A classification algorithm based on density applied to municipality zoning and infrastructure performance measure

Li He
PhD at Polytechnique Montréal, Canada
Strategic information advisor - Metropolitan Regional Transportation Authority - Montréal

Co-authors:
Martin Trépanier, Bruno Agard (Polytechnique Montréal & CIRRELT)
Marcela Munizaga, Benjamin Bustos (Universidad de Chile)
1 - Problematic, objective and basic model (spatio-temporal classification method)

2 - Proposed improvement (Density-based classification method)
   **Method 1: Municipality zoning algorithm** – by using spatio-temporal classification result
   **Method 2: Infrastructure performance measure algorithm** – with the first daily transition time

3 - Result and analysis Visualization of data

4 - Conclusion
1 – Problematic, objective and basic model
Objective (before): classification of the daily spatio-temporal behaviors

Part 1: Problematic and objective, basic model

User 1
Home

User 2
Home

User 3
Home

The Euclidean distance of coordinates

School

Home

Work

Super market

Transaction time
06:30 07:00 16:00 18:00
Objective (in this research): classification of zones

Suppose that users are randomly located in three zones respectively, how to classify these zones?

Part 1: Problematic and objective, basic model
Basic model – classification of users’ behaviors

Daily space-time path of each cluster
Basic model – Daily space-time path of 100 users

Part 1: Problematic and objective, basic model
Basic model - Define work location by cutting space-time path

Part 1: Problematic and objective, basic model

P.m. peak hours
Cut at 11 a.m. to see work locations

A.m. peak hours
Cut at 3 a.m. to see home locations

11 h
3 h
Basic model - Home / work location for 100 users

Part 1: Problematic and objective, basic model
Basic model - Home / work location for all the users

Part 1: Problematic and objective, basic model

3 a.m.  
11 a.m.
2 - Proposed improvement
### Methodology – Basic model and proposed improvement

<table>
<thead>
<tr>
<th>Type of algorithm</th>
<th>Object (based on whom the density is calculated)</th>
<th>Clusters <em>a priori</em></th>
<th>Extra process</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic model</td>
<td>Home location</td>
<td>NA (For all the users together)</td>
<td>NA</td>
<td>Home location density</td>
</tr>
<tr>
<td>Algorithm 1</td>
<td>Home location</td>
<td>For each clusters of spatio-temporal classification</td>
<td>Choose a dominant cluster for a zone</td>
<td>Classification of zones based on users’ clusters</td>
</tr>
<tr>
<td>Algorithm 2</td>
<td>Location of first daily transaction</td>
<td>For each period of time</td>
<td></td>
<td>Classification of zones based on first daily transaction</td>
</tr>
</tbody>
</table>
Methodology – Choose a dominant cluster for a zone

One user on all zones

Zone A

User 1
User 1 on zone A: 100%

Zone B

User 1 on zone B: 50%

User 1 on zone C: 20%

Zone C

Part 2: Proposed improvement
Methodology – Choose a dominant cluster for a zone

One user on all zones

Zone A
User 1
User 1 on zone A: 100%

Zone B
User 1 on zone B: 50%

Zone C
User 1 on zone C: 20%

User 1 on all zones

All users on all zones

Zone A
User 1
User 1 on zone A: 50%
User 2 on zone A: 40%
User 3 on zone A: 10%

Zone B
User 2

Zone C
User 3
Methodology – Choose a dominant cluster for a zone

One user on all zones

Zone A
User 1
User 1 on zone A: 100%
User 1 on zone B: 50%
User 1 on zone C: 20%

Zone C

All users on all zones

Zone A
User 1
User 1 on zone A: 50%
User 2 on zone A: 40%
User 3 on zone A: 10%

Cluster of user 1 represents zone A

Zone B
User 2

Zone C
User 3

Part 2: Proposed improvement
Methodology – Application for all the zone

100 X 100 = 10000 zones
3 - Result and analysis
Result – Classification of zones based on the first transaction time – aggregated
Part 3: Result and analysis

Result – Classification of zones based on the first transaction time – disaggregated

Infrastructure performance measure

Line 6
Conclusion

• Contribution: Density-based classification algorithm for city zones -- municipality zoning and infrastructure performance measure

• Limitation: Only available for smart card users

• Perspective: Applied to more cases – fare system, cash deposit location selection...