New Laws of City Growth

Diego Rybski,
Hernán D. Rozenfeld, José S. Andrade Jr.,
Michael Batty, H. Eugene Stanley,
Xavier Gabaix, Hernán A. Makse

FU Berlin,
Institute of Meteorology,
21.3.2011, 15:00
How to define cities?

Different definitions of cities affect statistical properties of urban activity.

Method to define cities by the US census bureau: MSAs, places, counties, etc.

Conventional Method to define (large) “cities” in the USA: **Metropolitan Statistical Areas (MSA)**.

MSAs are subject to socio-economical factors and constructed manually, done only for the largest cities (about 300) and cannot be easily applied in other countries.
Outline

I. City Clustering Algorithm (CCA)

II. City size (Zipf's Law)

III. City growth (Gibrat's Law)

City growth (Gibrat's Law) & CCA:

City size (Zipf's Law):
I. City Clustering Algorithm (CCA)
A new definition of cities

We define a new way to construct cities:
- unbiased
- automated
- fast
- can be easily used in any country
- based only on location of population
- allows studying cities at different level of observation

City Clustering Algorithm (CCA)
Data

*Great Britain* (England, Scotland, and Wales):
58.7 millions in 2007
0.23 million km²
grid of 200m x 200m

*USA*:
303 millions in 2007
7.44 million km²
59456 sites (FIPS)
A new definition of cities

The map is gridded. The populated cells are identified.

We define: “a city” = “a cluster of connected populated cells, with maximal size”
A new definition of cities

The burning algorithm.
Pick a populated cell and burn it.

Stauffer *Introduction to percolation theory* (‘84).
A new definition of cities

The burning algorithm.

Find the populated neighbors of the burnt cell and burn them.
A new definition of cities

Find the populated neighbors of the burnt cells and burn them.

The burning algorithm.
Recursively, continue identifying the populated neighbors of the burnt cells and burning them.

The burning algorithm.
A new definition of cities

The burning algorithm.

When all burnt cells have no populated neighboring cells, the cluster is completed.

Our first cluster!
A new definition of cities

The burning algorithm.

Repeat the procedure Until all populated cells are burnt.

Our three clusters!
A new definition of cities
CCA in Great Britain

CCA applied to Greater London
CCA in the USA

- Boston
- New York City
- Jersey City
- Newark
- Philadelphia
- Baltimore
- Washington DC
II. City size (Zipf's Law)
Does the city size distribution obey Zipf's Law?

Zipf's Law

\[ P(S) \sim S^{-\zeta - 1}, \quad \zeta = 1 \]

The distribution of sizes follows a power-law with \( \zeta = 1 \)

Zipf's law has been documented for words, firms, size of exports, and many more.
Zipf's Law

Understanding the origin of this regularity is an ongoing task.

Typically, studies use MSAs for the top 200 cities, i.e. Eeckhout ('07)

Eeckhout ('07)
Uses data on all administrative cities
Finds a very good log-normal fit

Distribution of city size using the CCA?
Zipf's Law for the USA
Zipf's Law for the GB
Comparison with MSA: Northeastern USA
Correlations between MSA and CCA
Correlations with Places

(a) Scatter plot showing the correlation between $\log(S_{\text{CCA}})$ and $\log(S_{\text{Places}})$.

(b) Scatter plot showing the correlation between $\log(S_{\text{Places}})$ and $\log(S_{\text{MSA}})$.
Correlations with Area

USA

Slope = 0.958

GB

Slope = 1.065
III. City growth (Gibrat's Law)
City growth

\[ S_0 \quad \text{Population of a city at time 0.} \]

\[ S_1 \quad \text{Population of a city at time 1.} \]

\[ S_1 = R(S_0)S_0 \quad \rightarrow \quad R \quad \text{growth factor} \]

\[ r(S_0) \equiv \ln R(S_0) = \ln(S_1/S_0) \quad \rightarrow \quad r \quad \text{growth rate} \]

\[ \langle r(S_0) \rangle \sim S_0^{-\alpha} \]

\[ \sigma(S_0) = \sqrt{\langle r(S_0)^2 \rangle - \langle r(S_0) \rangle^2} \]

\[ \sigma(S_0) \sim S_0^{-\beta} \]
City growth in the USA (1990-2000)

\[ \langle r(S_0) \rangle \sim S_0^{-\alpha}, \quad \alpha = 0.28 \]

\[ \sigma(S_0) \sim S_0^{-\beta}, \quad \beta = 0.20 \]

Are not in agreement with Gibrat's Law (stating that average growth rate and standard deviation are constant)
City growth in the GB (1981-1991)

\[ \langle r(S_0) \rangle \sim S_0^{-\alpha}, \quad \alpha = 0.17 \]

\[ \sigma(S_0) \sim S_0^{-\beta}, \quad \beta = 0.27 \]
Summary

- CCA constructs cities based only on geographical features
- Zipf's Law holds over a wide range, even for smaller cities
- Scale-invariant growth mechanisms at different geographical scales (violation of Gibrat's Law)
- Power-law standard deviation is due to long-range spatial correlations in the growth (not shown)
- How about other countries?
Thank you for your attention.

http://www.rybski.de/diego/
http://www.pik-potsdam.de/members/rybski/