Urban Diversity Through an Information-Theoretic Lens

Phil Chodrow

July 4, 2016

MIT Human Mobility and Networks Laboratory | Operations Research Center

Introduction

Two questions for today:

Two questions for today:

1. What is urban diversity, and can we measure it in spatially-aware ways?

Neighborhoods and Diversity

Two questions for today:

1. What is **urban diversity**, and can we measure it in spatially-aware ways?
2. What are **neighborhoods**, and how can we distinguish neighborhoods from each other?

Two questions for today:

1. What is **urban diversity**, and can we measure it in spatially-aware ways?
2. What are **neighborhoods**, and how can we distinguish neighborhoods from each other?

Central Claim

These two questions are deeply related, and we should approach both using *information theory*.

Three Dimensions of Urban Diversity

Global diversity is the extent to which different groups are equally represented in the city as a whole.

Three Dimensions of Urban Diversity

Global diversity is the extent to which different groups are equally represented in the city as a whole.

Spatial evenness *"or clustering, refers to the extent to which groups are similarly distributed in residential space."* [12]

Three Dimensions of Urban Diversity

Global diversity is the extent to which different groups are equally represented in the city as a whole.

Spatial evenness *"or clustering, refers to the extent to which groups are similarly distributed in residential space."* [12]

Spatial exposure *"refers to the extent that members of one group encounter members of another group (or their own group, in the case of spatial isolation) in their local spatial environments."* [12]

Three Dimensions of Urban Diversity

Global diversity is the extent to which different groups are equally represented in the city as a whole.

Spatial evenness *“or clustering, refers to the extent to which groups are similarly distributed in residential space.”* [12]

Spatial exposure *“refers to the extent that members of one group encounter members of another group (or their own group, in the case of spatial isolation) in their local spatial environments.”* [12]

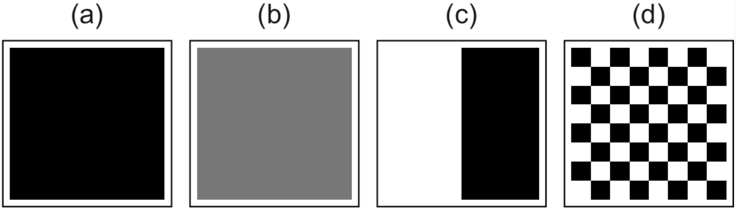
No global diversity $\implies$ perfect evenness $\implies$ maximal exposure.

Neighborhoods, Exposure, and the Checkerboard

Claim: Spatial exposure is about the number, scale, and pattern of neighborhoods.

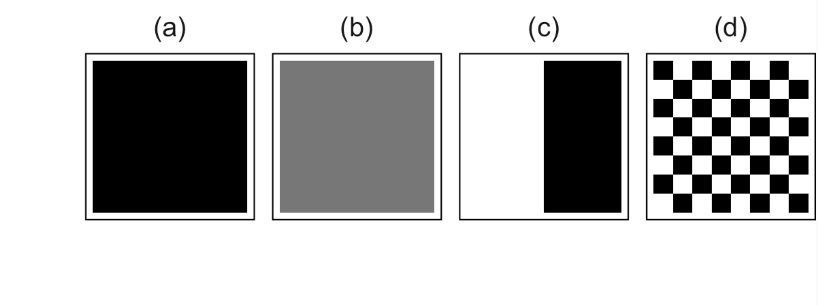
Neighborhoods, Exposure, and the Checkerboard

Claim: Spatial exposure is about the number, scale, and pattern of neighborhoods.



Neighborhoods, Exposure, and the Checkerboard

Claim: Spatial exposure is about the number, scale, and pattern of neighborhoods.



City	Exposure	Neighborhoods
(c)	Low	Few, large
(d)	High	Many, small

Diversity and Information Theory

Summary: Information Theory and Diversity

Entropy Measures Diversity

$H(Y) \triangleq -\mathbb{E}_Y[\log p(Y)]$

Summary: Information Theory and Diversity

Entropy Measures Diversity

$H(Y) \triangleq -\mathbb{E}_Y[\log p(Y)]$

Mutual Information Measures Evenness

$I(X, Y) \triangleq \mathbb{E}_{X,Y}\left[\log \frac{p(X,Y)}{p(X)p(Y)}\right]$

Summary: Information Theory and Diversity

Entropy Measures Diversity

$H(Y) \triangleq -\mathbb{E}_Y[\log p(Y)]$

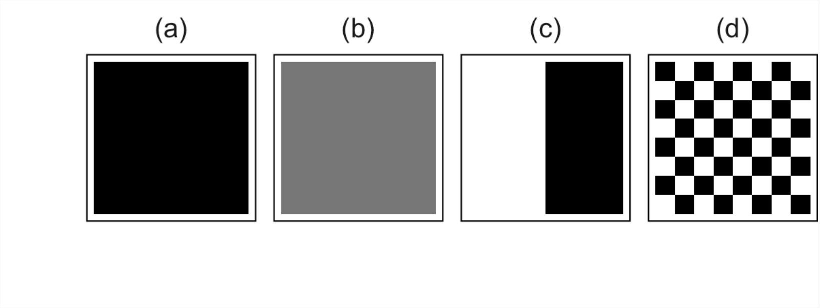
Mutual Information Measures Evenness

$I(X,Y) \triangleq \mathbb{E}_{X,Y}\left[\log \frac{p(X,Y)}{p(X)p(Y)}\right]$

Mean Local Information Measures Exposure

$J(X,Y) \triangleq \mathbb{E}_X[\text{trace } J_Y(X)]$

Back to the checkerboard



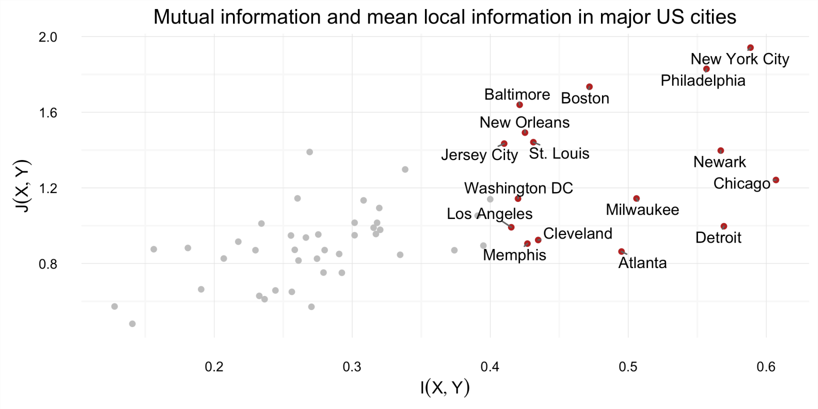
City	$H(Y)$	$I(X, Y)$	$J(X, Y)$
(a)	0.0	0.0	0.0
(b)	0.7	0.0	0.0
(c)	0.7	0.7	0.6
(d)	0.7	0.7	2.7

A Three-Dimensional Characterization of Urban Diversity

City	$H(Y)$	$I(X,Y)$	$J(X,Y)$
Albuquerque	1.08	0.19	0.66
Detroit	1.12	0.57	1.00
Philadelphia	1.31	0.56	1.83

Data accessed from the American Community Survey of the U.S. Census [1].

Visualizing Diversity Profiles in Major Cities



- Philadelphia has greater exposure / more neighborhood structure than Detroit.

Identifying Natural Neighborhoods

Key Property of Mutual Information

Suppose we cluster the locations X into groups with labels C.

“Additive Organizational Decomposability” (Chain Rule)

$$I(X, Y) = I(C, Y) + I(X, Y|C)\ .$$

Key Property of Mutual Information

Suppose we cluster the locations X into groups with labels C.

“Additive Organizational Decomposability” (Chain Rule)

$$I(X, Y) = I(C, Y) + I(X, Y|C) \ .$$

We can use this to identify neighborhoods!

Key Property of Mutual Information

Suppose we cluster the locations X into groups with labels C.

Recipe for Neighborhood Identification

$$I(X, Y) = \underbrace{I(C, Y)}_{\text{maximize this}} + \overbrace{I(X, Y|C)}^{\text{minimize this}} .$$

Key Property of Mutual Information

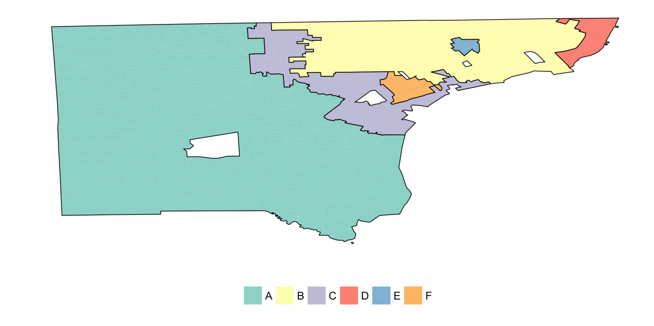
Suppose we cluster the locations X into groups with labels C.

Recipe for Neighborhood Identification

$$I(X, Y) = \underbrace{I(C, Y)}_{\text{maximize this}} + \overbrace{I(X, Y|C)}^{\text{minimize this}} .$$

This optimization is hard to do exactly, but we can use a greedy approach based on hierarchical clustering.

Practical Example: Racial Neighborhoods in Detroit



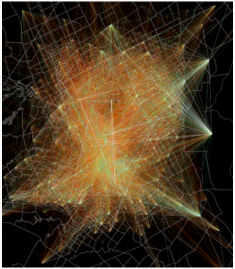
	A	B	C	D	E	F
White	36	2	7	2	1	1
Black	5	29	4			
Hispanic	2		1			2
Asian	2				1	
Other	2	1	1			

Percentage of population by race and neighborhood.

Wrapping Up

Generality and future work

- **Other demographics**: education level, income, occupation type, etc.
- **Other dimensions**: time
    - Evolution of urban demographics (timescale: decades)
    - Dynamics of diversity in daily mobility (timescale: minutes)



Daily commuting flows in Riyad, Saudi Arabia.
Image credit: Shan Jiang and MIT HuMNetLab

Download our tools!

R Package for Information-Theoretic Analysis

https://github.com/PhilChodrow/compx

This project: analysis and presentations

https://github.com/PhilChodrow/spatial_complexity

THANK YOU! Questions? Feedback?

Supplementary Materials

References I

2014 American Community Survey 5-Year Estimates, Table B01003, 2016.

Luis M A Bettencourt, Joe Hand, and José Lobo.
Spatial Selection and the Statistics of Neighborhoods.
2015.

Roger Bivand, Tim Keitt, and Barry Rowlingson.
rgdal: bindings for the Geospatial Data Abstraction Library, 2014.

Roger Bivand and Nicholas Lewin-Koh.
maptools: Tools for reading and handling spatial objects, 2014.

References II

Roger Bivand and Colin Rundel.
rgeos: Interface to Geometry Engine - Open Source (GEOS), 2014.

Thomas M Cover and Joy A Thomas.
Elements of Information Theory.
John Wiley and Sons, New York, 1991.

Imre Csiszar and Paul C Shields.
Information Theory and Statistics: A Tutorial.
Foundations and Trends™ in Communications and Information Theory, 1(4):417–528, 2004.

References III

Barret A Lee, Sean F Reardon, Glenn Firebaugh, Chard R Farrell, Stephen A Matthews, and David O'Sullivan.
Beyond the Census Tract: Patterns and Determinants of Racial Segregation at Multiple Geographic Scales.
American Sociological Review, 73(5):766–791, 2008.

Douglas S Massey and Nancy A Denton.
The Dimensions of Residential Segregation.
Social Forces, 67(2):281–315, 1988.

R Core Team.
R: A Language and Environment for Statistical Computing, 2016.

References IV

Sean F Reardon and G Firebaugh.
Measures of multigroup segregation.
Sociological Methodology, 32:33–67, 2002.

Sean F. Reardon and David O'Sullivan.
Measures of Spatial Segregation.
Sociological Methodology, 34(1):121–162, 2004.

Relation of local and Fisher informations

Suppose that $p(y|x) > 0$ and that $p(y|x)$ is differentiable as a function of x for all $x, y \in \mathbb{R}^n \times \mathcal{Y}$. Define $B_r \triangleq B_r(x_0) \triangleq \{x \in \mathbb{R}^n \mid |x - x_0| \leq r\}$. Additionally, define the *local mutual information* in B_r as the mutual information between X and Y where X is restricted to B_r:

$$
\begin{aligned}
I_r(x_0) &\triangleq \mathbb{E}_X[D[p(\cdot|X)\|p(\cdot|X \in B_r)]|X \in B_r] &(1)\\
&= \int_{B_r} p(x|X \in B_r) D[p(\cdot|x)\|p(\cdot|X \in B_r)] d^n x\,. &(2)
\end{aligned}
$$

where $D[p\|q] \triangleq \sum_y p(y) \log \frac{p(y)}{q(y)}$ is the Kullback-Leibler divergence of q from p.

Relation of local and Fisher informations

Theorem

Under the stated conditions,

$$\lim_{r \to 0} \frac{I_r(x_0)}{r^2} = \frac{n}{2(n+2)} \operatorname{trace} J_Y(x_0) \ . \tag{3}$$

The Fisher information matrix J_Y is given by

$$J_Y(x) \triangleq \mathbb{E}_Y \left[\nabla_x S_Y(x) \nabla_x S_Y(x)^T \right] \tag{4}$$

$$S_y(x) \triangleq \log p(y|x) \ . \tag{5}$$

Information Measures Check the Boxes

Criterion	$H(Y)$	$I(X, Y)$	$J(X, Y)$
Organizational Equivalence	✔	✔	✔
Size/Density Invariance	✔	✔	✔
Additive Group Decomposability	✔	✔	✔
Additive Spatial Decomposability	NA	✔	*
Scale Interpretability	NA	✔	✔
Boundary Independence	NA	✔	✔
Exchanges	NA	✔	✔

* We contend that [12]'s version is not a workable criterion for explicitly spatial measures.

✔ In the theoretical definition; in practice we are reliant on the data as it is provided.

Some Comparative Clusters of Equal Information