




Synthetic Geosocial Network Generation

Ketevan Gallagher, Taylor Anderson,
Andrew Crooks, Andreas Züfle



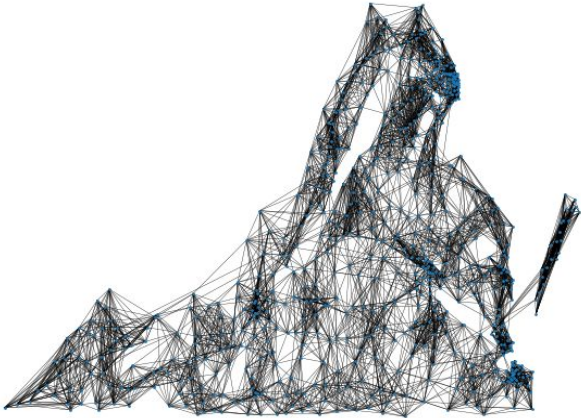
Background: Synthetic Social Networks

- ◎ Representations of people connect to one another
- ◎ Real world data difficult to find
- ◎ Three classical models: Erdős-Rényi, Barabási-Albert, Watts-Strogatz

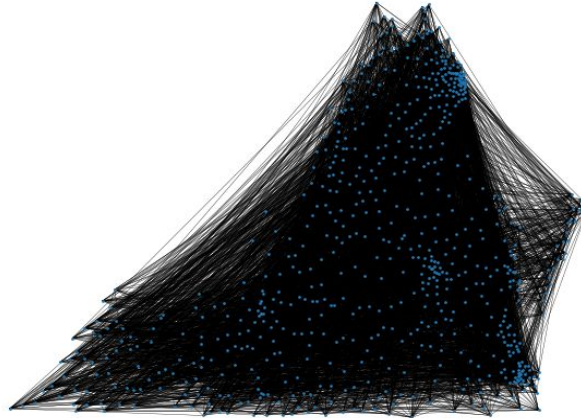
Background: Spatial Social Network Models

- © Real-world social networks exhibit spatial homophily
- © Incorporate spatial distances between nodes

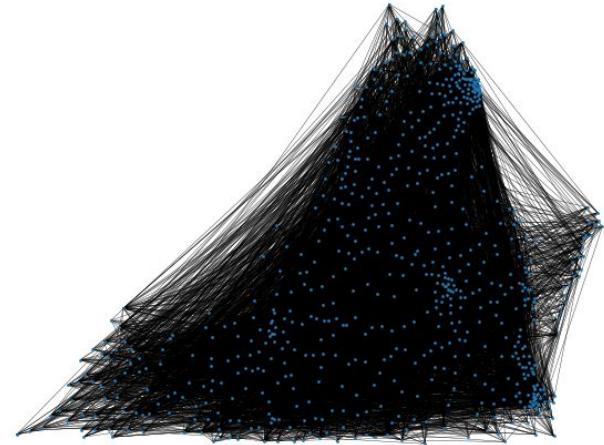
Ground Truth



Erdős-Rényi



Barabási-Albert



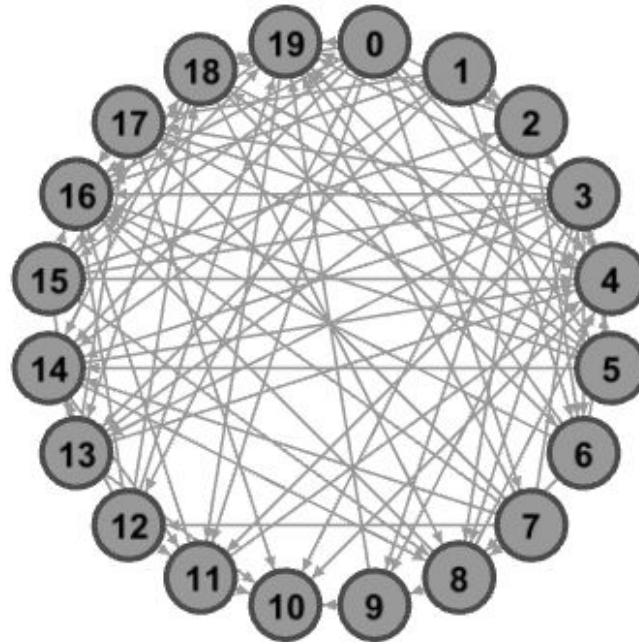
Goal

Generate social networks that exhibit spatial autocorrelation



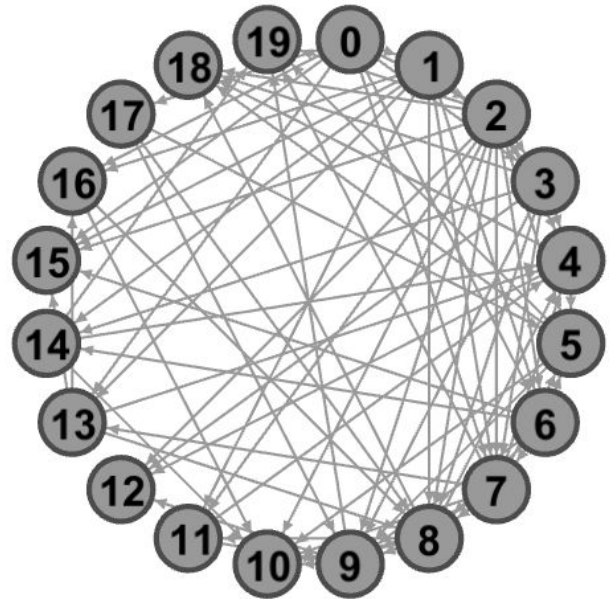
Classical Erdős-Rényi Model

© Nodes connected randomly



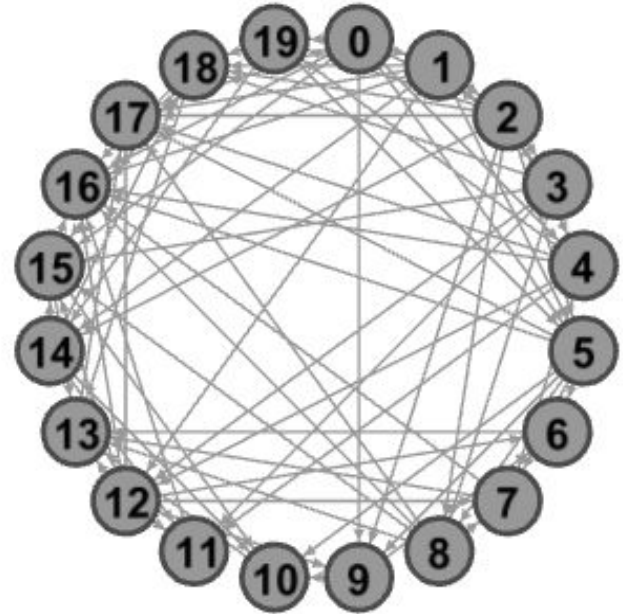
Classical Barabási-Albert Model

- ◎ Scale free network growth
- ◎ Preferential attachment
- ◎ Nodes are added iteratively



Classical Watt-Strogatz Model

- ⦿ Nodes connect to k nearest neighbors, and then have chance to rewire
- ⦿ Exhibits small-world properties



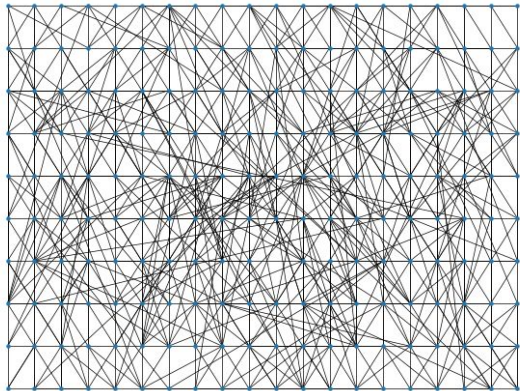
Existing Work

- © Generating and Analyzing Spatial Social Networks[1]
 - Assumes that nodes are in a uniform lattice
 - When real-world location data is incorporated, large variety of distances leads to sparse graphs or long runtimes

[1] M. Alizadeh, C. Cioffi-Revilla, and A. Crooks. Generating and analyzing spatial social networks. *Computational and Mathematical Organization Theory*, 23:362–390, 2017.

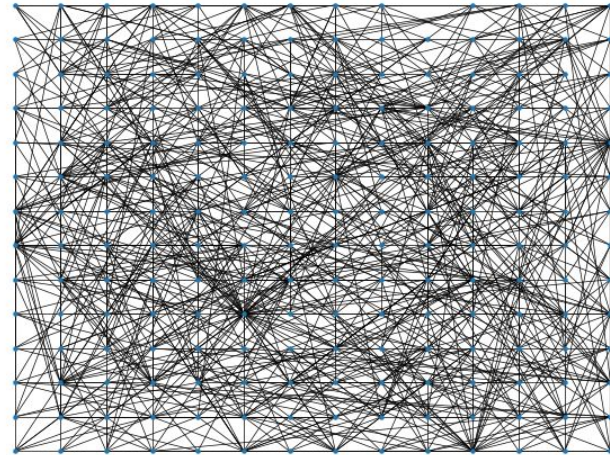
Spatial Erdős-Rényi Model

- © $p(d) = Cd^{-\alpha}$
- d : distance between nodes
 - C : normalizing coefficient
 - α : distance-decay exponent



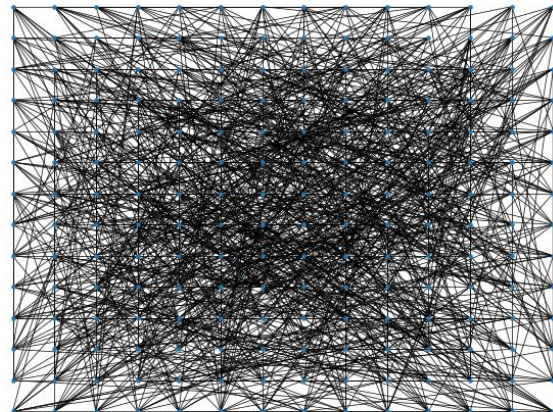
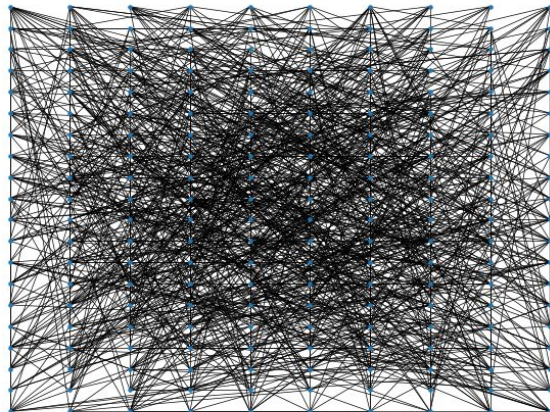
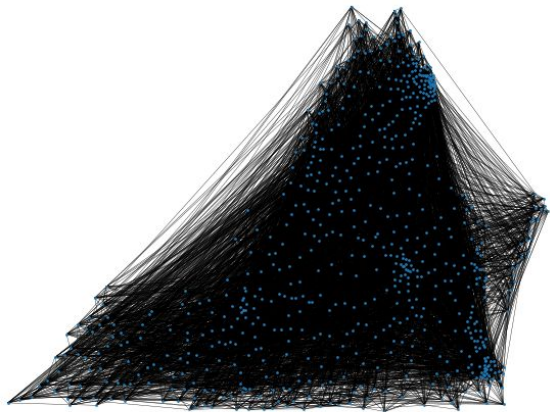
Spatial Barabási-Albert Model

- ◎ Starts with clique of m nodes
- ◎ $p(d) = kd^{-\alpha}$
 - k : degree of node being considered for connection
- ◎ Long runtime



Spatial Watts-Strogatz Model

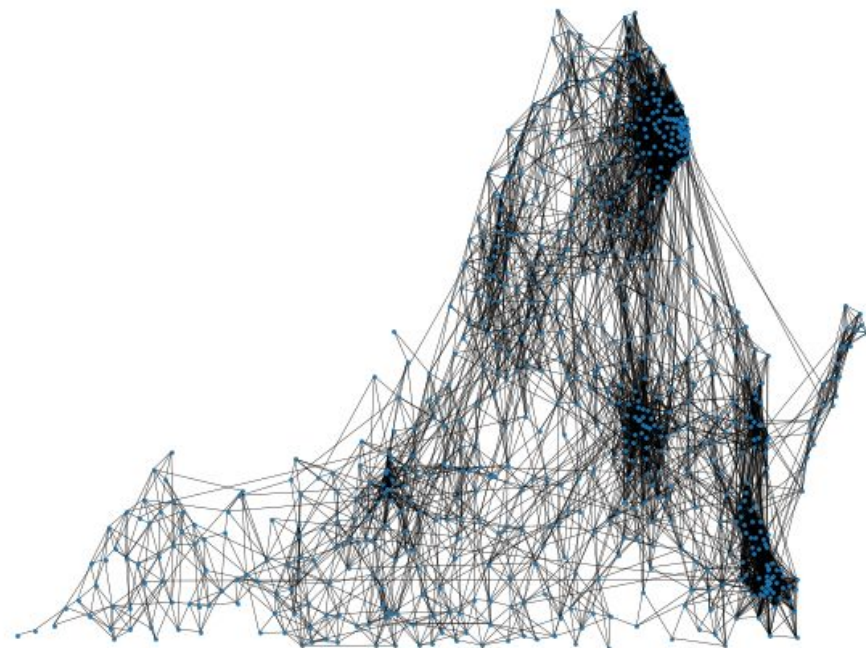
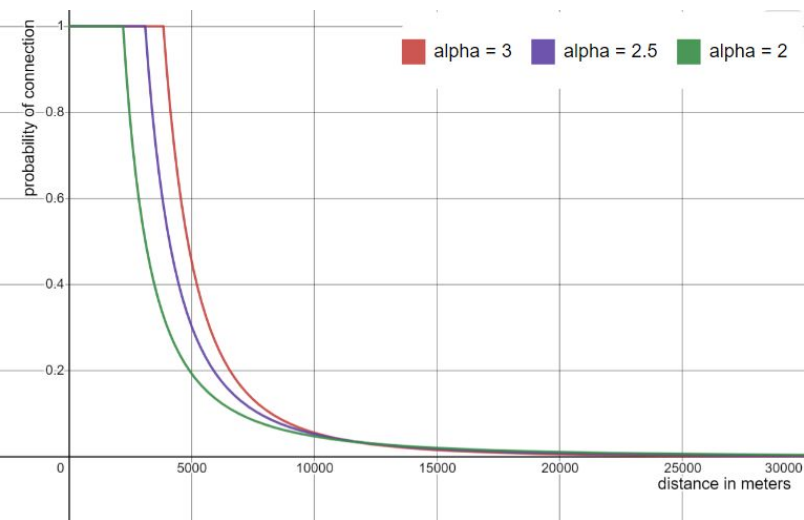
- © Two versions: initially connect based on distance, and rewire based on distance
- © Due to node IDs, random



Geosocial Erdős-Rényi Model

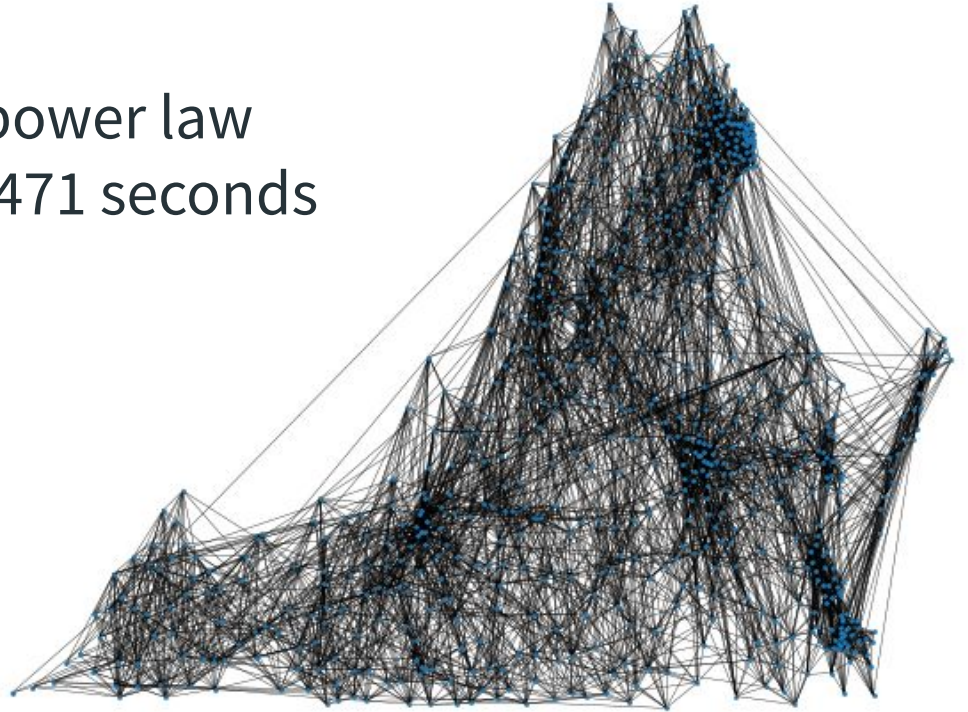
© Implemented scaling factor

- $p(d) = \min(1, sd^{-\alpha})$
- $\alpha = 3$, average degree 20



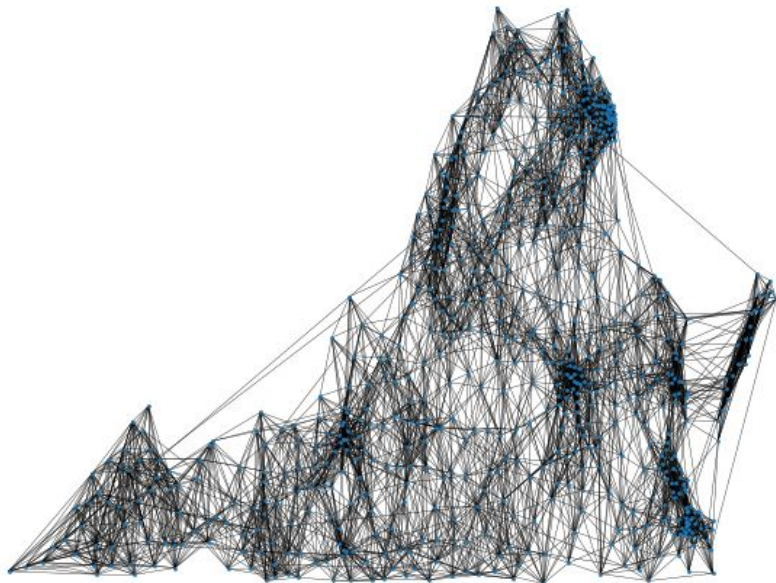
Geosocial Barabási-Albert Model

- © Node processing order
- © Normalize weights from power law
- © 103.1283 seconds vs 49.5471 seconds



Geosocial Watts-Strogatz Model

- © Nearest neighbors defined as nodes that are closest spatially

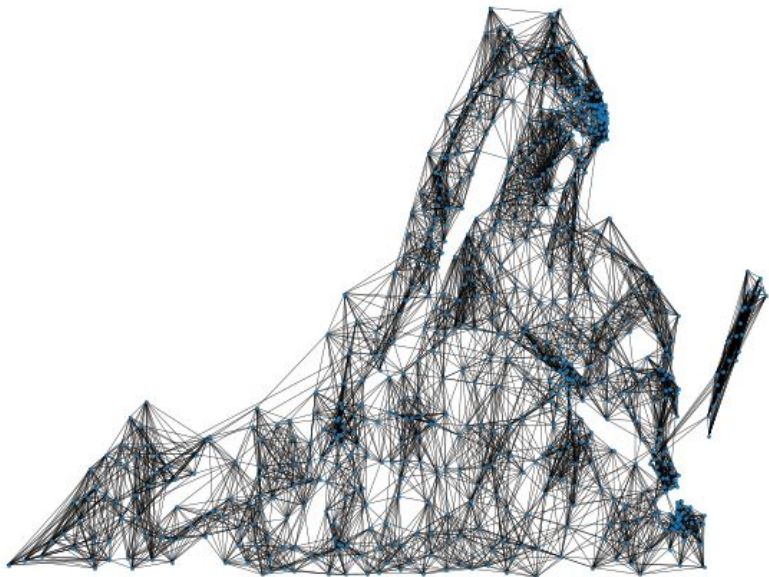


Datasets

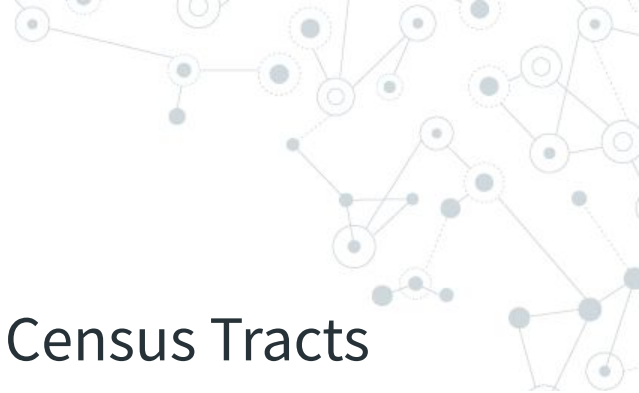
- ◎ Virginia ZIP code data
 - From Facebook
 - Each ZIP code given Social Connectedness Index
- ◎ Fairfax County Census Tracts
 - Mobility data
 - Highest estimated population flows for 1/4/2020
- ◎ Average degree 20, 14, $\alpha = 3$

Ground Truth Graphs

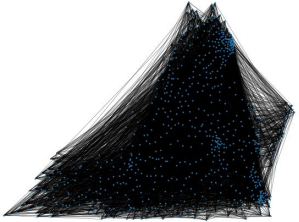
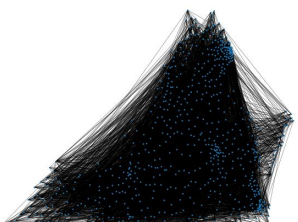
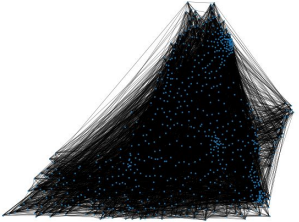
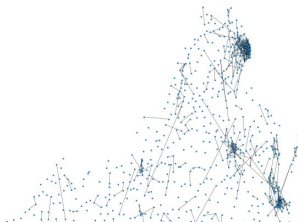
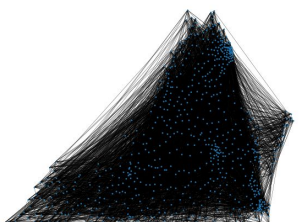
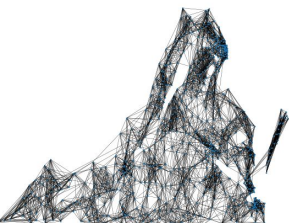
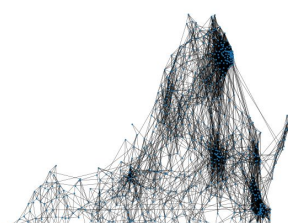
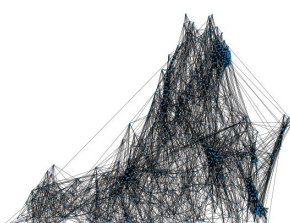
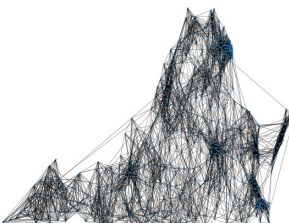
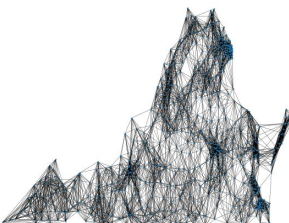
Virginia ZIP code



Fairfax Census Tracts

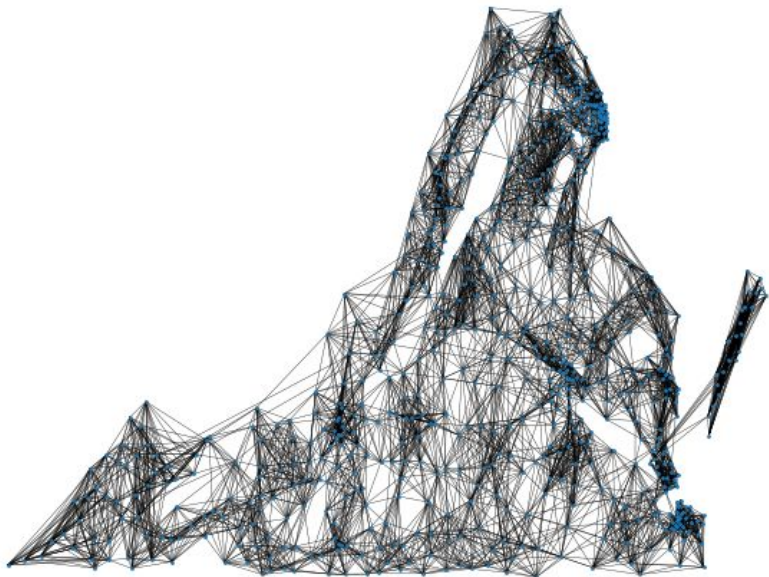


Qualitative Evaluation (Virginia)

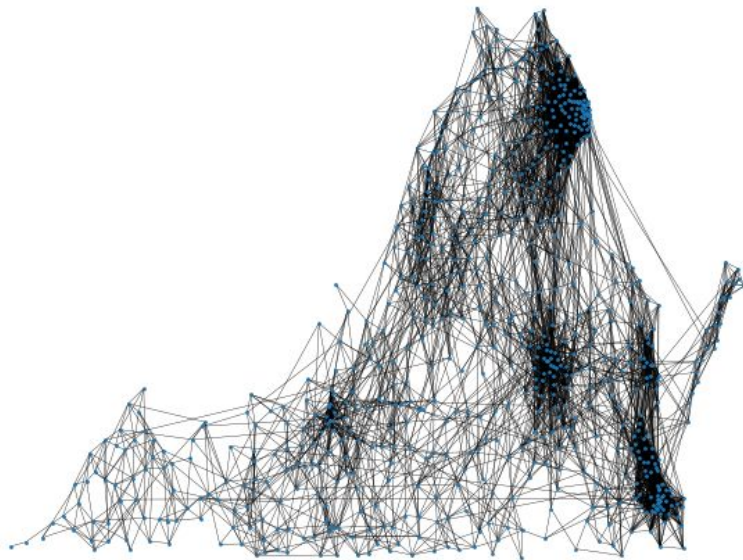
Classical ER	Classical BA	Classical WS	Spatial ER	Spatial WS
 A network visualization showing a dense, triangular structure with many overlapping edges, representing a classical Erdős-Rényi model.	 A network visualization showing a dense, triangular structure with many overlapping edges, representing a classical Barabási-Albert model.	 A network visualization showing a dense, triangular structure with many overlapping edges, representing a classical Watts-Strogatz model.	 A network visualization showing a sparse, triangular structure with many overlapping edges, representing a spatial Erdős-Rényi model.	 A network visualization showing a dense, triangular structure with many overlapping edges, representing a spatial Watts-Strogatz model.
Ground Truth	Geosocial ER	Geosocial BA	Geosocial WS	KNN
 A network visualization showing a complex, irregular structure with many overlapping edges, representing the ground truth network.	 A network visualization showing a complex, irregular structure with many overlapping edges, representing a geosocial Erdős-Rényi model.	 A network visualization showing a complex, irregular structure with many overlapping edges, representing a geosocial Barabási-Albert model.	 A network visualization showing a complex, irregular structure with many overlapping edges, representing a geosocial Watts-Strogatz model.	 A network visualization showing a complex, irregular structure with many overlapping edges, representing a K-Nearest Neighbors model.

Qualitative Evaluation (Virginia)

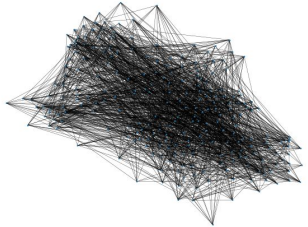
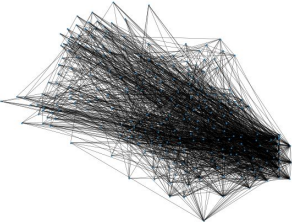
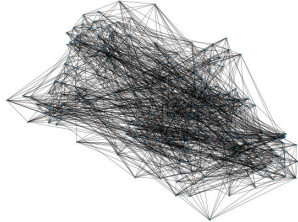
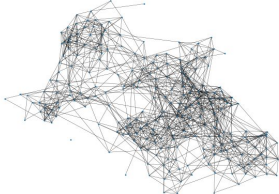
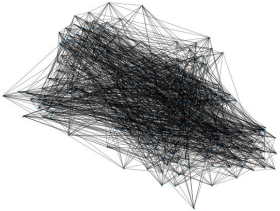
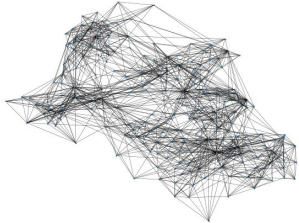

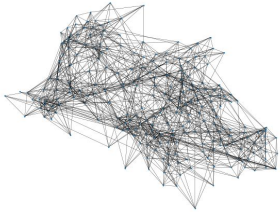
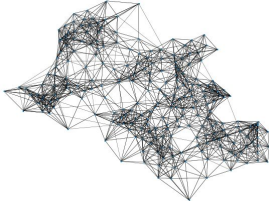
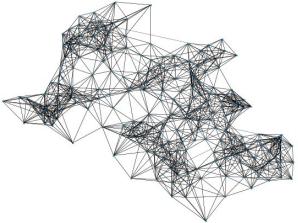
Ground Truth



Geosocial Erdős-Rényi



Qualitative Evaluation (Fairfax County)

Classical ER	Classical BA	Classical WS	Spatial ER	Spatial WS
 A dense, highly interconnected network graph with many overlapping edges, appearing as a solid black mass.	 A network graph with a dense core and many edges extending outwards, resembling a fan or a star-like structure.	 A network graph with a dense core and many edges extending outwards, similar to Classical BA but with a slightly different internal structure.	 A network graph with a sparse, irregular structure, showing clear clusters and gaps.	 A network graph with a dense, irregular structure, showing clear clusters and gaps, similar to Spatial ER.
Ground Truth	Geosocial ER	Geosocial BA	Geosocial WS	KNN
 A network graph with a sparse, irregular structure, showing clear clusters and gaps, representing the ground truth.	 A network graph with a sparse, irregular structure, showing clear clusters and gaps, similar to Ground Truth.	 A network graph with a sparse, irregular structure, showing clear clusters and gaps, similar to Ground Truth.	 A network graph with a sparse, irregular structure, showing clear clusters and gaps, similar to Ground Truth.	 A network graph with a sparse, irregular structure, showing clear clusters and gaps, similar to Ground Truth.

Virginia ZIP Code Statistics

	Ground Truth	Classic ER	Classic BA	Classic WS	Spatial ER	Spatial WS	Geosocial ER	Geosocial BA	Geosocial WS	KNN
Avg. Degree	19.59	19.78	19.84	19.77	2.94	20.00	20.05	19.73	20.00	20.00
Std. Dev. Degree	3.40	3.83	18.49	2.70	4.29	1.37	14.70	5.06	4.36	4.50
Max. Degree	32.00	32.60	163.25	29.40	22.55	25.00	69.40	37.50	33.90	34.00
Radius of Gyration	55292	513022	489469	509785	12539	510990	87529	133198	68856	37152
Std. Dev Radius of Gyration	51072	132646	134790	132887	23579	130288	67073	107183	64809	15330
Avg. Length of Edge	20934	206714	208432	201354	7739	188125	18440	29880	18792	16087
Std. Dev. Length of Edge	21175	166824	161757	168058	14522	167691	25926	45530	22785	12822
# Triangles	26357	1251	6742	4717	1258	22709	38421	13283	23522	27596
Jaccard Index	1	0.0078	0.0118	0.0136	0.1167	0.0266	0.3615	0.1660	0.2367	0.2560

Fairfax Census Tract Statistics

	Ground Truth	Classic ER	Classic BA	Classic WS	Spatial ER	Spatial WS	Geosocial ER	Geosocial BA	Geosocial WS	KNN
Avg. Degree	13.58	13.62	13.77	13.69	12.48	14.00	13.99	13.63	14.00	14.00
Std. Dev. Degree	8.69	3.12	11.52	2.24	4.68	1.35	4.96	4.07	3.11	3.16
Max. Degree	72.00	22.95	82.75	20.65	24.45	18.35	26.95	26.30	22.70	22.00
Radius of Gyration	11836	29287	29047	26435	9049	28719	9675	12637	7434	4401
Std. Dev Radius of Gyration	6935	6843	8267	7847	5597	7770	5505	8532	4814	970
Avg. Length of Edge	4109	13852	14036	9506	2731	12338	2842	3594	2462	2068
Std. Dev. Length of Edge	3968	9930	10414	9008	3018	9951	3049	4587	2345	1333
# Triangles	3074	401	1606	900	2491	3251	3176	2075	3687	4597
Jaccard Index	1	0.0163	0.0317	0.0643	0.2814	0.0295	0.2861	0.1179	0.1880	0.2038

Conclusions and Future Work

- © Geosocial Erdős-Rényi performs well when compared to other geosocial and spatial graphs
- © Topographical distances instead of Euclidean distances

Acknowledgements

This work is supported by National Science Foundation Grant #2109647 titled “Data-Driven Modeling to Improve Understanding of Human Behavior, Mobility, and Disease Spread” and by the Aspiring Scientists Summer Internship Program (ASSIP) at George Mason University.

Thank you to Dr. Andrew Crooks, Dr. Taylor Anderson, and Dr. Andreas Züfle

