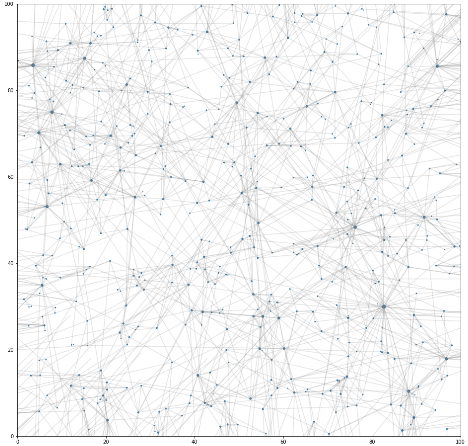
Spatial Networks

Generalized Network Models and Quantifying Spatial Embedding Strength

October 15, 2019

Introduction

What is a network model?
What are the uses of network models?
When to consider (or not) spatial generalizations?



"Simple" Spatial Generalization

How do we turn a non-spatial model into a spatial one?

Inclusion of a deterrence function:

$$h(r_{i,j}) = r_{i,j}^{-\beta}, \beta > 0$$

$r_{i,j}$ is the distance between nodes v_i and v_j
Use deterrence function to decrease likelihood of nodes that are far away to be adjacentskyp

Network Models

Barabsi-Albert (preferential attachment (PA)) model

Power-law degree distribution, "rich get richer" scheme

Model description:

- Start with a "seed" network (e.g. 10-clique)
- Evolve the network: at each time step add a new node. New node creates m edges to nodes with probability proportional to the degree k_j of existing node k_j
- Continue until T nodes added

Probability of forming edge from node v_t at time t to existing node v_j:

$$p(v_t, v_j) = \frac{k_j}{\sum_i k_i}$$

Network Models

Spatial preferential attachment (SPA) model

Differences:

- Network is embedded in $[0,1] \times [0,1]$ 2-dimensional space. Every new node is assigned a position uniformly at random
- Edges are created proportional to degree k_j of existing node and distance $r_{t,j}$.

New probability:

$$p(v_t, v_j) = \frac{k_j}{\sum_i k_i} h(r_{t,j}) \, .$$

Network Characteristics

- Mean local clustering coefficient: $c_i = \frac{2T(v_i)}{k_i(k_i-1)}$
- Mean geodesic distance: $L = \frac{\sum_{i \neq j} d(v_i, v_j)}{n(n-1)}$
- Mean edge length
- Degree assortativity: $r = \frac{\sum_{i,j}(A_{i,j} - k_i k_j / 2m) k_i k_j}{\sum_{i,j}(k_i \delta_{i,j} - k_i k_j / 2m) k_i k_j}$

Network Models

SPA Characteristics

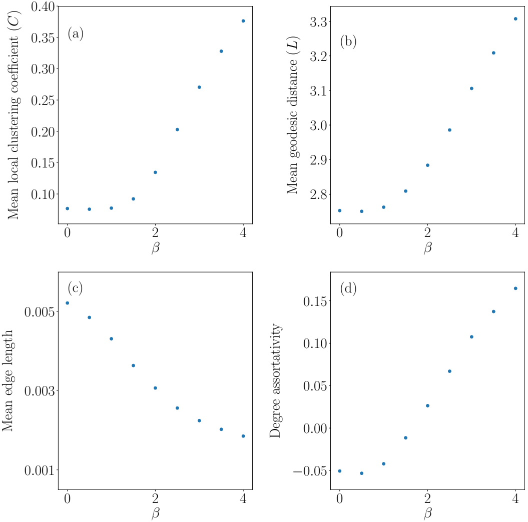


Figure: Metrics for $n = 10000$

Network Models

SPA Characteristics

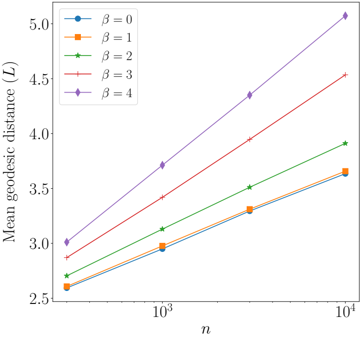


Figure: log-log n vs. mean geodesic distance

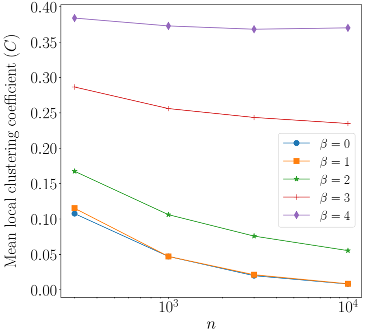


Figure: log-log n vs. clustering

Network Models 2

Inherent fitness model

Nodes have "inherent fitness" - propensity to gain edges

Model description:

- n total nodes are each given a fitness value from a distribution, e.g. $f(w) = \lambda e^{-\lambda w}\,, \quad w \geq 0$
- Each pair of nodes is considered and an edge formed between them depending on function g, e.g. Heavistep $g(v_i, v_j) = w_i + w_j \geq \theta$, θ threshold parameter

Differences:

- Network is embedded in $[0,1] \times [0,1]$ 2-dimensional space. Every new node is assigned a position uniformly at random
- New probability of forming edges between pairs of nodes $g(v_i, v_j) = (w_i + w_j)h(r_{i,j}) \geq \theta$

Network Models 2

GF Model

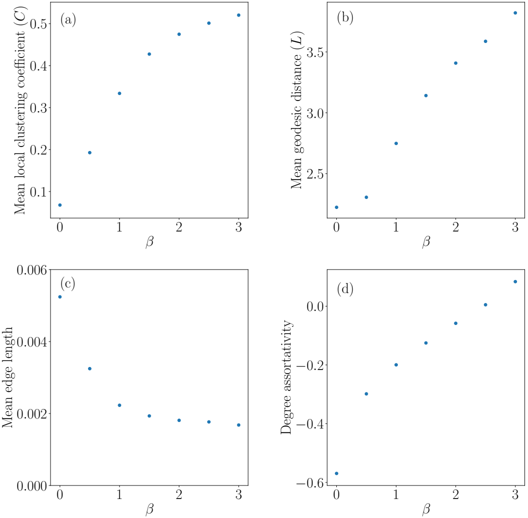
Also consider using a different fitness distribution and edge forming probability:

$$w_i \sim N(0,1)$$

$$g(v_i, v_j) = |w_i - w_j| h(r) \geq \theta$$

Network Models

GF Characteristics



Network Models

GF Characteristics

Closeness centrality:

$$C(v_i) = \frac{n-1}{\sum_{i \neq j} d(v_i, v_j)}$$

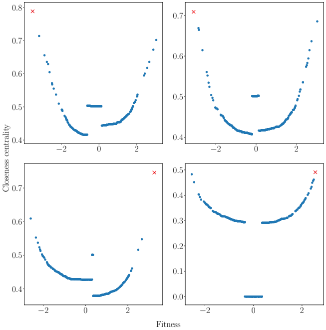


Figure: Closeness centrality for $\beta = 0$

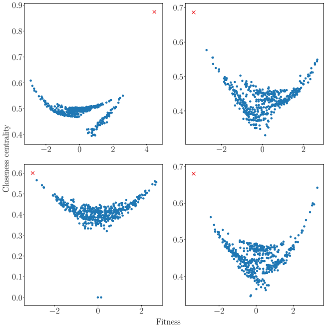


Figure: Closeness centrality for $\beta = 0.5$

One of the most classic null models - preserve degree sequence

Model description:

- Fix a degree sequence (often taken from an existing network)
- Assign edge stubs based on degree sequence, so each node is guaranteed to have
- Match edge stubs uniformly at random

Network Models 3

Configuration Model

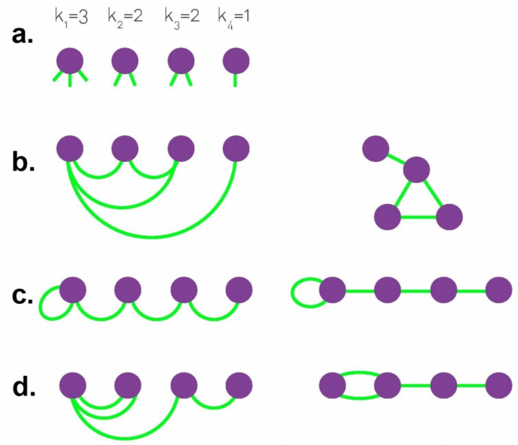


Figure: Taken from Albert-Lszl Barabsi - http://networksciencebook.com/

Differences:

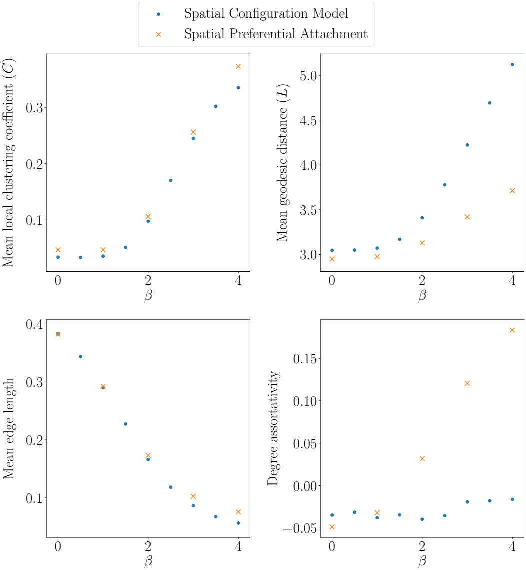
- n nodes given random location in $[0,1] \times [0,1]$ with stubs from degree sequence
- Stub matching is proportional to distance from nodes, i.e.

$$p(v_i, v_j, t) = \frac{u(v_j, t)h(r_{i,j})}{\sum_{l \neq i} u(v_l, t)h(r_{i,l})}$$

where $u(v_j, t)$ is the number of stubs that v_j has at time t

Network Models 3

Spatial Configuration Characteristics



Network Models 3

Concluding Thoughts

Generally spatial models may be useful as null models or for creating families of networks to model systems

As β increases, the importance of distances between nodes increases, but can we quantify this?

Spatial Strength

Idea

Quantity to measure "how strongly spatial embedding effects network structure"

Idea: have a centrality that captures whether nodes are adjacent due to spatial embedding, or due to other network structure reasons

Spatial Strength

Definition

$N(v_i)$ is the neighborhood of v_i, then
Normalized mean edge length:

$$L(v_i) = \frac{\sum_{v_j \in N(v_i)} r_{i,j}}{k_i} \frac{1}{\langle L \rangle}$$

Mean neighbor degree:

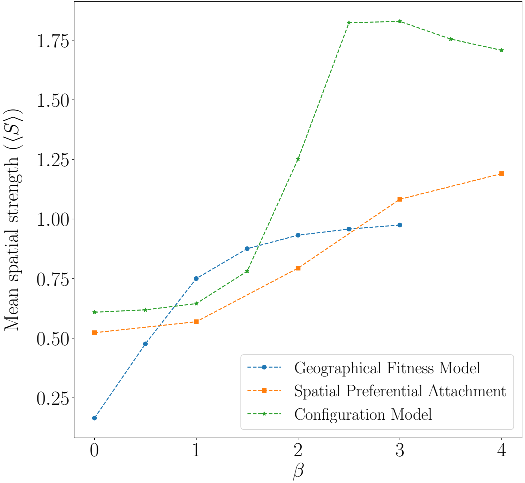
$$K(v_i) = \frac{\sum_{v_j \in N(v_i)} k_j}{k_i} \frac{1}{\langle k \rangle}$$

Finally, spatial strength centrality:

$$S(v_i) := \frac{1}{L(v_i) K(v_i)}$$

Spatial Strength

Spatial strength of models



Spatial Strength

Visual examples

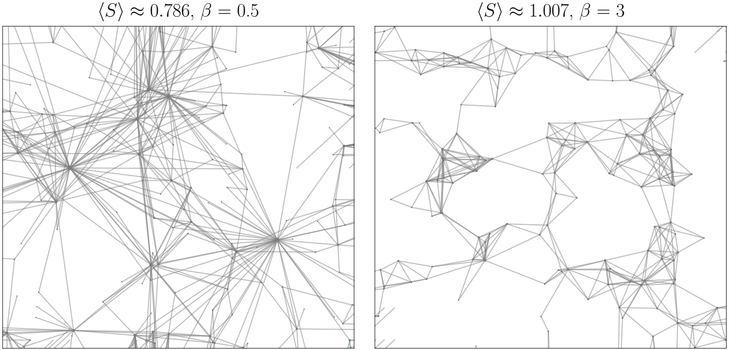


Figure: Examples for GF model

Spatial Strength

Visual examples 2

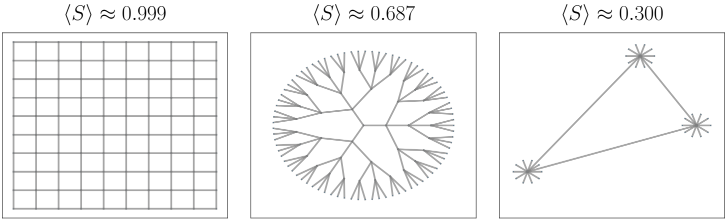


Figure: Toy models

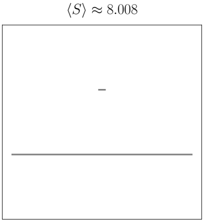


Figure: Showing mean spatial strength is unbounded

Spatial Strength

Mean spatial strength on data

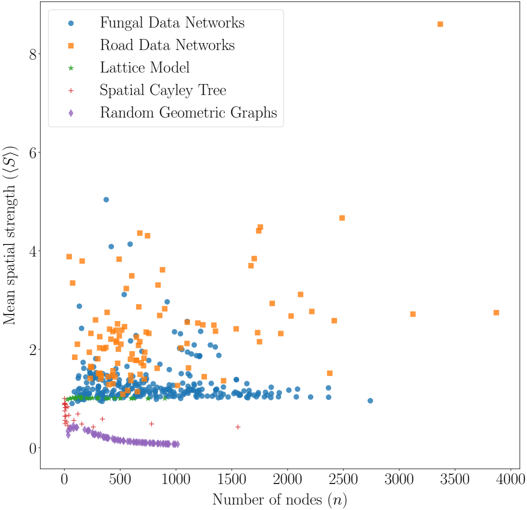


Figure: Mean spatial strengths for different data organized by number of nodes

Other Points

- There are other ways to think about spatial strength
- Spatial strength assumes closer nodes should be attached and doesn't consider underlying topology
- Try to figure out influence of deterrence function on edge formation probability
- Further considerations for spatial configuration model

The end!