Characterizing Patient Mobility

Abstract

People traveling to seek healthcare has been investigated as part of health geographic and social-justice studies in past. These studies have either focused on specific segments of the population or tried to characterize spatial accessibility of healthcare in certain regions. But there has been no study that investigates how people travel within a given entity across county or regional borders to seek healthcare. In this paper we characterize mobility pattern of patients across counties in the state of Texas by encoding the patient movement in a network model and studying its topological properties. We observe that these topological properties capture the geography of the region and are better indicators than the pure geographic distance measures. In doing so we uncover a specific class of networks that differ from the observed workforce-based mobility networks as well as other well know real-world networks. We develop a simple model that can generate directed graphs that mimic the observed patient mobility networks with the specific properties of degrees and clustering.

Keywords: networks, healthcare, mobility

Domain

Healthcare Access

Version 1.0

A GIS Method to Assess Distance Effects on Hospitalizations

By

Ge Lin

RESEARCH PAPER 2002-15

Distance from the Primary Health Center: A GIS Method to Study Geographical Access to Health Care

S. Kohli, K. Sahlén, Å. Sivertun, O. Löfman, E. Trell, and O. Wigertz

Questions

- Are there any global patterns in patient traveling to seek healthcare across a given region?

Questions

- Are there any global patterns in patient traveling to seek healthcare across a given region?
- How much does geography and demography play in deciding where patients travel to seek healthcare?

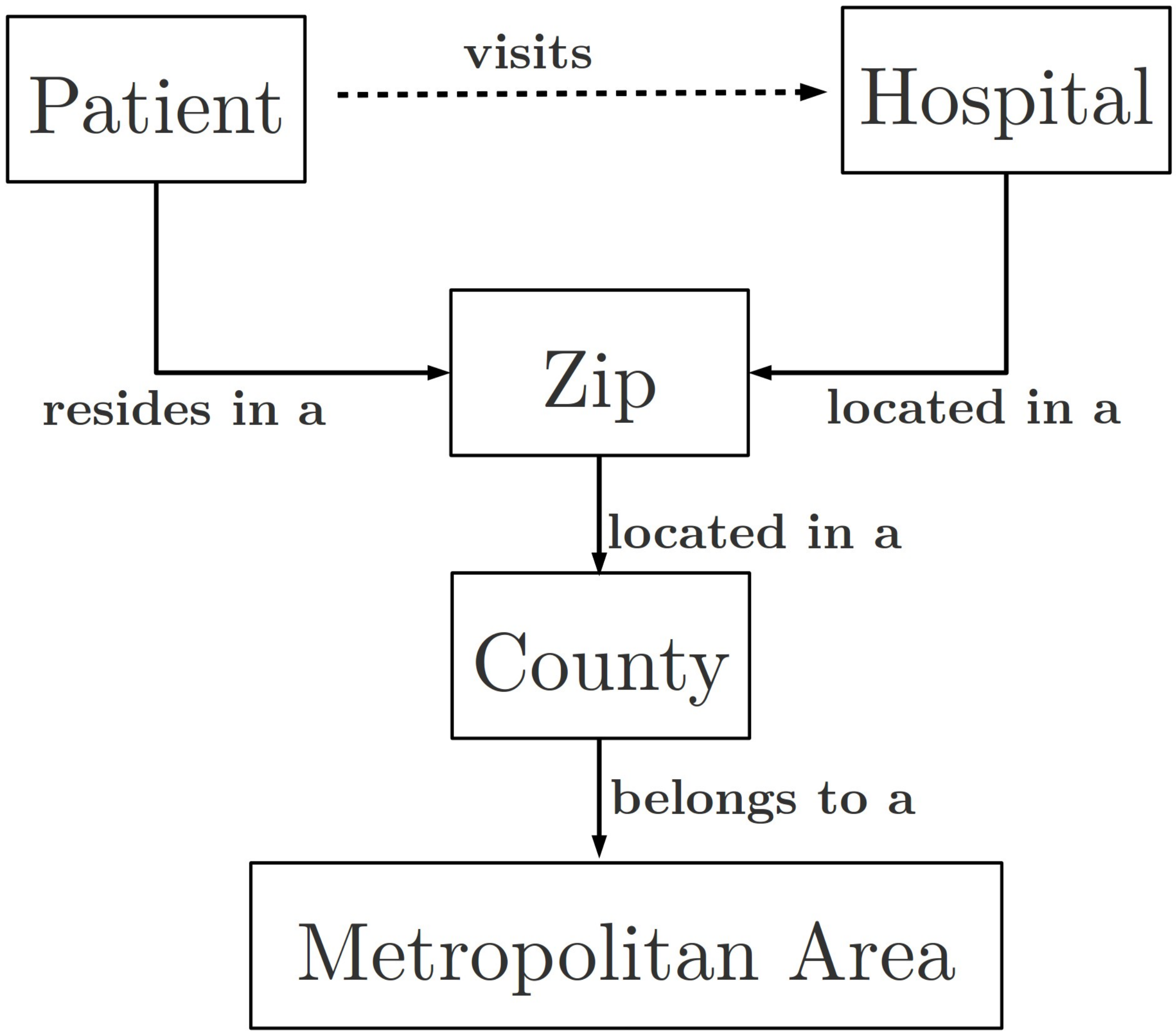
Questions

- Are there any global patterns in patient traveling to seek healthcare across a given region?
- How much does geography and demography play in deciding where patients travel to seek healthcare?
- How does the mobility pattern differ from work force mobility?

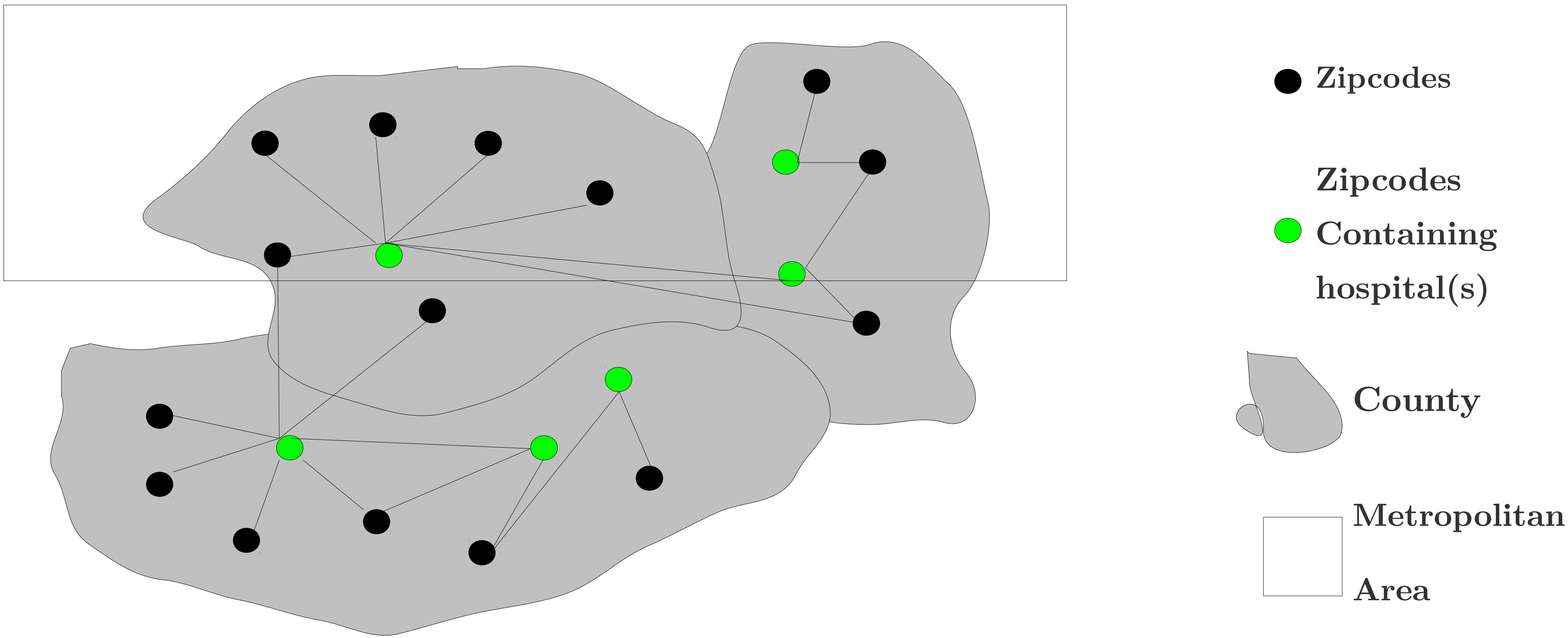
About

- Analysis of patient mobility based on hospital discharge data from 2004-2008 in TX
- Geographical distance based characterization across the state
- Network analysis at three levels -
  - Zip to Zip
  - County to County
  - Metropolitan area to Metro/Micropolitan area
- A simple model

Overview – Entity Model

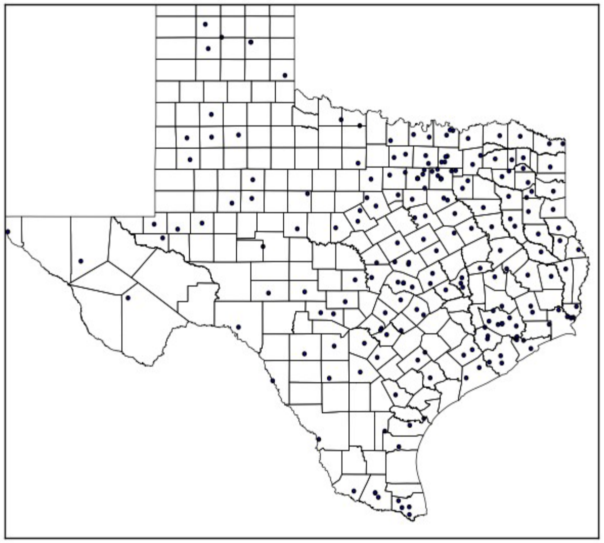


Overview – Entity Model Example

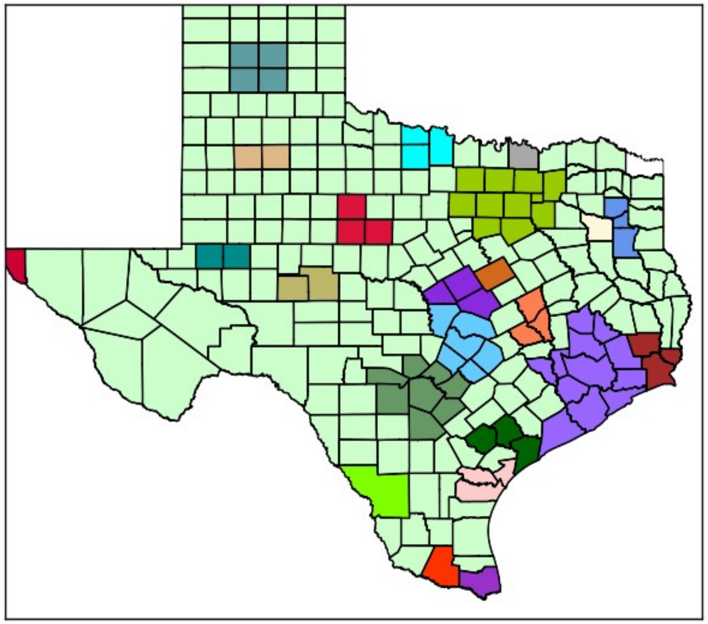


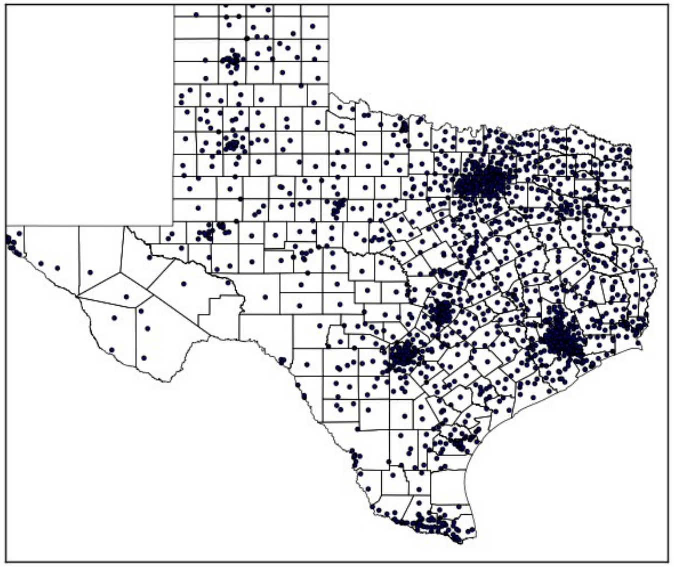
Overview - Geography

Servicing Hospitals



Metropolitan Areas in TX



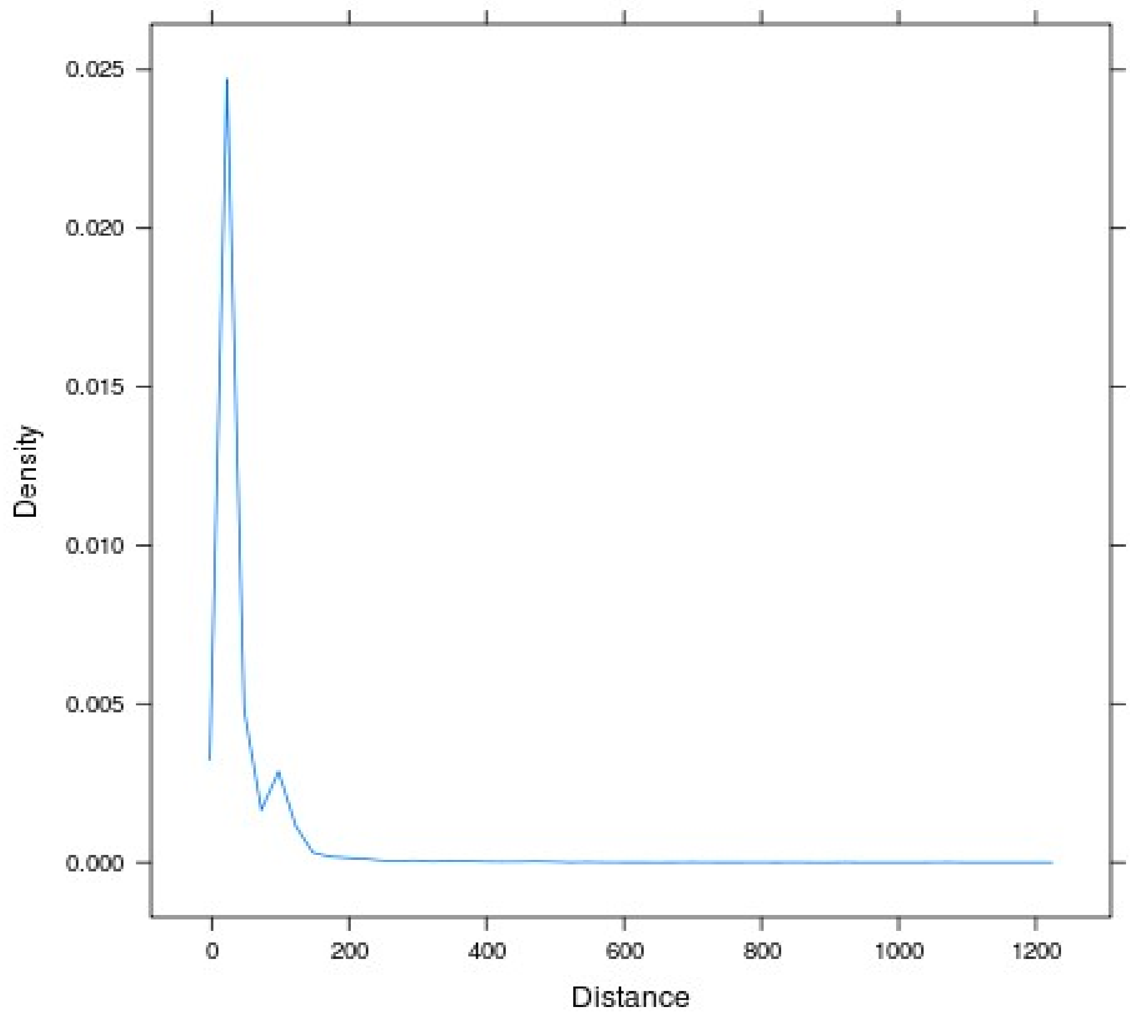


Zips Serviced

Observations

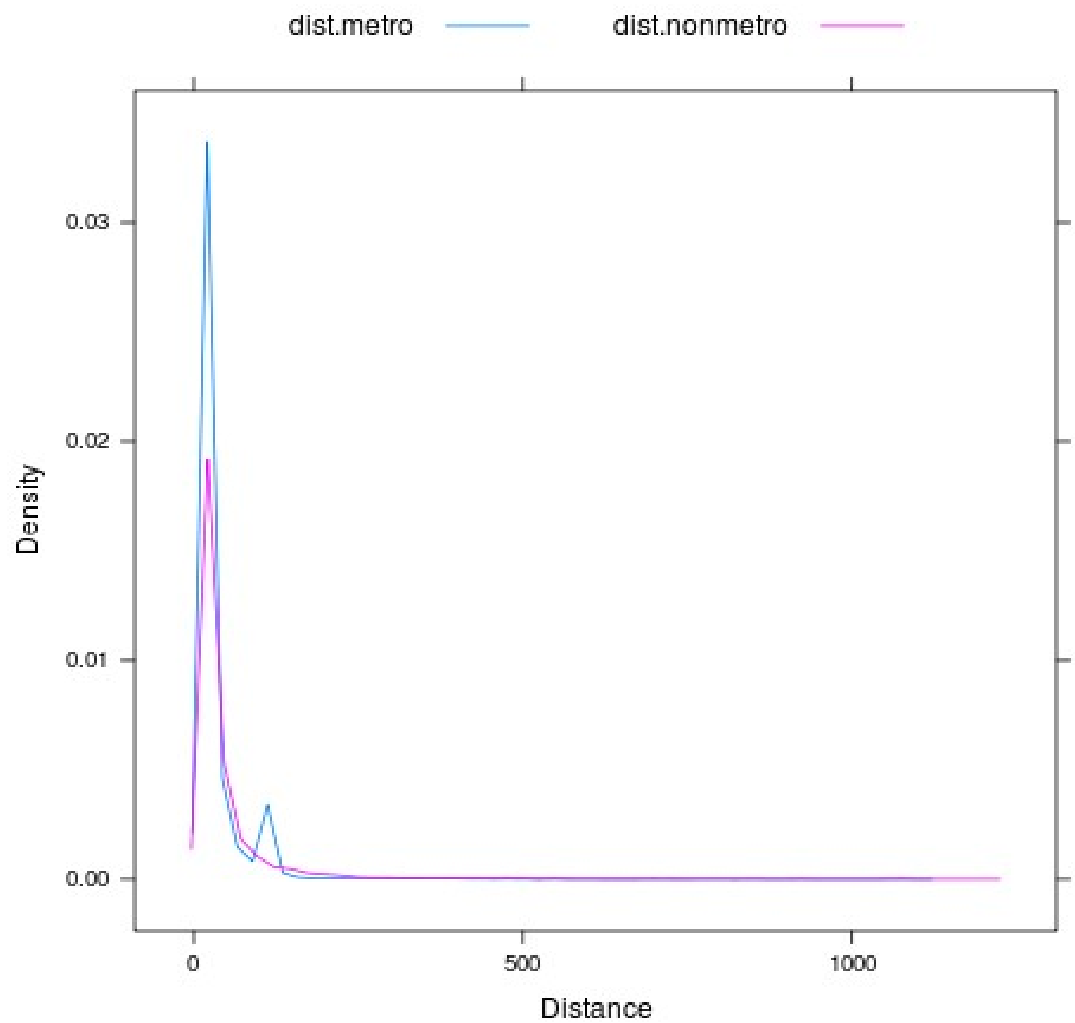
Geographic Distance Travelled

The total data was about 12,708,160 entries with the minimum distance traveled was 0, while maximum was 1221 and the mean was around 35.5km. The distribution was weakly bimodal with a strong peak around 0-35km and a second peak around 130-150km



Geographic Distance Travelled

The counties belonging to the Metropolitan areas (79 of them) are representative of the overall geographic distance histogram

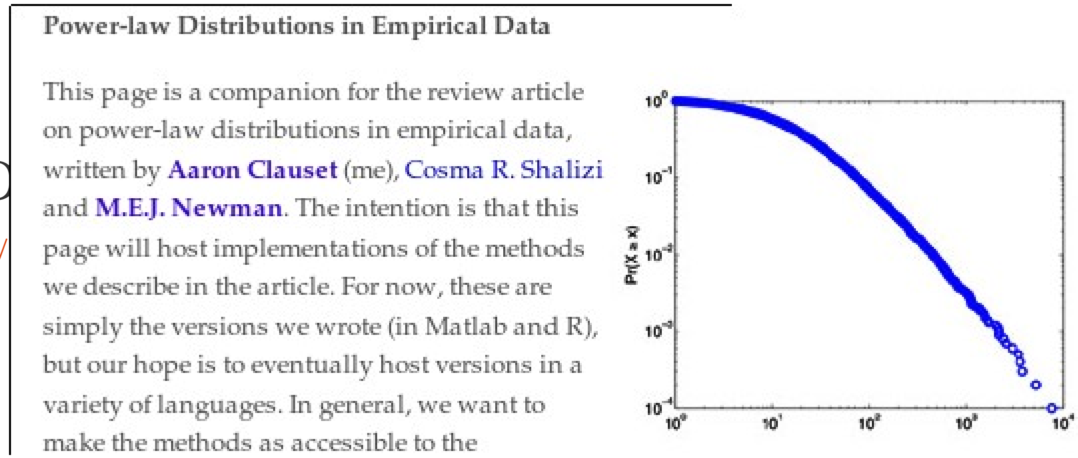


Network Based Study

- Zip based: Started with a directed weighted network of nodes as zips, based on the patient's zipcode and the Hospital's zipcode
- County based: Derived a directed weighted network of nodes as counties, based on patient's county and the Hospital's county
- Metropolitan based: Derived a directed weighted network of nodes as metro or micropolitan areas based on county (above)

Fitting Degree Distribution

- Based o tions in empirical data - http://arxiv.org/



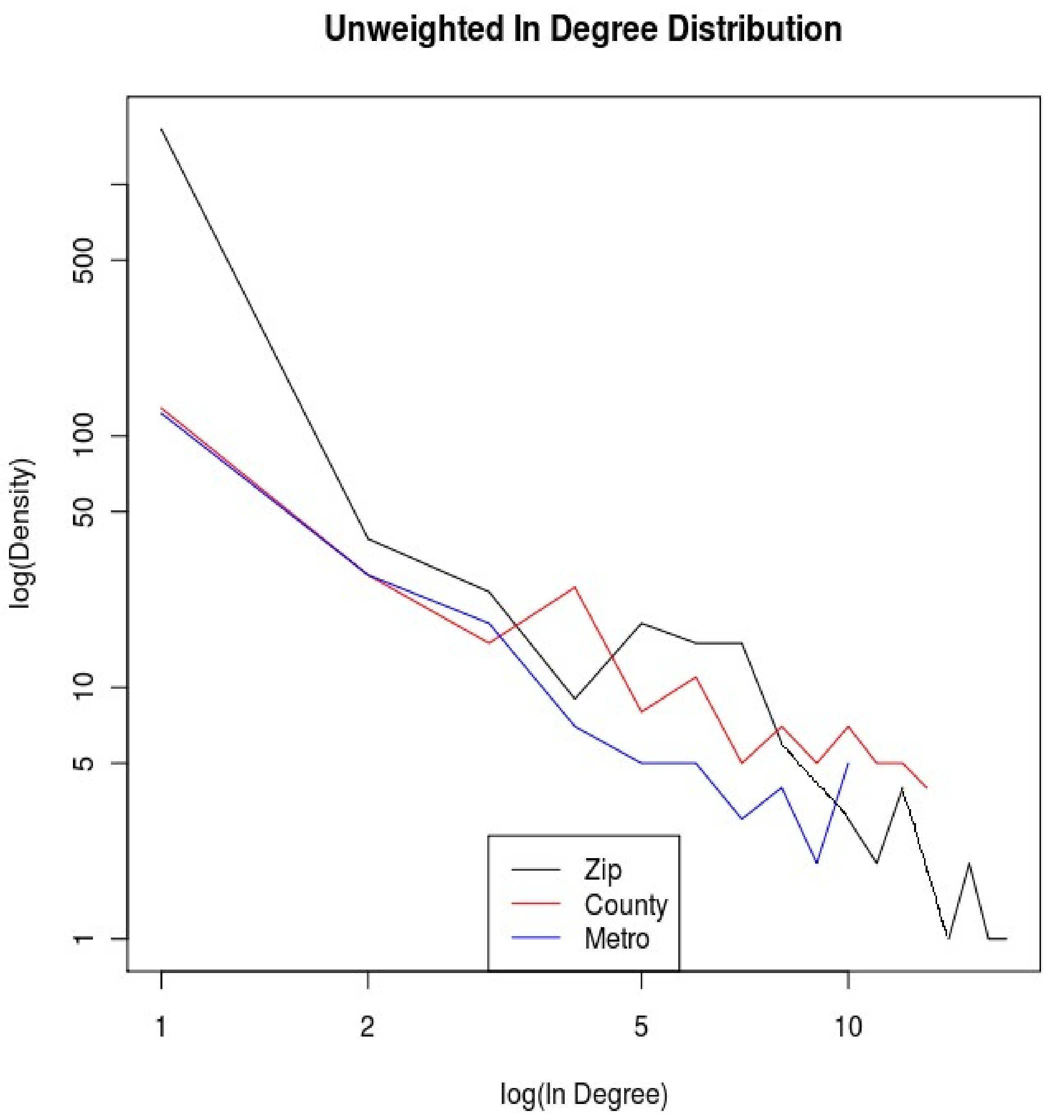
Power-law Distributions in Empirical Data

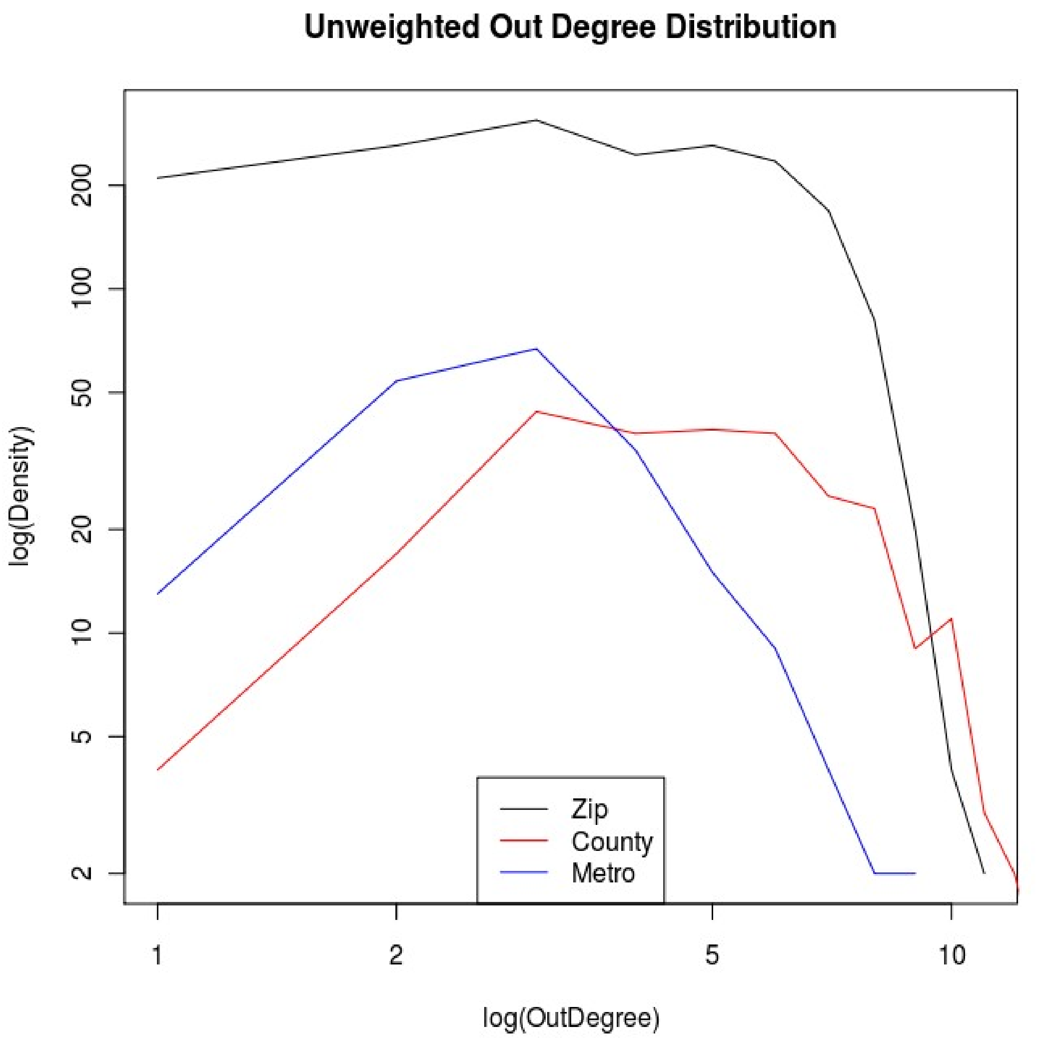
This page is a companion for the review article on power-law distributions in empirical data, written by **Aaron Clauset** (me), Cosma R. Shalizi and **M.E.J. Newman**. The intention is that this page will host implementations of the methods we describe in the article. For now, these are simply the versions we wrote (in Matlab and R), but our hope is to eventually host versions in a variety of languages. In general, we want to make the methods as accessible to the

- Fitted Zeta, Yule, Discrete Exponential, Poisson, Negative Binomial, Geometric, Discrete Weibull

- Compared AICs and picked the lowest one to report as best fit

Degree Distribution





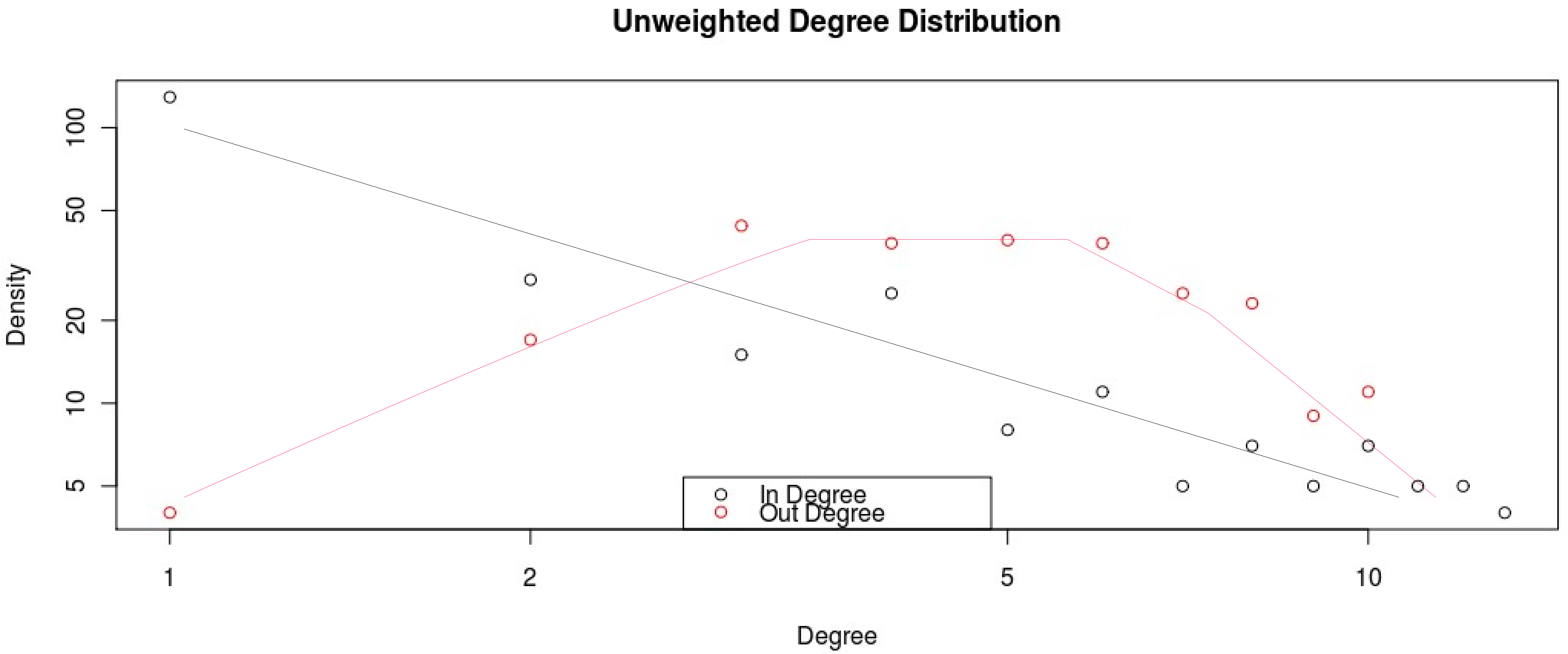
Network-based DB/WN

	In Degree	Out Degree
Zip 2 Zip (binary)	Yule; 1.18	Neg_Binom; 2.24, 36.66
County 2 County (binary)	Yule; 1.22	Neg_Binom; 4.58, 49.35
Metro 2 Metro (binary)	Yule; 1.25	Neg_Binom; 4.54, 28.32
Zip 2 Zip (weighted)	Yule; 1.101	Neg_Binom; 0.544,7072
County 2 County (weighted)	Yule; 1.098	Neg_Binom; 0.41,55221
Metro 2 Metro (weighted)	Yule; 1.103	Neg_Binom; 0.32,42572

County-based DBN

- Jurisdiction, manageable dataset and stable
- We were interested in characterizing mobility patterns, hence we looked at-
  - Weighted and unweighted degree distribution,
  - Clustering Coefficients
- Instead of properties like centrality measures, graph spectrum etc. which are more applicable in understanding cascading behavior

County - Network Based Study

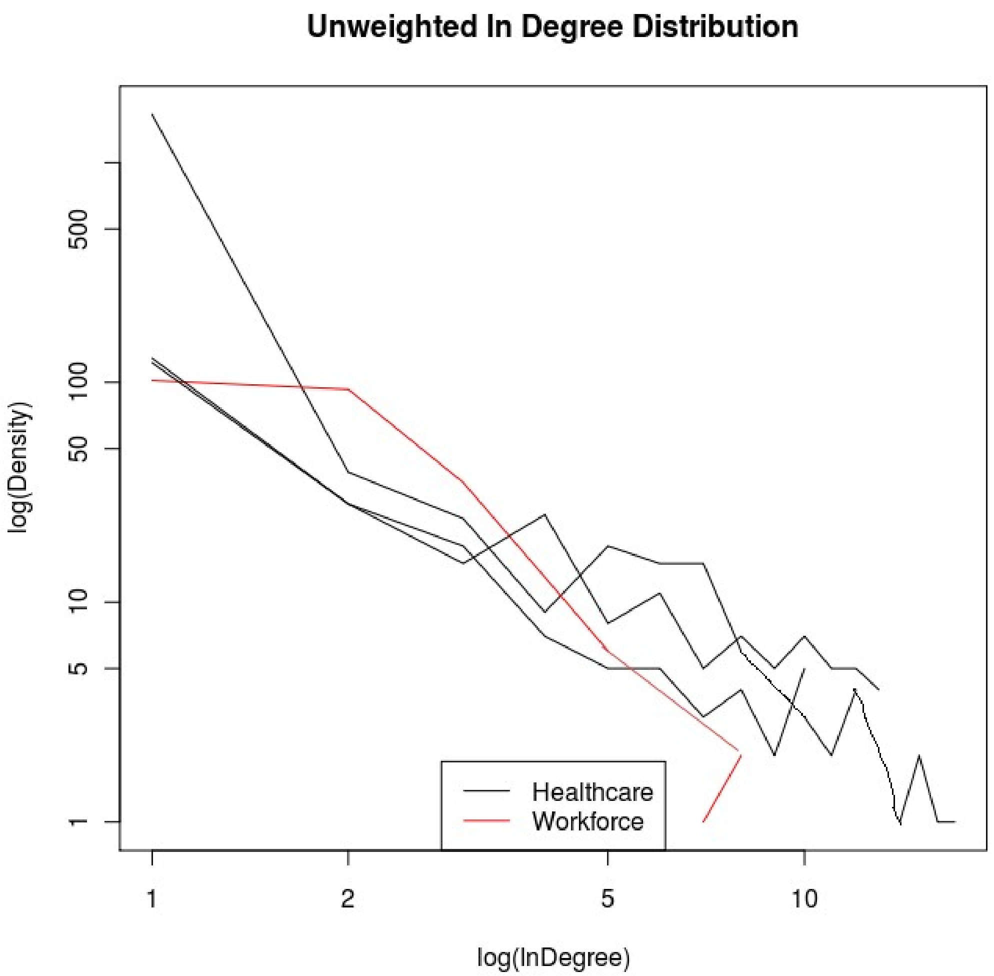


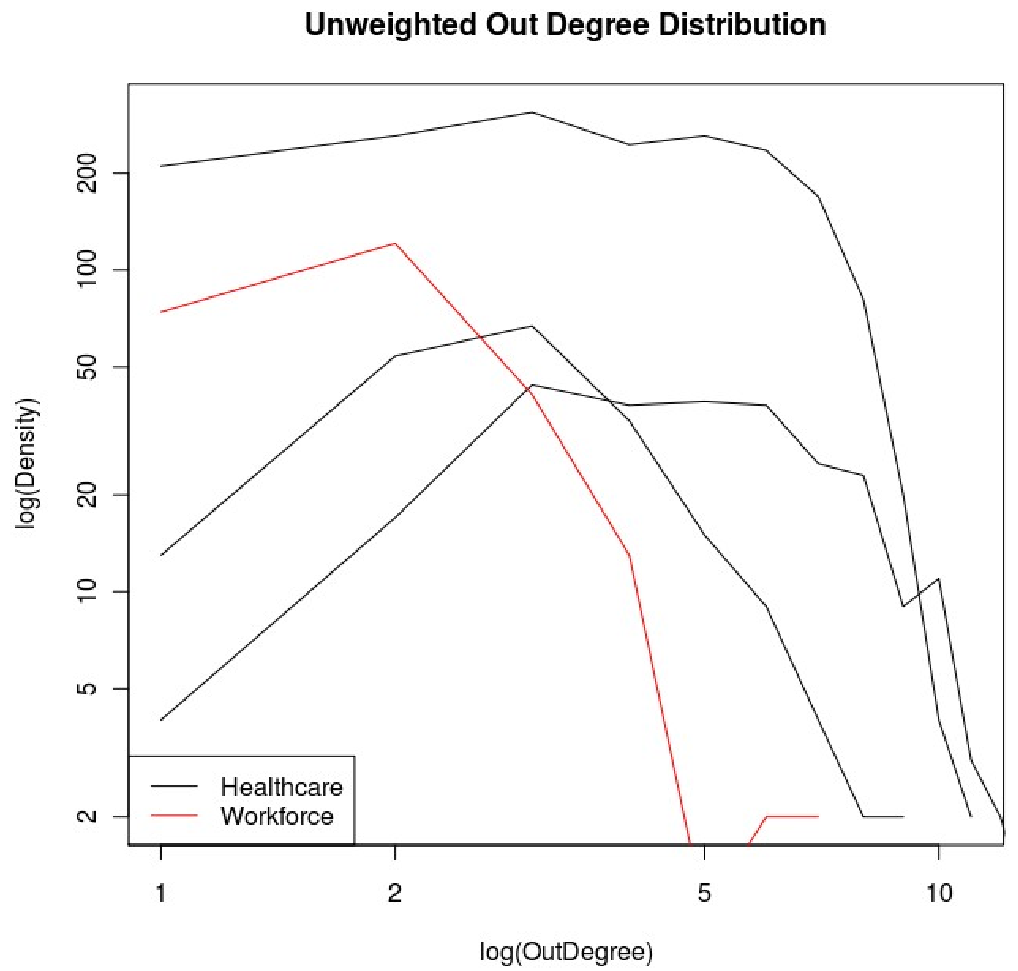
- In Degree best fitted with Yule Distribution - exponent=1.319441 with AIC 1806.458; Out Degree best fitted with Negative Binomial with S=4.577, mu=49.3504 with AIC 2302.647

Degree Distribution County

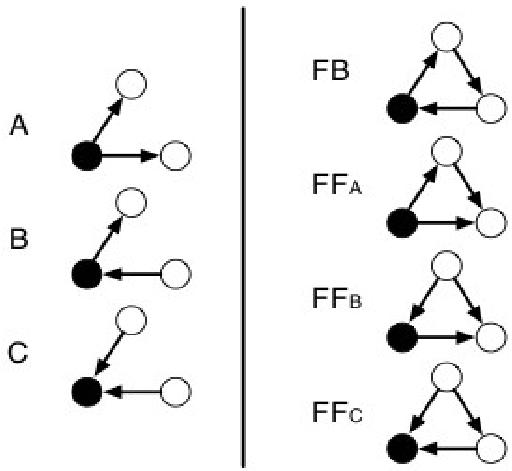
- This was markedly different from other well studied mobility network:
  - Work mobility network (of county-county) in TX were both Negative Binomial:
    - InDegree – NegBinom: 2.378, 31.449
    - OutDegree – NegBinom: 3.664, 31.44

Degree Distribution County





Transitivity County-based DWN



Feedback – Cycles

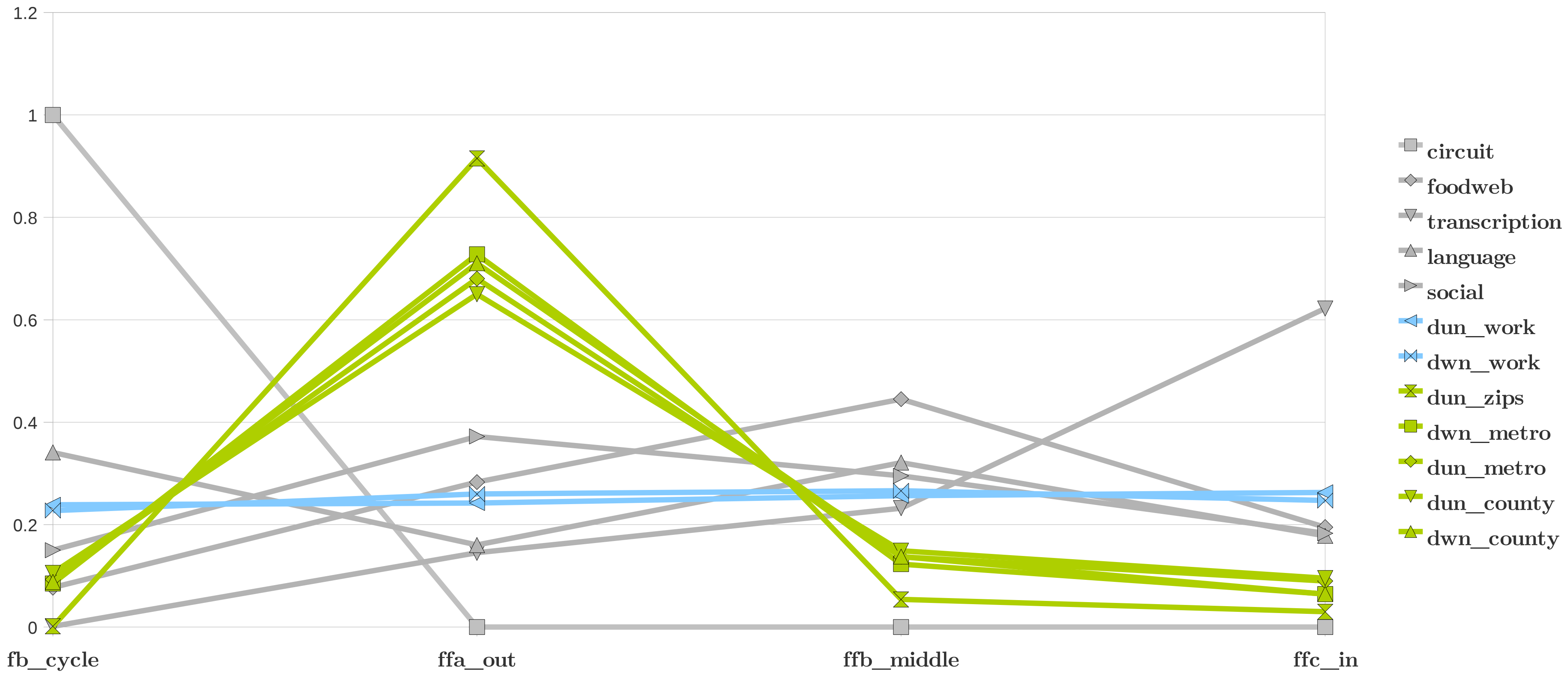
Out

Middleman

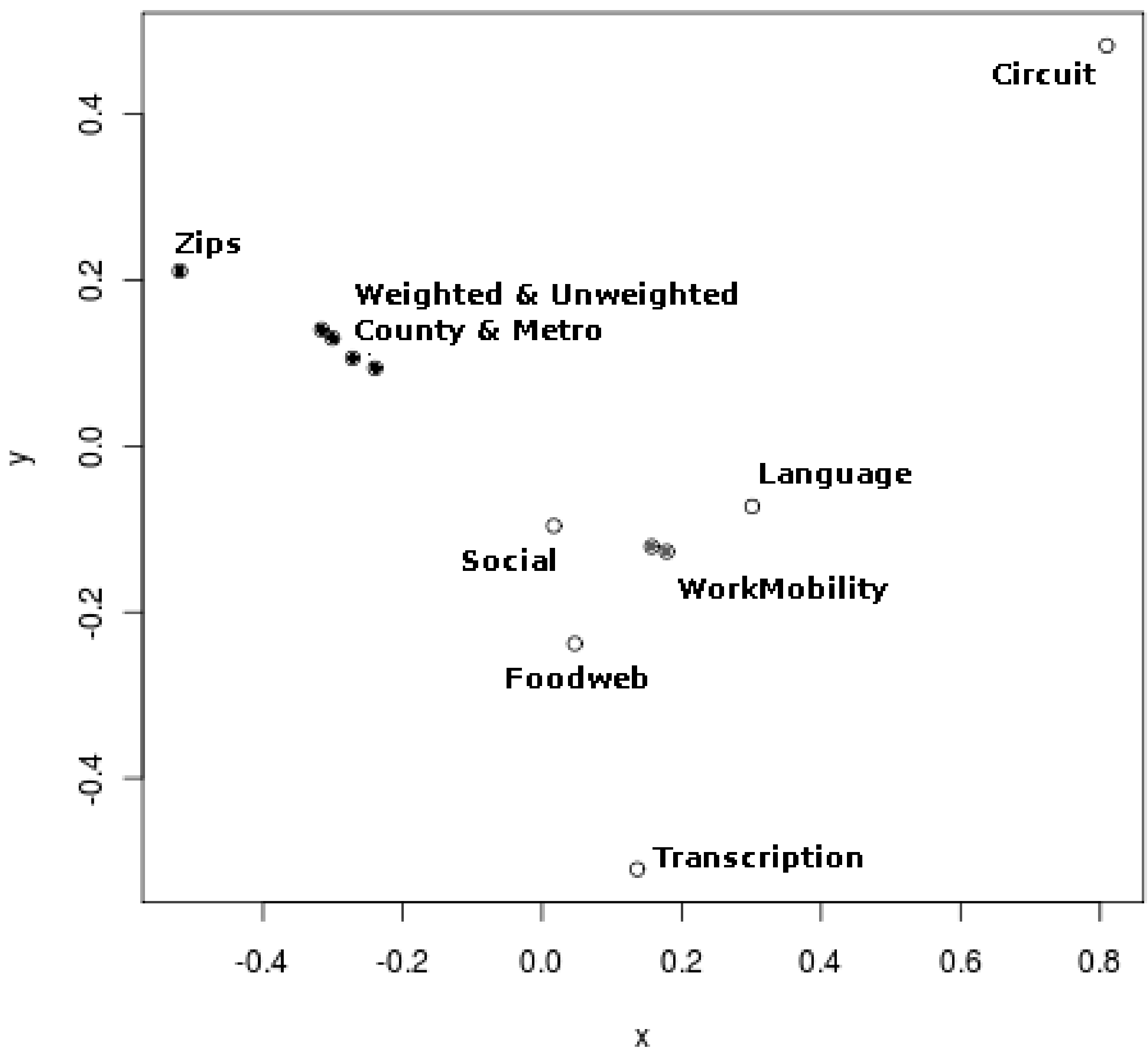
In

Following [1] computed Clustering Coefficients of County-based DWN and saw that the outward feed-forward loop was more predmoninant. We then compare it to results from [2] to see how the clustering signatures compared

Transitivity County-based DB/WN



Clustering Signature Comparison



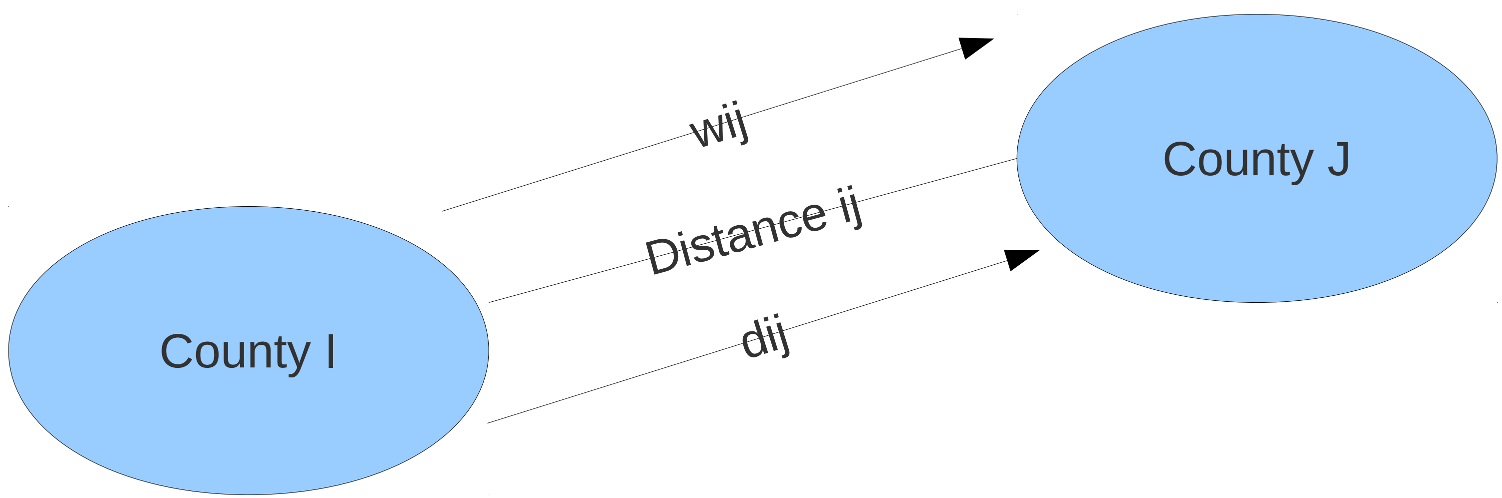
Comparing (using multi-dimensional scaling) with DUN from [2] yielded a distinct signature of the patient mobility network. We also included the Work based county to county mobility network to see if there were any commonality with our patient mobility network...

Model

Gravity Model

- Gravity Model is of the form

$$Patients_{ij} = G \frac{Population_i^{a1} * Population_j^{a2}}{Distance_{ij}^{a3}}$$



Logistic Gravity Model

- We looked at the best way to model the directed binary (unweighted) network using a Logistic Regression

$$p(d_{ij}=1)=\frac{1}{1+\exp(-A^T X)}$$

$$A=\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

model1

$$X=\begin{bmatrix} 1 \\ \ln(Population_i) \\ \ln(Population_j) \\ \ln(1+GeoDist_{ij}) \end{bmatrix}$$

model2

$$X=\begin{bmatrix} 1 \\ \ln(Population_i) \\ \ln(Population_j) \\ TopoDist_{ij} \end{bmatrix}$$

model3

$$X=\begin{bmatrix} 1 \\ \ln(Population_i) \\ \ln(Population_j) \\ \ln(1+TopoDist)_{ij} \end{bmatrix}$$

Model Performance

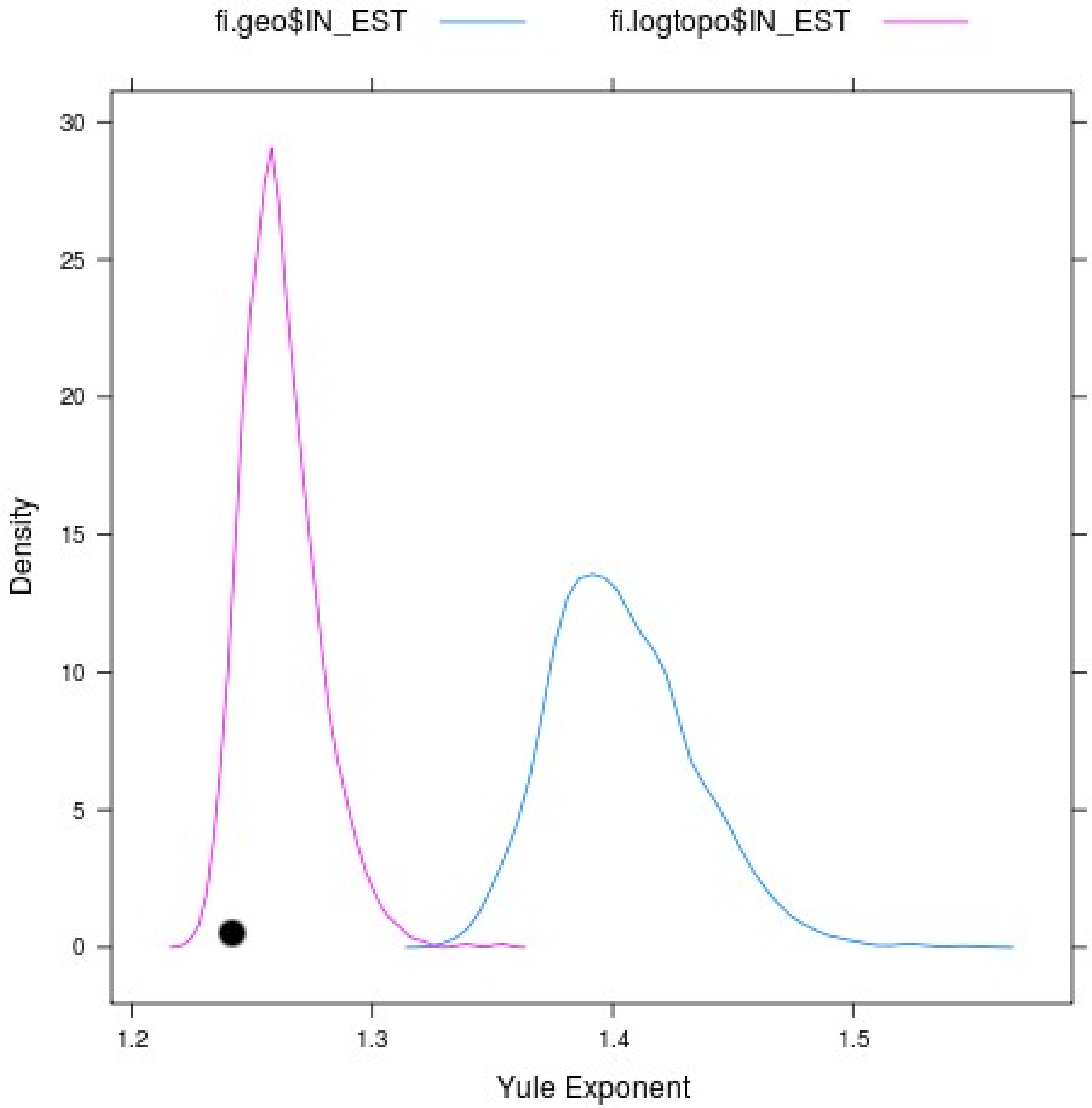
	Model1 (geo)	Model2 (topo)	Model3 (topo)
AIC	31048	31172	30518
R2	39.8	39.2	40.34

Model Performance

Ran the model with parameters from fitted models (3) 1000 times – starting with a fixed network (with 254 nodes) and geography/topology based on county adjacency and weighting the nodes based on the real TX County populations, randomly.

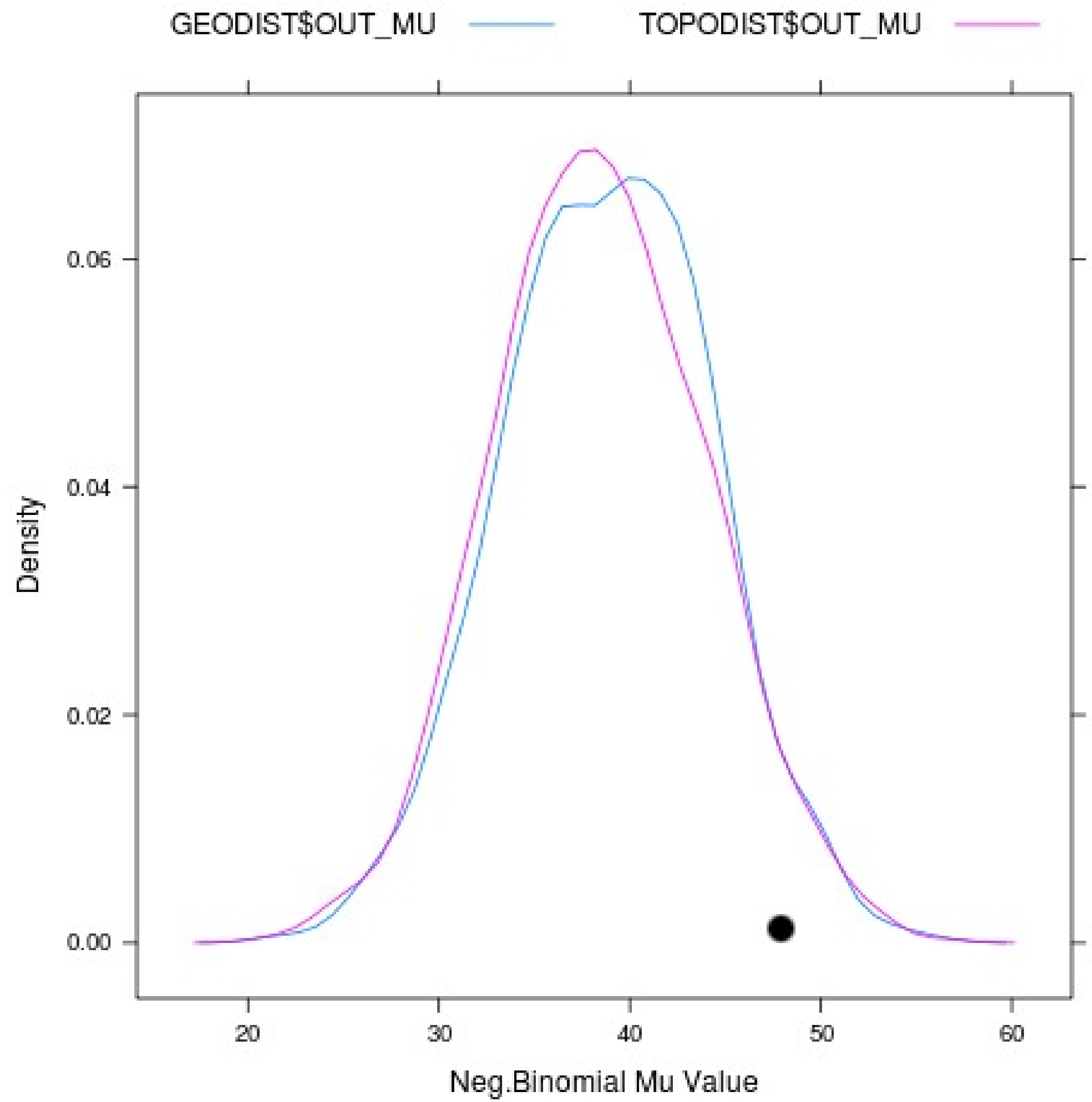
Plotted the degree distributions and clustering signatures to observe the model performance

Model Performance – In Degree Dist

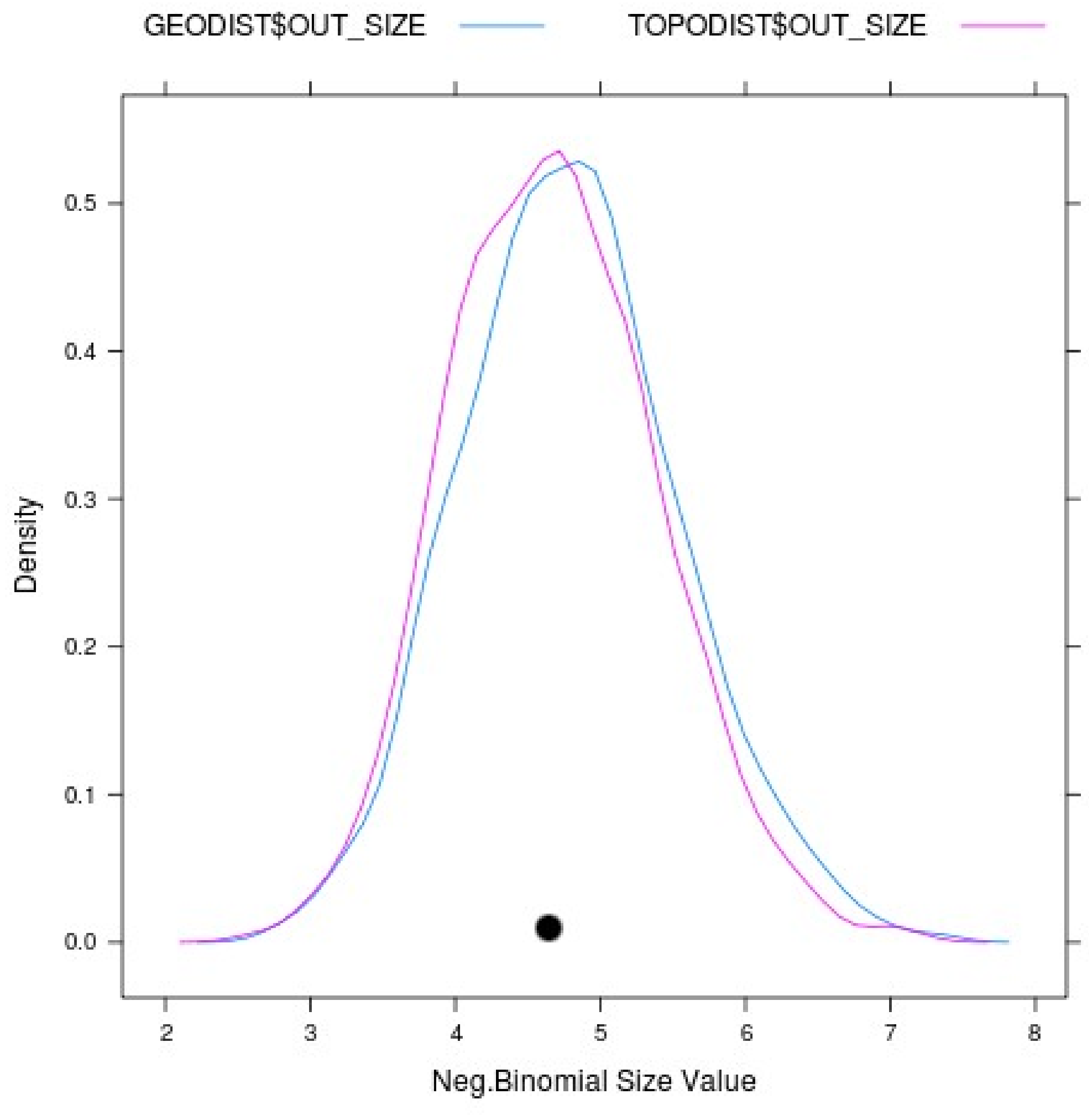


Comparing Model1 and Model3.

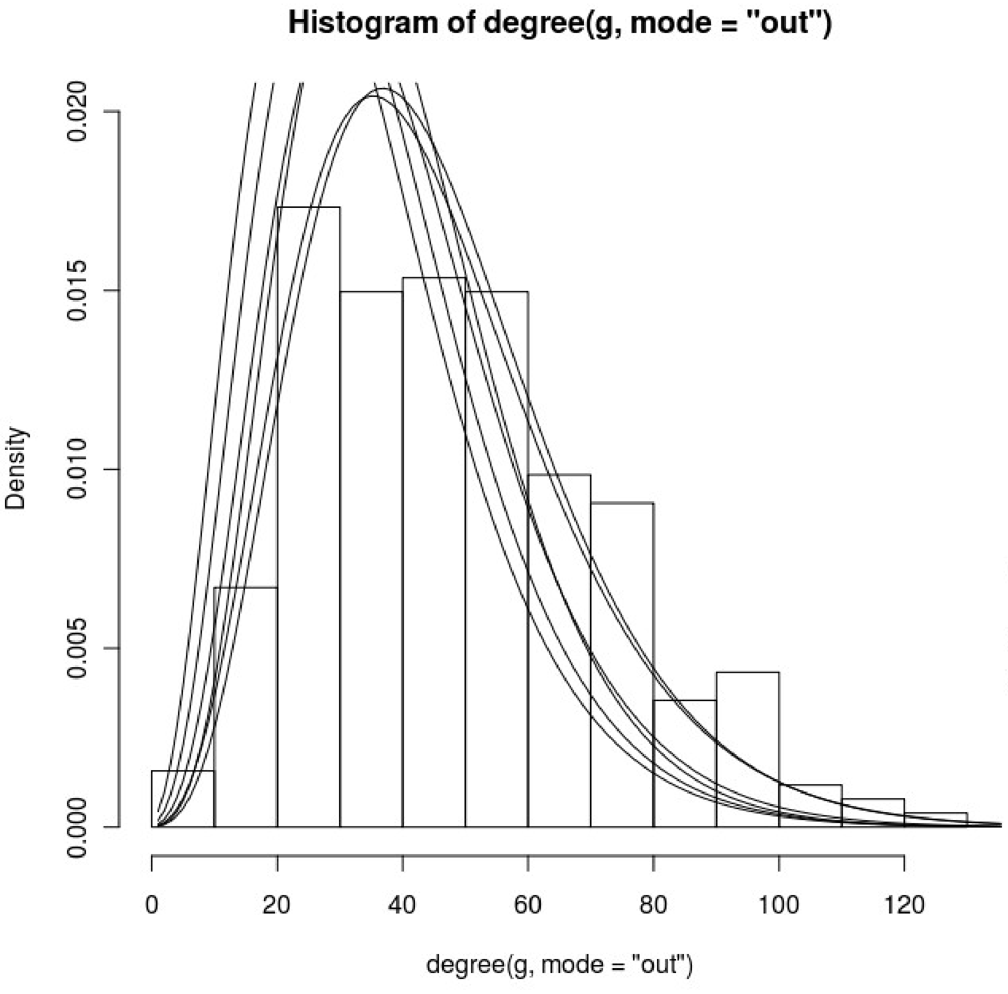
Model Performance – Out Degree Dist

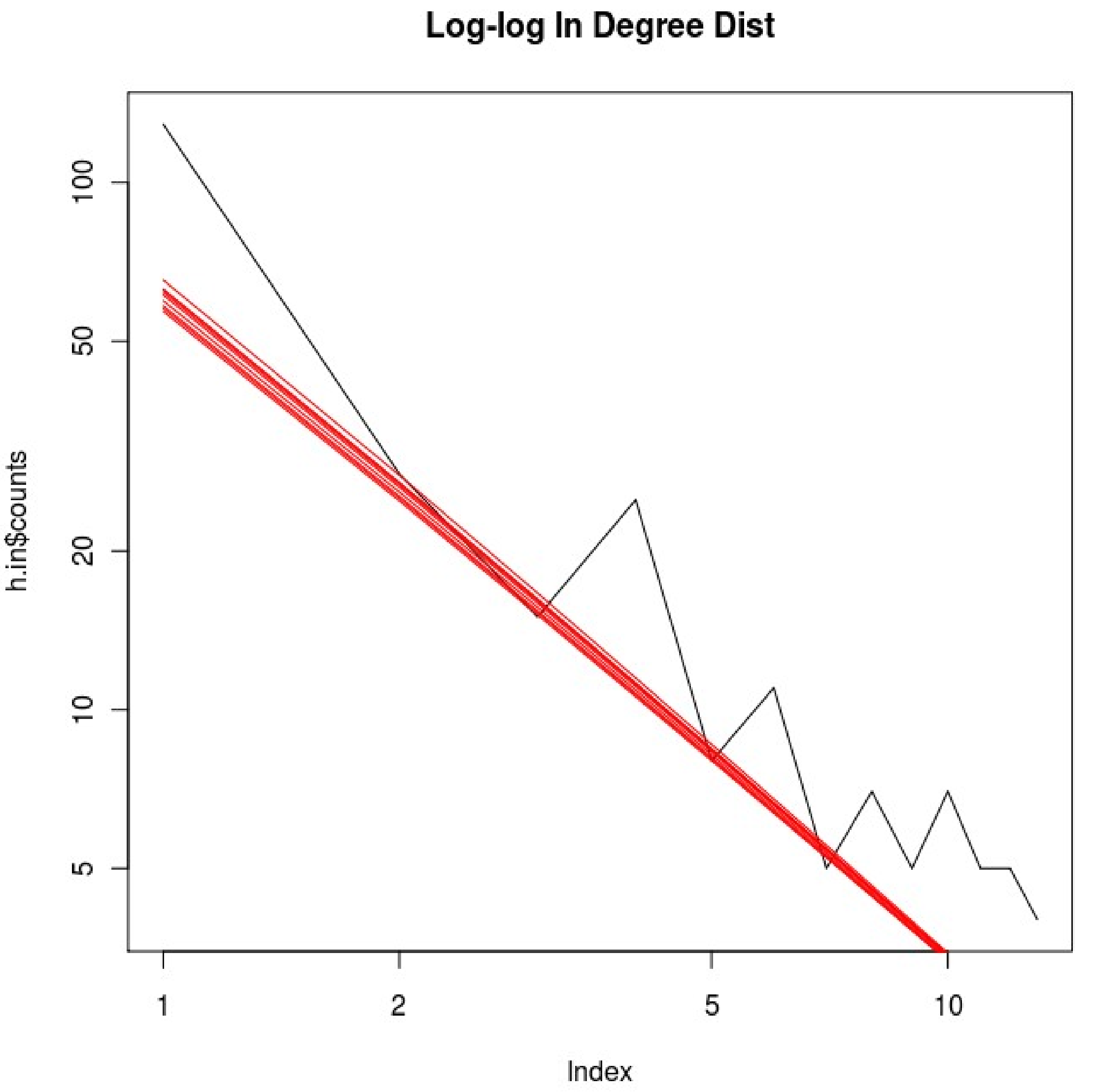


Model Performance – Out Degree Dist

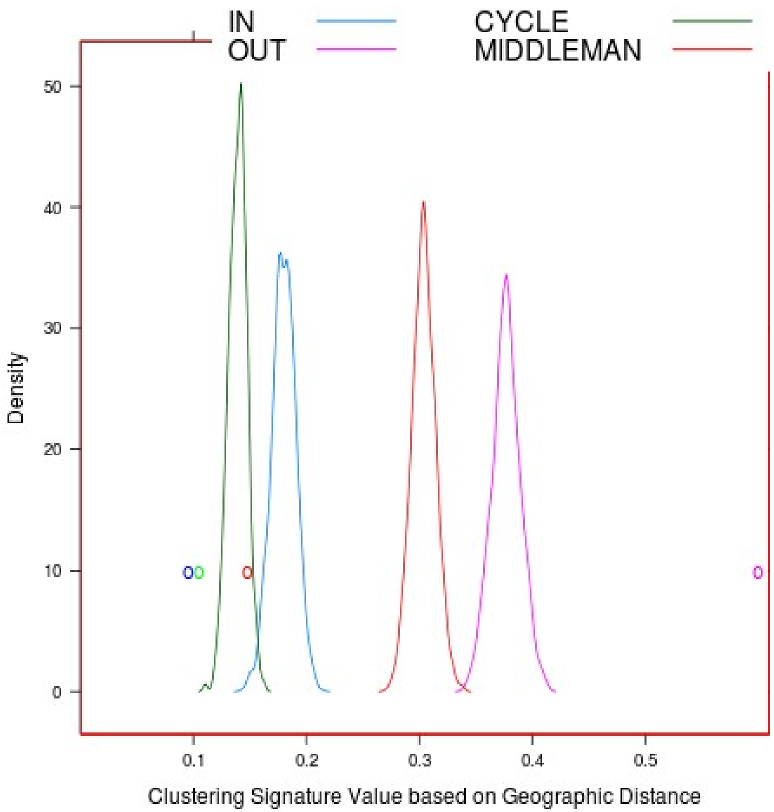


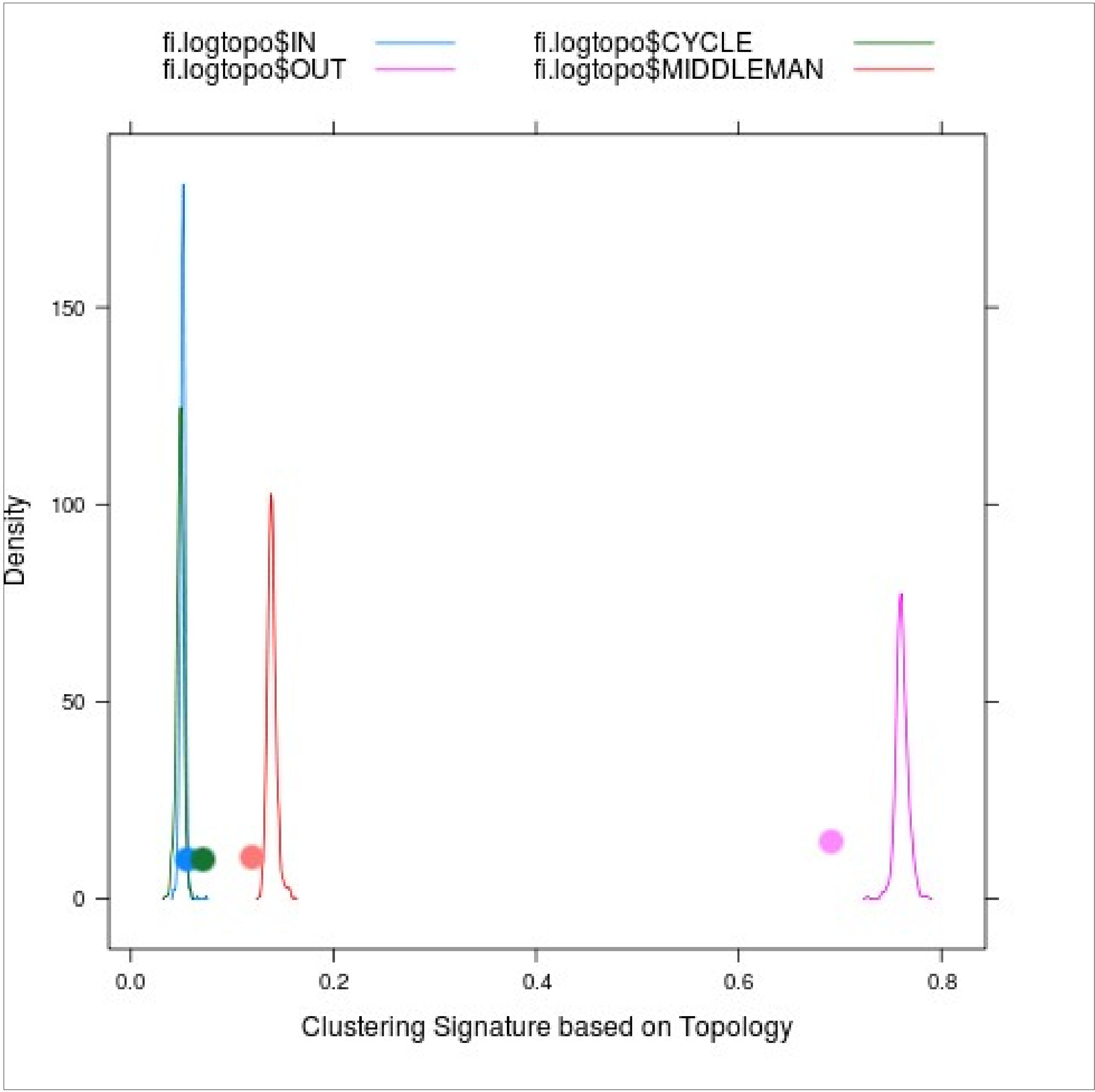
Model Performance –Degree Dist





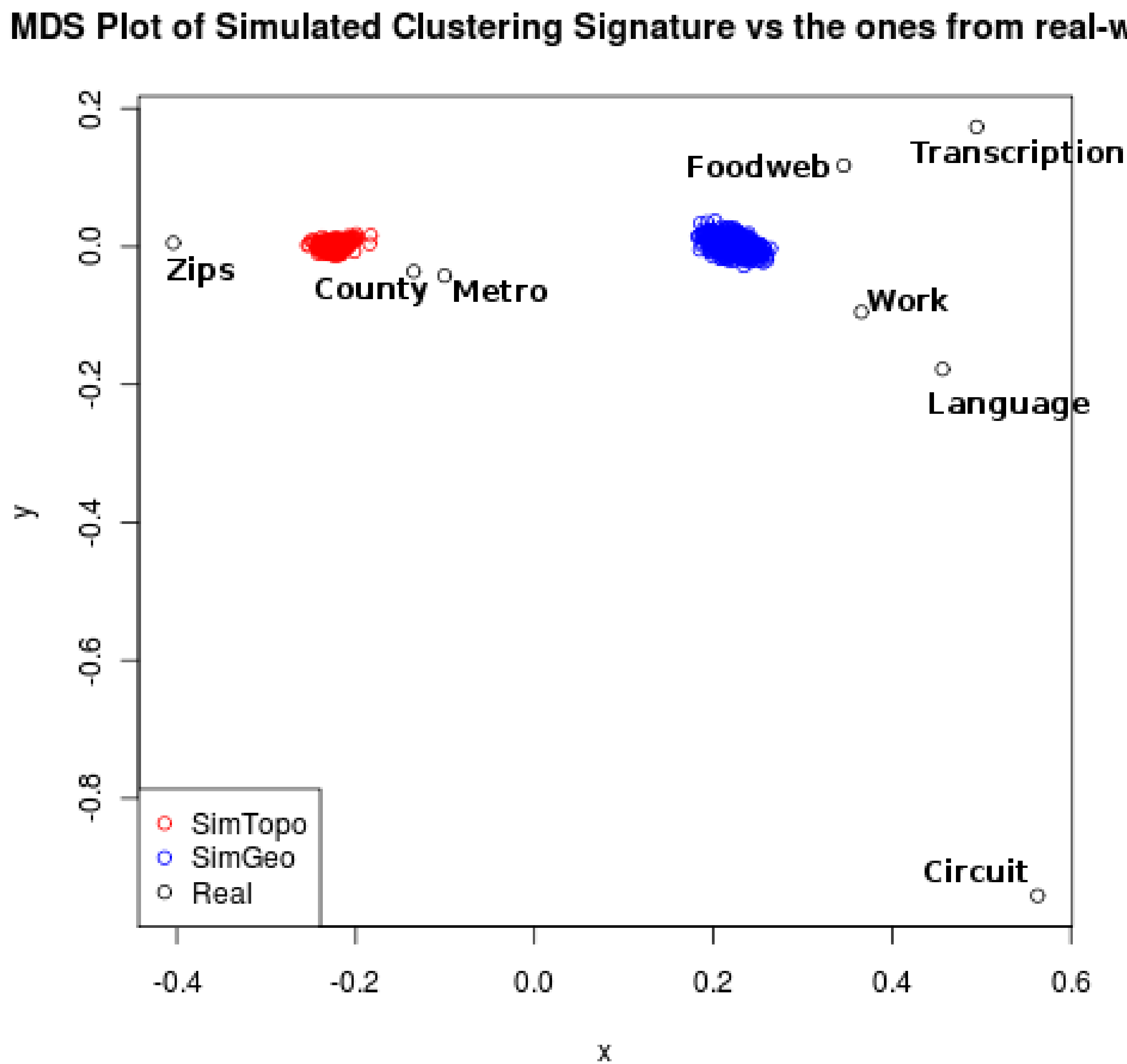
Model Performance – Clust Sig





Synthetic Model Fitting

Used the results from Network distance based simulation to get at the MDS plot to view how the simulation performed as compared to other Clustering Signatures-



Summary -1

- Network based analysis of mobility yields more insights over pure geographic analysis
- Patient mobility network is markedly different from other observed directed networks as well as from work force mobility network
- Identification of a simple model that generates a characteristic patient mobility network for future work- simulations/analysis

Summary - 2

- Topology captures and does better than geography in describing patient mobility
- A new model for network generation
  - Start with an underlying lattice (undirected) network
  - Weighting the nodes following an approximate geometric or discrete exponential distribution
  - Yields the target directed binary network with a unique clustering signature and a power-law indegree and homogenous outdegree

Math

Mathematical Derivations

- Our core model is a lattice-like graph N and 3 random variable : *P* – population or weight of each node N from a power-law like distribution, *D* – network distance in N and G – binary random variable denoting mobility between two nodes on N

$$pr(P=w)=(1-e^{-\lambda})e^{-\lambda P_{min}} \quad e^{-\lambda w}$$

$$p(d_{ij}=1)=\frac{1}{1+\mathrm{e}^{(-A^T X)}}$$

$$N \sim Lattice(dimesion=1, count=n, neighbors=k, probability=p)$$

$$pr(Deg_I=k)=\frac{Gamma(k)}{Gamma(k+\wp)}$$

$$pr(Deg_O=k)=NB(k,p)$$

References

[1] Clustering in complex directed networks - http://arxiv.org/abs/physics/0612169

[2] Clustering signatures classify directed networks - http://pre.aps.org/abstract/PRE/v78/i3/e036112

[3] Power law distributions in empirical data - http://arxiv.org/abs/0706.1062

[4] Workforce Mobility http://www.census.gov/population/www/cen2000/commuting/index.html#TX

[5] Geo

[5] TX Metropolitan Areas http://en.wikipedia.org/wiki/List_of_Texas_metropolitan_areas