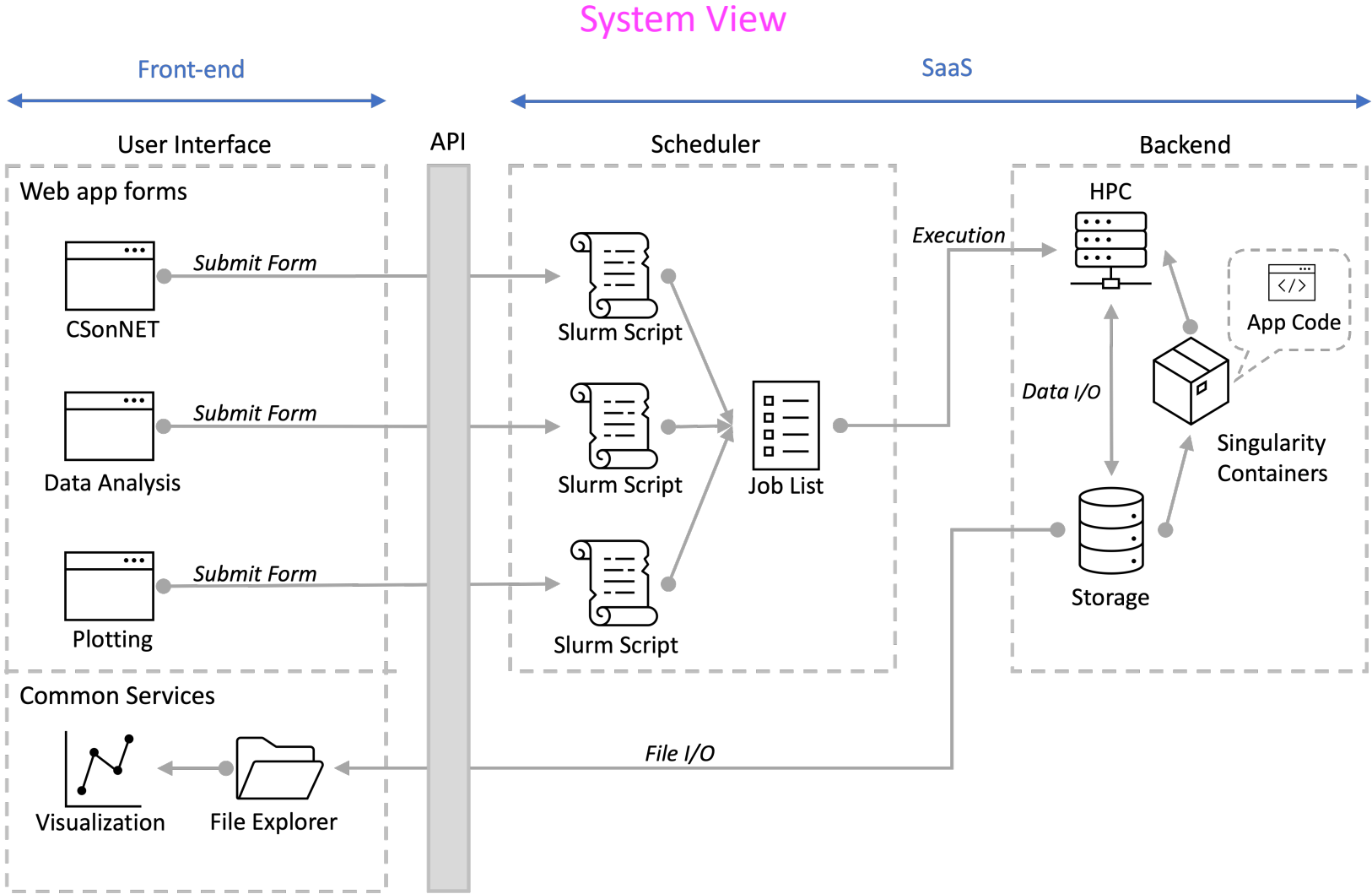
STRONGLY CONNECTED COMPONENTS	WEAKLY CONNECTED COMPONENTS	NODE IN DEGREE
Num: 1	Num: 1	Average: 27.69096967
Size Smallest: 198	Size Smallest: 198	Max: 100
Num Smallest: 1	Num Smallest Size: 1	Min: 1
Frac Smallest: 1	Frac Smallest: 1	
Size Largest: 198	Size Largest: 198	
Num Largest: 1	Num Largest Size: 1	
Frac Largest: 1	Frac Largest: 1	

NODE OUT DEGREE	SCORE	NODE BETWEENNESS CENTRALITY
Average: 27.69096967	Smallest: 0	Average: 121.65151515
Max: 100	Num Nodes Smallest: 198	Max: 2916.2901040932
Min: 1	Frac Nodes Smallest: 1	Min: 0
	Largest: 29	
	Num Nodes Largest: 30	
	Frac Nodes Largest: 0.1515151515	
	Kcore Size 90% Nodes: 17	

EDGE BETWEENNESS CENTRALITY	NODE AUTHORITY SCORE HITS	NODE EIGENVECTOR CENTRALITY
Average: 33.741310721	Average: 0.0026719064	Average: 0.0026722193
Max: 664.846335133	Max: 0.2104085663	Max: 0.2104080754
Min: 2	Min: 0.0000005138	Min: 0.0000005138

NODE HUB SCORE HITS	NODE PAGE RANK
Average: 0.0058719064	Average: 0.0026550261
Max: 0.2104085663	Max: 0.0465792351
Min: 0.0000005138	Min: 0.0009289792

Network Simulation as a Service (NetSimS)

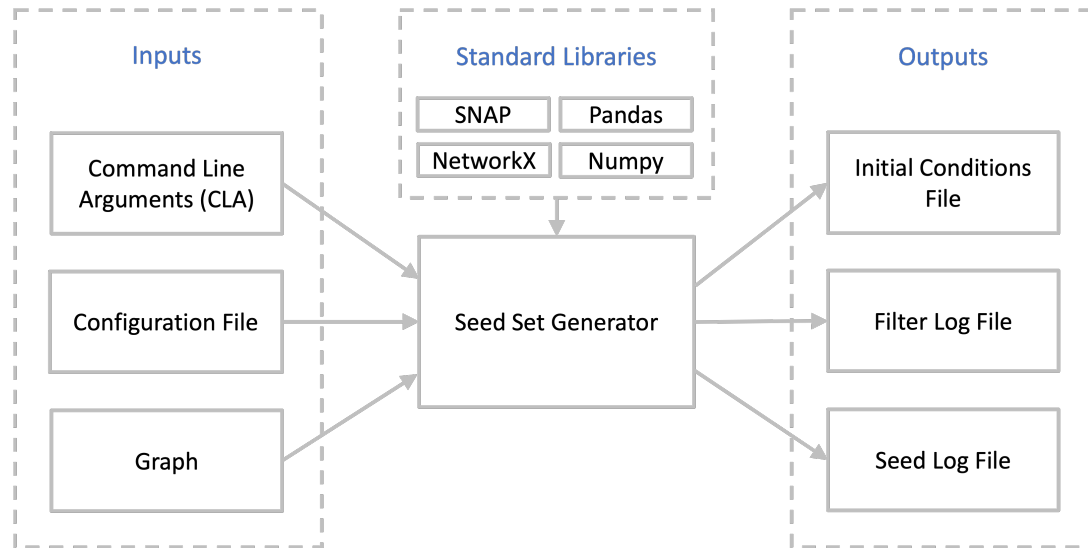


Web Interface

The screenshot displays the Net.Science v1.2.3 dev web interface. The top navigation bar includes 'Home', 'About', 'Files', 'My Jobs', and 'Chris'. The main content area is divided into two columns. The left column, titled 'Input', contains a text input for 'input_file' with a 'SELECT FILE' button and a file path. Below this is the 'Dynamics Model' section, which includes a 'Behaviour Model' dropdown set to 'SEIR', a 'Sub Model' dropdown set to 'stochastic exposed fixed infectious', and three numerical input fields for 'Edge probability', 'Exposed transition probability', and 'Infectious duration'. The right column, titled 'Stochasticity', contains a 'Seed' input field set to '0'. Below this is the 'Composition Of Simulation' section, which includes 'Simulation Timing' with 'Iterations' and 'Time Steps' input fields, and 'Initial Conditions (Seeding)' with a 'Seeding Method' dropdown set to 'Custom', 'Number Nodes' and 'State' input fields, and a 'Node Selection Method' dropdown. At the bottom of the right column, there are two property configuration boxes. The first box is for 'Degree' with 'Ordering' and 'Weight' dropdowns, and 'Min', 'Max' input fields. The second box is for 'Clustering Coefficient' with 'Ordering' and 'Weight' dropdowns, and 'Min', 'Max' input fields. A blue plus sign button is located between the two property boxes.

- 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



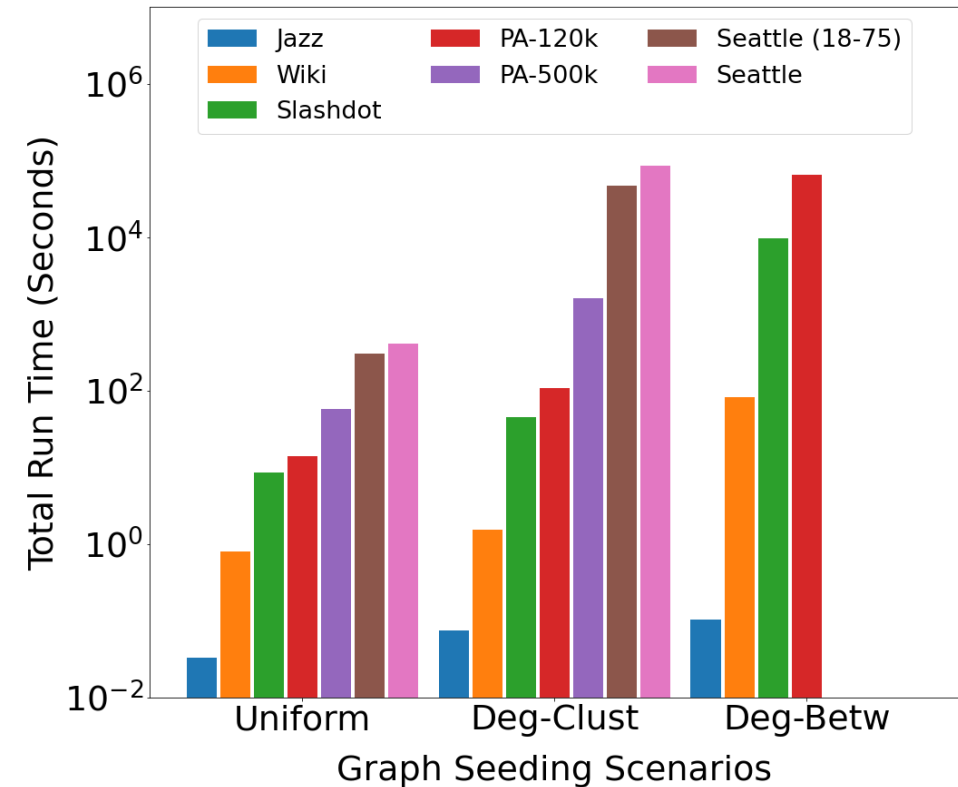
```
seed.config.01
1 #graph_metric,sort_method,metric_min,metric_max,metric_weight
2 outdeg,high,3,4,1
3 betwcnt,high,0.2,6,1
4 clustcf,high,0.4,0.8,1
5
```

- 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.

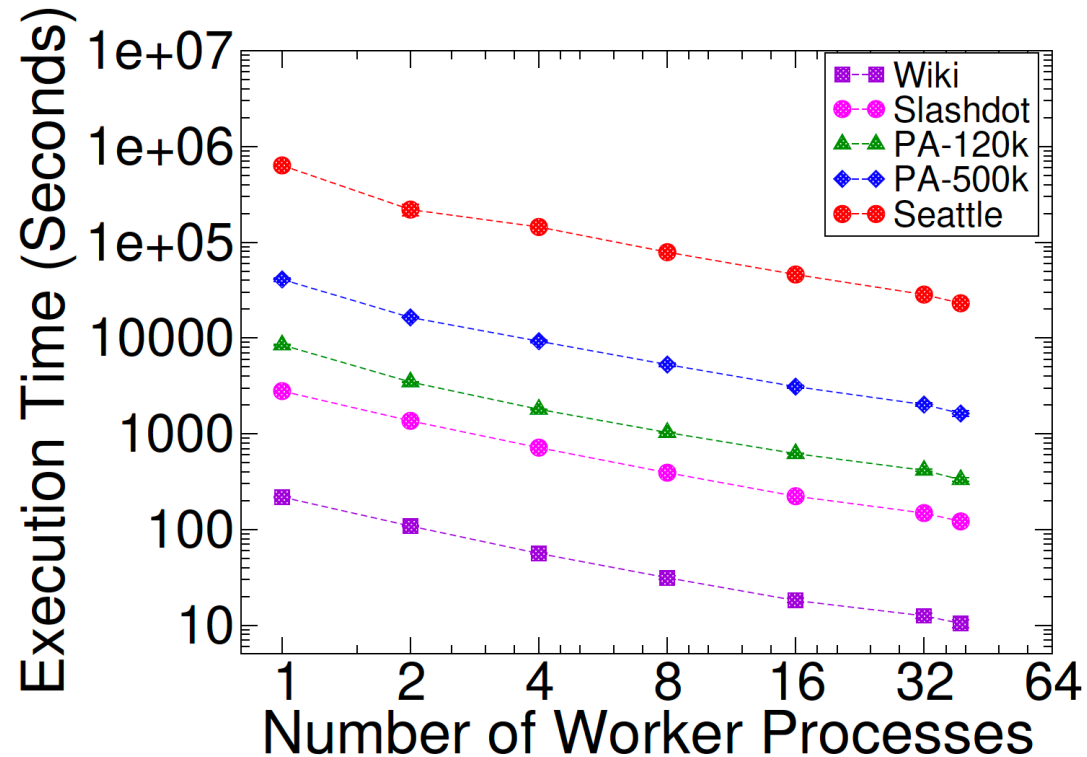
Performance Evaluation – Network Initialization

Network	Structural Properties				
	<i>Num. Nodes</i>	<i>Num. Edges</i>	<i>Avg. Degree</i>	<i>Max. Degree</i>	<i>Largest k-core</i>
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

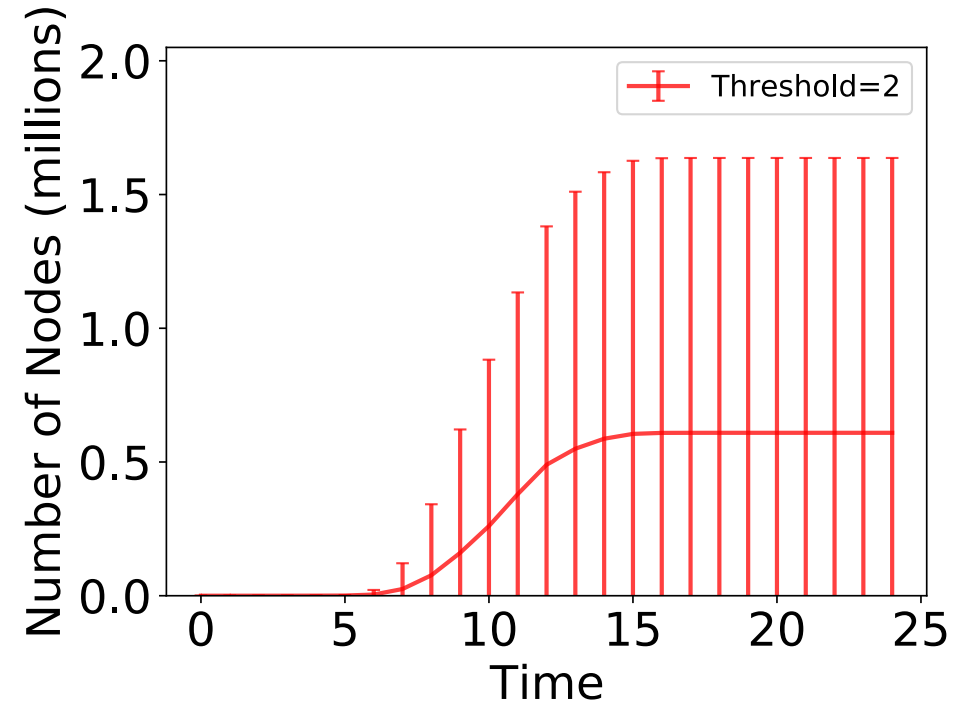
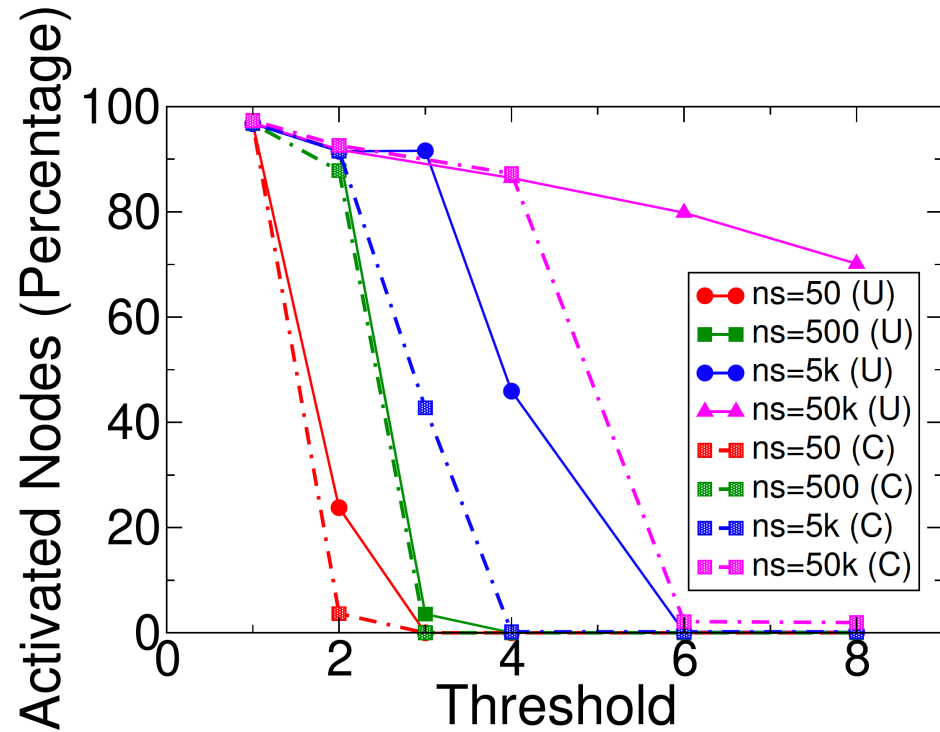


Strong Scaling Study – Agent-Based Simulation



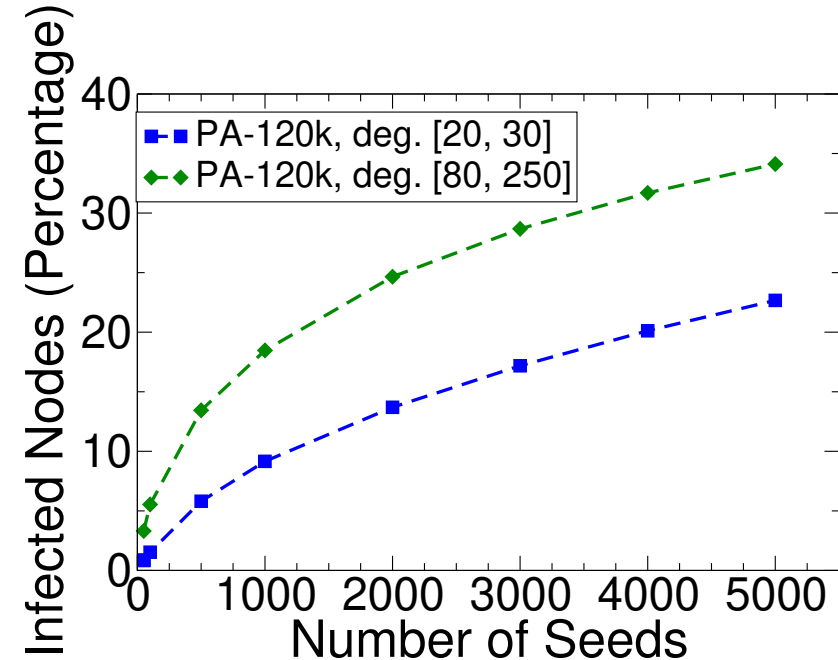
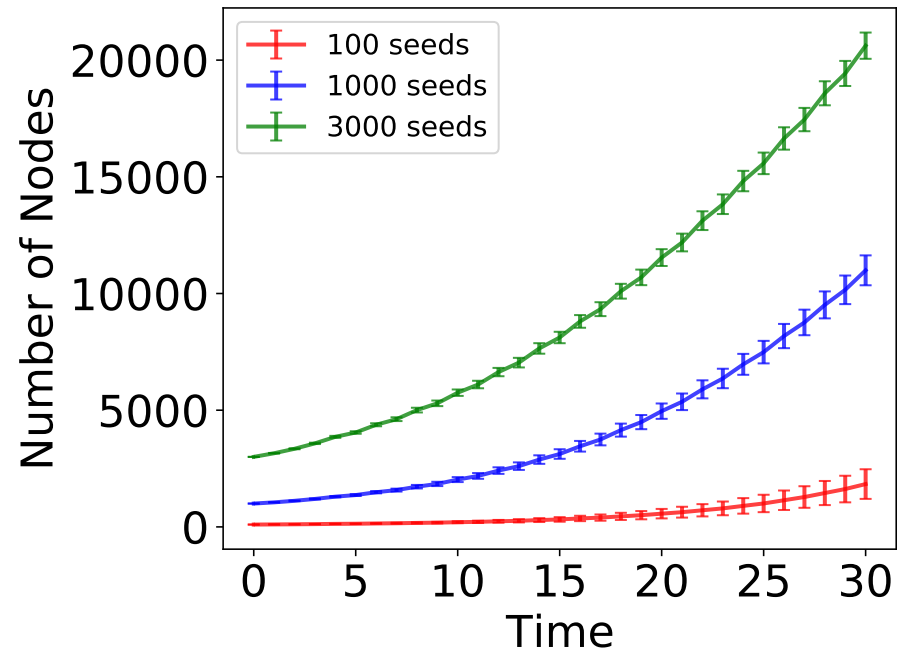
- Uses SIR model
 - Transmission probability: 0.004
 - Infectious duration: 4 days
- Y-Axis: total time to run 100 iterations of 100 timesteps
- 300 nodes initialized as infected (I).
- The system exhibits strong scaling.

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 Symphony [North et al., 2013]
 - NetLogo [Wilensky, 2011]
- Geared towards parallel and high-performance computing
 - Repast HPC [Collier et al., 2013]
 - μ sik [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>