Day 3. GIS-Based Trade Area Analysis and Application in Regional Planning

Fahui Wang
CNU
June 27, 2007 (Wednesday)
Outline

- Measures of Distance
- Computing Network Distance and Time
- Case Study 3A: Measuring Distances between Counties and Major Cities in Northeast China
- Basic Methods for Trade Area Analysis
- Gravity Models for Delineating Trade Areas
- Case Study 3B: Defining Hinterlands of Major Cities in Northeast China
Measures of Distance

- **Euclidean distance**
  \[ d_{12} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \]

- **Geodetic distance** through a great circle of the earth
  \[ d_{12} = r \times a \cos[\sin b \times \sin d + \cos b \times \cos d \times \cos(c - a)] \]

- **Manhattan distance**
  \[ d_{12} = |x_1 - x_2| + |y_1 - y_2| \]

- **Network distance**: shortest-path through a real-world road network
Computing Network Distance and Time

- *label setting algorithm* by Dijkstra (1959)
- *valued-graph* (or *L*-matrix) method

- NODEDISTANCE command in ArcInfo computes distances between nodes on the network
label setting algorithm
**L-matrix method**

Two-step connection 1–3
\[(1,1) + (1,3) = 0 + M = M\]
\[(1,2) + (2,3) = 116 + 113 = 229\]
\[(1,3) + (3,3) = M + 0 = M\]
\[(1,4) + (4,3) = M + 76 = M\]
\[(1,5) = (5,3) = 155 + M = M\]
Case Study 3A (Part 1): Measuring Euclidean and Manhattan Distances

- Generating county centroids
- Computing Euclidean distances
- Optional: Computing Manhattan distances
Case Study 3A (Part 2): Measuring Railway Distances

three segments:

(1) a county to its closest node
(2) network distance between nodes
(3) a city to its closest node
Implementation in ArcInfo

- Building network topology (line & node)
- Computing air distances between counties/cities and their nearest nodes
- Identifying unique origin and destination nodes
- Computing network distances between nodes (see AML)
- Joining & summing up distance segments together
AML Program

- `Dbaseinfo sum_fid.dbf tmp /*convert to INFO file ‘tmp’`
- `Pullitems tmp fm_node near_fid /*extract the item ‘near_fid’ to create INFO file ‘fm_node’ for origin nodes`
- `Pullitems city4.pat to_node near_fid /*extract the item ‘near_fid’ to create INFO file ‘to_node’ for destination nodes`
- `ap /* access the arcplot module`
- `netcover railne railroute /* set up the route system`
- `centers fm_node /* define the origin nodes`
- `stops to_node /* define the destination nodes`
- `nodedistance centers stops rdist 3000000 network ids`
- `q /*exit`

*Use the command IMPEDANCE if computing travel time*
Table Joins in Computing Travel Distances
join w/o air distance matrix?

### One-to-one

<table>
<thead>
<tr>
<th>County</th>
<th>O_Node</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>1</td>
</tr>
<tr>
<td>X2</td>
<td>1</td>
</tr>
<tr>
<td>X3</td>
<td>2</td>
</tr>
<tr>
<td>X4</td>
<td>3</td>
</tr>
</tbody>
</table>

### Many-to-many

<table>
<thead>
<tr>
<th>Link</th>
<th>County</th>
<th>City</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1A</td>
<td>X1</td>
<td>A</td>
<td>10.24</td>
</tr>
<tr>
<td>X2A</td>
<td>X2</td>
<td>A</td>
<td>20.45</td>
</tr>
<tr>
<td>X3A</td>
<td>X3</td>
<td>A</td>
<td>18.99</td>
</tr>
<tr>
<td>X4A</td>
<td>X4</td>
<td>A</td>
<td>7.66</td>
</tr>
<tr>
<td>X1B</td>
<td>X1</td>
<td>B</td>
<td>5.44</td>
</tr>
<tr>
<td>X2B</td>
<td>X2</td>
<td>B</td>
<td>5.78</td>
</tr>
<tr>
<td>X3B</td>
<td>X3</td>
<td>B</td>
<td>3.22</td>
</tr>
<tr>
<td>X4B</td>
<td>X4</td>
<td>B</td>
<td>5.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Link</th>
<th>O_Node</th>
<th>D_Node</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>110</td>
<td>1</td>
<td>10</td>
<td>11.24</td>
</tr>
<tr>
<td>210</td>
<td>2</td>
<td>10</td>
<td>24.45</td>
</tr>
<tr>
<td>310</td>
<td>3</td>
<td>10</td>
<td>23.44</td>
</tr>
<tr>
<td>120</td>
<td>1</td>
<td>20</td>
<td>7.88</td>
</tr>
<tr>
<td>220</td>
<td>2</td>
<td>20</td>
<td>9.34</td>
</tr>
<tr>
<td>320</td>
<td>3</td>
<td>20</td>
<td>5.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D_Node</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>A</td>
</tr>
<tr>
<td>20</td>
<td>B</td>
</tr>
</tbody>
</table>
Trade area analysis

- “the geographic area from which the store draws most of its customers”
- analog method
- regression model
- proximal area method
  - consumers-based ("NEAR" tool)
  - stores-based (Thiesscn polygons)
Gravity models: Reilly’s Law

Breaking point $x$:

\[
\begin{align*}
    d_{1x} + d_{2x} &= d_{12} \\
    S_1 / d_{1x}^2 &= S_2 / d_{2x}^2
\end{align*}
\]

\[
\begin{align*}
    d_{1x} &= d_{12} / (1 + \sqrt{S_2 / S_1}) \\
    d_{2x} &= d_{12} / (1 + \sqrt{S_1 / S_2})
\end{align*}
\]
Gravity models: Huff model

- Huff (1963) model:
  \[ P_{ij} = S_j d_{ij}^{-\beta} / \sum_{k=1}^{n} (S_k d_{ik}^{-\beta}) \]

- Reilly’s law is a special case of Huff model (\(\beta=2\) and \(n=2\))

- Extension 1: multiplicative competitive interaction (MCI) model
  \[
  P_{ij} = (\prod_{l=1}^{L} A_{ij}^{\alpha_l}) d_{ij}^{-\beta} / \sum_{k \in N_i} [\prod_{l=1}^{L} A_{lk}^{\alpha_l}] d_{ik}^{-\beta}
  \]

- Extension 2: multinomial logit (MNL) model
  \[
  P_{ij} = (\prod_{l=1}^{L} e^{\alpha_l A_{ij}}) e^{-\beta_{ij} d_{ij}} / \sum_{k \in N_i} [(\prod_{l=1}^{L} e^{\alpha_{ilk} A_{lk}}) e^{-\beta_{ik} d_{ik}}]
  \]
Deriving the $\beta$ Value in the Gravity Models

- **Special case:**

  \[ T_{ij} = a O_i D_j d_{ij}^{-\beta} \]

  \[ \ln \left[ \frac{T_{ij}}{(O_i D_j)} \right] = \ln a - \beta \ln d_{ij} \]

- **General case:**

  \[ T_{ij} = a O_i^{\alpha_1} D_j^{\alpha_2} d_{ij}^{-\beta} \]

  \[ \ln T_{ij} = \ln a + \alpha_1 \ln O_i + \alpha_2 \ln D_j - \beta \ln d_{ij} \]
Case Study 3B: Defining Hinterlands of Major Cities in Northeast China

Part 1. Defining Proximal Areas

- Extracting distances between counties and their closest cities
- Identifying the closest cities
- Mapping the proximal areas

![Map of Northeast China showing proximal areas for major cities Harbin, Dalian, Shenyang, and Changchun.](image)
Case Study 3B (Part 2): by Huff Model

- Measuring potential
- Identifying cities with the highest potential
- Mapping hinterlands of major cities
Discussion:

- How to measure $d$?
  - in what modes?
  - Physical vs. behavioral distance
- How to measure $S$?
- How to measure $\beta$?
- Applications in business geography