Understanding and Rewiring Cities using Big Data

Bruno Lepri
Some cities are alive, others less so

ALIVE
NEW YORK
DEAD

How do you capture death & life of cities?
The systematic acceleration of urban life

- Social
  human-interactions

- Economical
  GDP, wages, patents

- Issues
  violent crimes, contagious diseases, pollution

The theory: Jane Jacobs

One of the most influential books in city planning

- planning models that dominated mid-century planning
- American housing policy (HOPE VI)
- Melbourne, Toronto etc.

The theory: not tested!

- Not empirically tested until 2015
- Tested in Seoul, from costly surveys collected in years
- Theory from 1961!

The theory: Jane Jacobs

One of the **most** influential books in city planning

- **Death**: caused by the elimination of pedestrian activity
- **Life**: created by a vital urban fabric at all times of the day
Jacobs’ diversity conditions

<table>
<thead>
<tr>
<th>LAND USE</th>
<th>SMALL BLOCKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGED BUILDINGS</td>
<td>DENSITY</td>
</tr>
</tbody>
</table>

Diversity => Urban vitality

There are 4 diversity conditions

To be ensured in each city’s district
(10,000+ inhabitants)
Land Use Mix

2+ primary uses (contemporarily)

JACOBS’ VIEW: People come for different purposes, continuously
EFFECT: “sidewalk ballet” and “eyes on the street”
Small blocks

<table>
<thead>
<tr>
<th>LAND USE</th>
<th>SMALL BLOCKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGED BUILDINGS</td>
<td>DENSITY</td>
</tr>
</tbody>
</table>

City blocks should be small/short

JACOBS’ VIEW: improves *walkability*

EFFECT: Increase face-to-face interactions
Aged buildings

Buildings mixed (age and types)

JACOBS’ VIEW: To ensure economic diversity
EFFECT: high-/low-income residents new/small enterprises
## Density

<table>
<thead>
<tr>
<th>LAND USE</th>
<th>SMALL BLOCKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGED BUILDINGS</td>
<td>DENSITY</td>
</tr>
</tbody>
</table>

Concentration of people and enterprises

**Jacobs’ View:** People have a reason to live in a district

**Effect:** Attract people
Necessary, diversity conditions

**All** four factors are necessary
Border Vacuums

- Patches of land dedicated to one single use
- They could be either bad and good:
  - **Parks** are good for pedestrian activity
  - But they are *exposed to criminality* and deprivation if not well managed (e.g. night)
“Operationalize” Land Use Mix

For district $i$:

$$LUM_i = - \sum_{j \in N} \frac{P_{i,j} \log(P_{i,j})}{\log |N|}$$

$P_{i,j}$: % square footage of land use $j$

$N$: {residential, commercial, recreation}

**EFFECT**: The higher, the better.

Ref: R. Cervero. Land-use mixing and suburban mobility. University of California Transportation Center, 1989
“Operationalize” Small blocks

<table>
<thead>
<tr>
<th>LAND USE</th>
<th>SMALL BLOCKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGED BUILDINGS</td>
<td>DENSITY</td>
</tr>
</tbody>
</table>

Street intersections are a proxy for:
- small blocks
- peoples’ interactions

For district $i$:

$$\frac{|\text{Intersections}_i|}{\text{area}_i}$$

**EFFECT**: The higher, the better
Aged buildings

Aged buildings are supposed to be a proxy for new, small enterprises.

For district \( i \):

\[
\frac{1}{|F_i|} \sum_{j \in F_i} \text{employees}(j)
\]

\( F_i \): set of companies

**EFFECT**: The higher, the worse
“Operationalize” Density

For district $i$:

Employment density:

$$\frac{|\text{Employed people}_i|}{\text{area}_i}$$

Population density:

$$\frac{|\text{Population}_i|}{\text{area}_i}$$

**EFFECT**: The higher, the better
“Operationalize” Vacuums

<table>
<thead>
<tr>
<th>LAND USE</th>
<th>SMALL BLOCKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGED BUILDINGS</td>
<td>DENSITY</td>
</tr>
</tbody>
</table>

Distance to highways for district $i$:

$$\frac{1}{|B_i|} \sum_{j \in B_i} \text{dist}(j, \text{closest}(j, V))$$

$B_i$: the set of the blocks  
$V$: the set of highways

**EFFECT:** The higher, the better
Diversity → Vitality
“Operationalize” Vitality

- Mobile phone Internet activity as a proxy for urban vitality
- We calculate the activity density in each district \( i \)

\[
\frac{1}{|H|} \sum_{h \in H} |internet\ records_i|
\]

\( H \): set of hours (180 days x 24h)

- Six Italian cities with \( 100,000+ \) inhabitants (e.g. Rome, Milan…)
- 6 months time span (in 2014)
Jacobs’ theory holds and is still valid
...But something is different
STEP 4

Why does it matter?
Web data and mobile phone records offer insights on how most urban dwellers experience entire cities
Why does it matter?

• Evaluate the districts vitality
• Know in advance the best places for retails
• Quantifying regulatory interventions
• We created the recipe for city that works
Which place looks safer?
The theory: Broken Windows

- City mismanagement
- Dirty places
- Poor infrastructure

Lead to misbehavior => Crime

The theory: Jane Jacobs + Oscar Newman

Two of the most influential books in city planning

- Lit streets
- Street-facing windows
- Physical demarcation private-public

Appearance and liveliness
The multi-modal approach

SAFETY PERCEPTION

LIVELINESS
Safety perception: MIT Place Pulse
Safety perception: MIT Place Pulse

Place Pulse
- New York
- Boston

Place Pulse 2
- Rome
- Milan

PROBLEM:
- Just thousands of examples
- Sparse in space and # votes
Safety perception: fix sparse votes

- AlexNet CNN trained on Places205
- Transfer on Place Pulse
- Data augmentation

<table>
<thead>
<tr>
<th>Model type</th>
<th>State of art</th>
<th>Our model</th>
</tr>
</thead>
<tbody>
<tr>
<td>NY - NY</td>
<td>0.687</td>
<td>0.718</td>
</tr>
<tr>
<td>Boston - Boston</td>
<td>0.718</td>
<td>0.744</td>
</tr>
<tr>
<td>NY - Boston</td>
<td>0.701</td>
<td>0.734</td>
</tr>
<tr>
<td>Boston - NY</td>
<td>0.636</td>
<td>0.693</td>
</tr>
</tbody>
</table>

Safety perception: fix sparse images

- From 6.7 to 100 images per km^2
- 360 degrees of images
- Prediction on all images
- Aggregation per district
Safety perception: aggregation
Liveliness: mobile phone data

- Anonymized mobile phone calls activity as a proxy for urban liveliness
- Call density in each district $i$

$$\frac{1}{|H|} \sum_{h \in H} |\text{call records}_i|$$

$H$: set of hours (180 days x 24h)

- Broken down by gender, age
- **3 months** time span (in 2015)
- Rome and Milan
Liveliness: metrics

- People around in district $i$
  \[
  \frac{|people_{i,24h}|}{|area_i|}
  \]
- Fraction females
  \[
  \frac{|females_{i,24h}|}{|people_{i,24h}|}
  \]
- People below 30 (and above 50)
  \[
  \frac{|people(< 30)_{i,24h}|}{|people_{i,24h}|}
  \]
Link: regression

SAFETY PERCEPTION

LIVELINESS

MULTI-MODAL APPROACH
The Linear Regression

\[ y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{29} x_{29} + \alpha \]

- Population density
- Perceived safety
- Livelihood metric
Spatial Regression

“everything is related to everything else, but near things are more related than distant things.”

Tobler's first law of geography

- We control for spatial auto-correlation

The multi-modal approach

**PLACE PULSE**

SAFETY PERCEPTION

(SPATIAL) OLS REGRESSION

LIVELINESS

MULTI-MODAL APPROACH 1 2
Safety perception and liveliness

1. Presence of people
2. Women, Young people around
3. Elements for safety
Safety perception <-> presence of people

<table>
<thead>
<tr>
<th>Urban metric</th>
<th>Beta coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population density</td>
<td>0.155**</td>
</tr>
<tr>
<td>Employees density</td>
<td>0.328**</td>
</tr>
<tr>
<td>Deprivation</td>
<td>-0.022</td>
</tr>
<tr>
<td>Distance from the center</td>
<td>-0.257**</td>
</tr>
<tr>
<td>Safety appearance</td>
<td>0.105**</td>
</tr>
</tbody>
</table>

$adj - R^2 = 0.91$

** p-value < 0.001; * p-value < 0.01;
Safety perception <-> women, young people

<table>
<thead>
<tr>
<th>Urban metric</th>
<th>Beta coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of women (from census)</td>
<td>0.001</td>
</tr>
<tr>
<td>Deprivation</td>
<td>-0.005</td>
</tr>
<tr>
<td>Distance from the center</td>
<td>-0.003</td>
</tr>
<tr>
<td>Safety appearance</td>
<td>0.020**</td>
</tr>
</tbody>
</table>

\[adj - R^2\] 0.65

** p-value < 0.001; * p-value < 0.01;
Safety perception <-> women, young people

<table>
<thead>
<tr>
<th>Urban metric</th>
<th>Beta coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of people &lt; 30 years (from census)</td>
<td>-0.001</td>
</tr>
<tr>
<td>Deprivation</td>
<td>0.032**</td>
</tr>
<tr>
<td>Distance from the center</td>
<td>-0.150**</td>
</tr>
<tr>
<td>Safety appearance</td>
<td>-0.048**</td>
</tr>
<tr>
<td>adj – $R^2$</td>
<td>0.66</td>
</tr>
</tbody>
</table>

** p-value < 0.001; * p-value < 0.01;
Safety perception <-> elements
Street facing windows and greenery contributes positively to the appearance of safety.
Why does it matter?
Now that we have new tools to measure aesthetics, we can estimate its consequences ... to understand the relative value of improving the aesthetics of neighbourhoods
Thanks