

Understanding road network dynamics : Link-based topological patterns

Susana Freiria, Bernardete Ribeiro, Alexandre O. Tavares.,
Journal of Transport Geography, Volume 46, June 2015, pp. 55-66

2015/06/19 理論談話会 M2 芝原

まとめると…

- ネットワーク指標はそれぞれ表す特徴が異なる.
- いろいろ混ぜてクラスター分けを行う手法を適用.
- その成分比からクラスターに何が効いているか分析.
- 高速道路よりも下道の方が, ネットワーク寸断に効く道がある.
- むしろ切った方がネットワーク全体には良い道もある.

0. Contents

1. Introduction
2. Literature review
3. Methodology
 - 3-1. Structural characterizaion of the network
 - 3-2. Biclustering technique
 - 3-3. Mean geodesic distance variation due to link interruptions
4. Case study
5. Results
 - 5-1. Application of biclustering technique
 - 5-2. Mean geodesic distance variation due to link interruptions
6. Discussion
7. Conclusion

1. Introduction

- 災害によるネットワークの寸断を想定
 - もっとも重要な道路を特定したい。
- ネットワークの構造的特徴から重要な街路の位置を特定する
 - 交通量が多いから切られると弱いという分けでもない。
(迂回路が多く接続しているとか)
 - ネットワークの指標が表す意味はそれぞれ異なっている。
(次数と媒介中心性など)
- ネットワーク指標を重ね合わせて評価しよう。

3. Methodology

■ 分析の流れ

1. ネットワークの構造を表す指標の設定
2. Biclustering techniqueの適用
3. リンク寸断時のgeodesic distanceの算出, Biclusterとの比較

3. Methodology

1. ネットワークの構造を表す指標の設定

- 次数5のリンクは次数3のリンクより大きい. 次数だけでは全体の連結性は表せない. . .
- どの指標が道路の連結性のどの部分を表現できるかを想定すべき.

1. リンクとリンクの接続関係に直す.
2. 指標を計算し, 何を評価しうるかを想定する.

3. Methodology

1. ネットワークの構造を表す指標の設定

都市空間

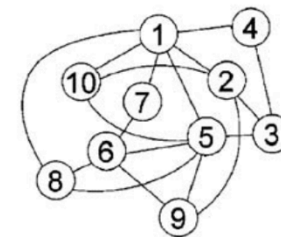
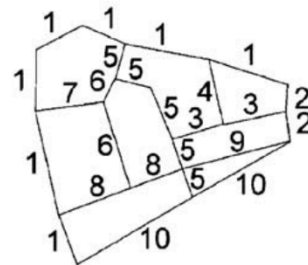
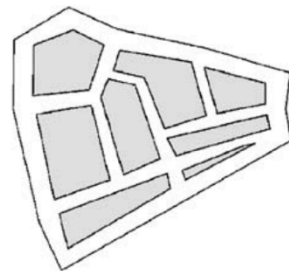
Primal graph

Dual graph

■ 本研究では同じ街路に同じIDを与えるnamed-street approachを用いる。

named-street approach

同じ名前の通りをつなぐ手法



■ ICNモデル (Porta et al., 2006) は逐次更新された歴史的都市では不自然な結果になってしまう (Masucci et al., 2014)

3. Methodology

1. ネットワークの構造を表す指標の設定

- 7つの指標を考える（目的に応じて変えるべき．実際に効いていたのは1. α index, 3. Bonacich Power, 6. Degreeの3つ）．

1. α index
2. Betweenness
3. Bonacich Power
4. Cluster Index
5. Cut level
6. Degree
7. fragmentation

3. Methodology

1. ネットワークの構造を表す指標の設定

- α index : ネットワーク内のcircuitの多さを示す

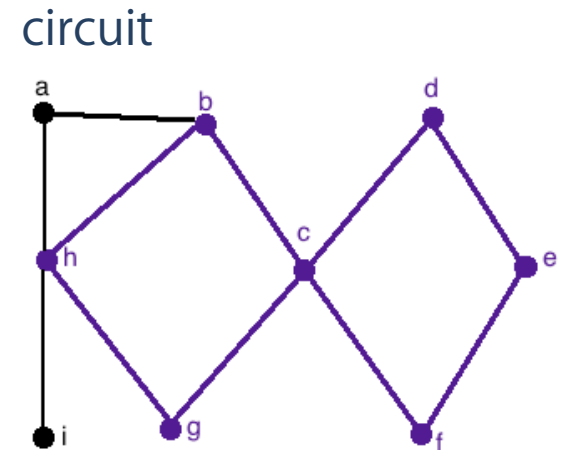
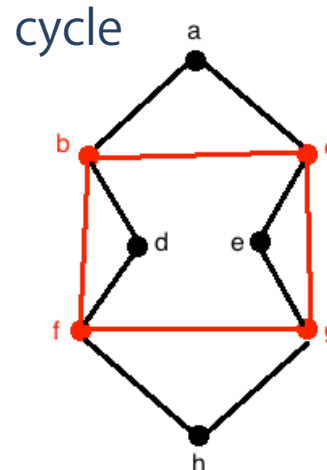
ノード*i*を起点・終点とするcircuit数

$$\alpha_i = \frac{\text{ノード } i \text{ を起点・終点とする circuit 数}}{\text{ノード } i \text{ を起点・終点とする, 完全グラフの circuit 数}}$$

$$= \frac{l - n + 1}{2n - 5}$$

l : リンク数

n : ノード数



3. Methodology

1. ネットワークの構造を表す指標の設定

■ Bonacich Power : 中心性の高いノードにつながるノードが高い中心性を持つ, とする指標

■ Degree : ノード*i*につながるリンクの数.

$$BP_i = \alpha(I - \beta R)^{-1} R1$$

$$d_i = \sum_j m_{ij}$$

α : スケールを合わせるための係数

R : グラフの隣接行列

I : 単位行列

m_{ij} : 隣接行列の*ij*成分

β : 隣接行列の重み付けパラメータ

1 : 全要素1の行列

3. Methodology

1. ネットワークの構造を表す指標の設定

- Betweenness : ネットワーク内の最短経路に位置する割合

$$C_B(n_i) = \sum_{j < k} \frac{g_{jk}(n_i)}{g_{jk}}$$

$g_{jk}(n_i)$: ノード j から ノード k への最短経路数のうち、
ノード n_i を通る経路数

g_{jk} : ノード j から ノード k への最短経路数

3. Methodology

1. ネットワークの構造を表す指標の設定

- Cluster Index : ノード周辺のリンクのつながり具合を表す指標

$$C_i = \frac{M_i}{k_i(k_i - 1) / 2}$$

M_i : ノード*i*のサブグラフに含まれるリンク数

k_i : ノード*i*の次数

3. Methodology

1. ネットワークの構造を表す指標の設定

- Cutpoints : コミュニティAとBのつながりを示す.

$$cut(A, B) = \sum_{i \in A, j \in B} w_{ij}$$

$i \in A, j \in B$: コミュニティ

w_{ij} : ij 間にリンクがあれば1, なければ0

3. Methodology

1. ネットワークの構造を表す指標の設定

- Cutpoints : つながっていないノードペアのネットワーク全体における比率

$$F_i = \frac{2 \sum_i \sum_{j < i} r_{ij}}{n(n-1)}$$

r_{ij} : ノード*i*からノード*j*に到達できれば1, できなければ0

3. Methodology

- Degreeはリンクの接続の強さを扱えない。全て等価に扱っている。
- Bonacich Powerはノードの重要性を表すが、重要な街路に接続する街路が重要であるという決め方。代替経路があれば、リンクが切れても大丈夫だが、その評価には使えない。
- α indexは全体のネットワーク連結性（代替経路の存在）を評価している。
- ケーススタディでは、これらの指標の多重共線性も確認する。

3. Methodology

2. Biclustering techniqueの適用

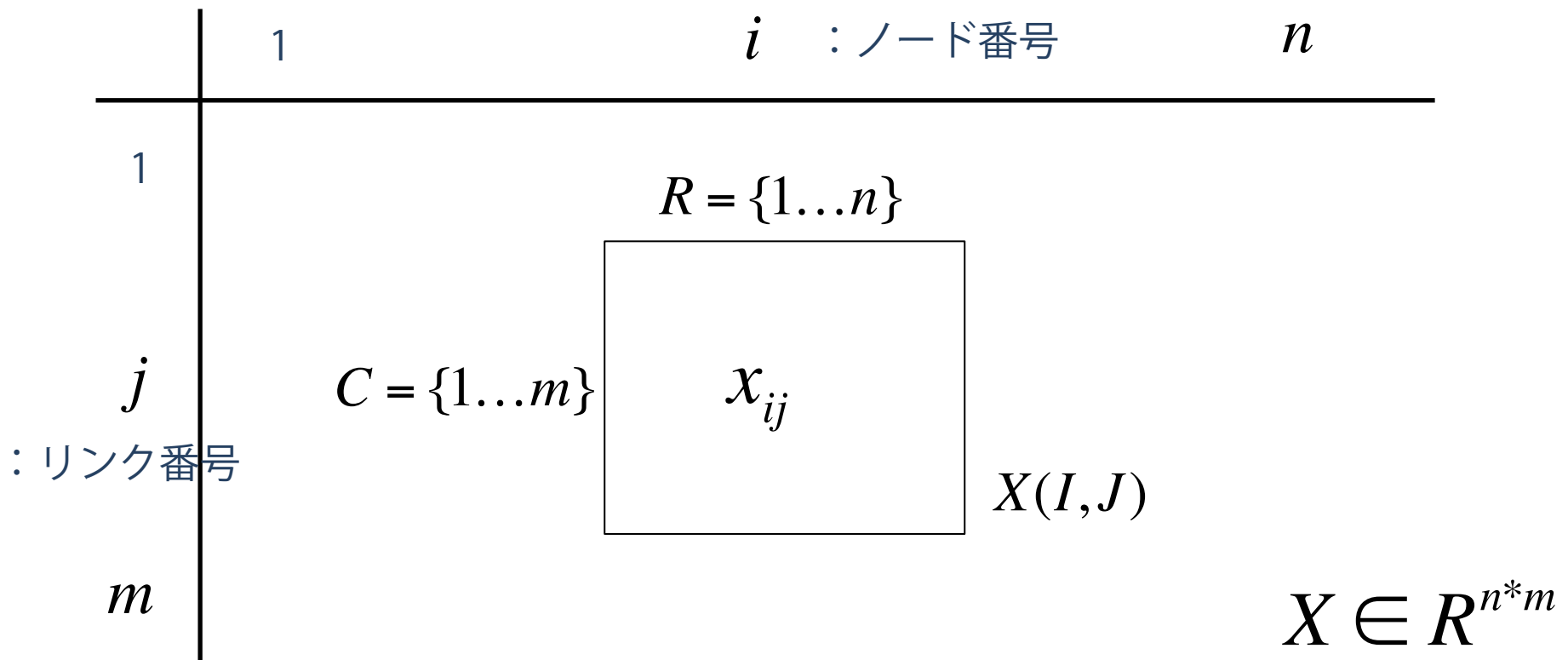
- 道路のコミュニティ分割（行列の列成分）
- 道路の特徴（行列の行成分）を同時に決定する。

- n 列（道路番号）， m 行の特徴量
- 列を近い関係のデータを分割する。
- Additive-relatedとmultiple-relatedの2種類がある。本研究ではadditive-relatedを用いる。

3. Methodology

2. Biclustering techniqueの適用

- n 個のノードそれぞれに m 個の属性があるとする。



3. Methodology

2. Biclustering techniqueの適用

■ 4つのパラメータを設定.

1. Noise threshold ϵ
2. Minimum percentage of rows N_r
3. Minimum numbers of columns N_c
4. Maximum bicluster overlap in percentage P_o

Table 1

Tests run: Phase II – parameter settings.

Parameter	Parameter settings
Noise threshold	5; 4; 3; 2; 1.5; 1; 0.5
Minimum number of columns	6; 5; 4; 3
Minimum percentage of rows	15; 10; 5; 4
Maximum bicluster overlap in percentage	25; 20; 10

3. Methodology

2. Biclustering techniqueの適用 $B_r = (I_e, J_r) \quad r = 1, \dots, t$

■ MSRS (Mean Square Residue Score) : 低いほど良い

$$MSRS = \frac{1}{nm} \sum_{i \in R, j \in C} (x_{ij} - x_{iC} - x_{Rj} + x_{RC})^2 \quad x_{iC} = \frac{1}{m} \sum_{j \in C} x_{ij} \quad x_{Rj} = \frac{1}{n} \sum_{i \in R} x_{ij}$$

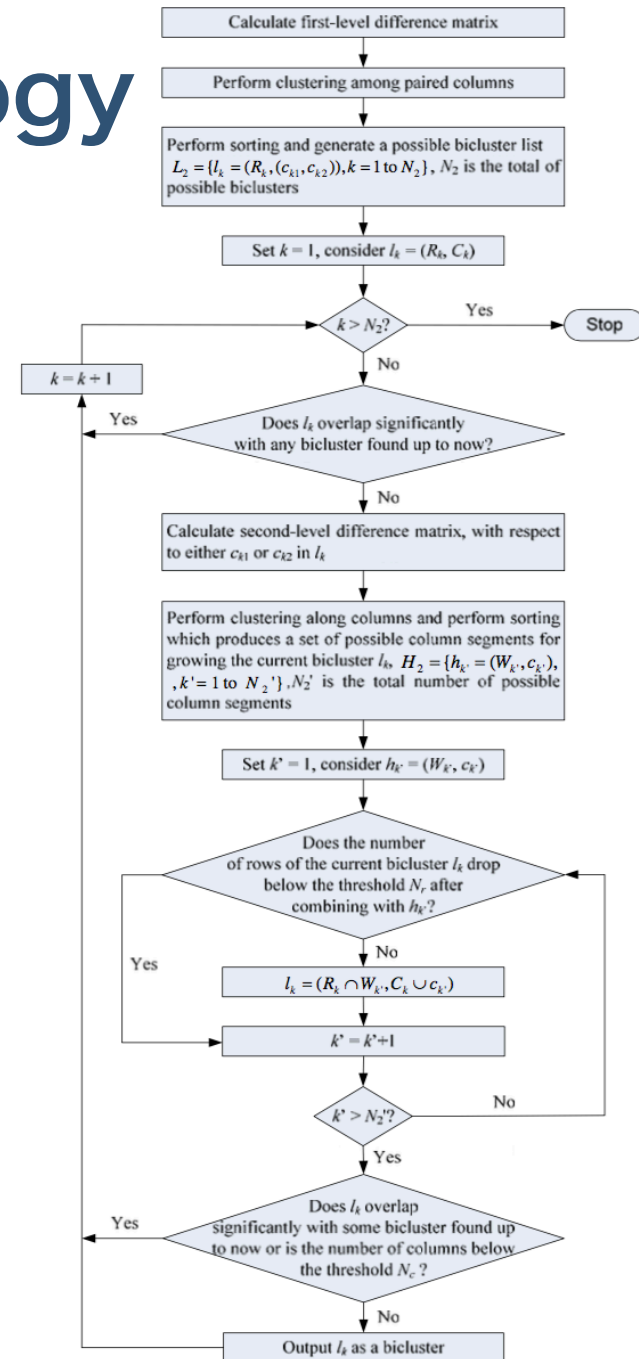
$$MSRS \leq \varepsilon \quad \text{なら, 正しいBicluster} \quad x_{RC} = \frac{1}{nm} \sum_{i \in R, j \in C} x_{ij}$$

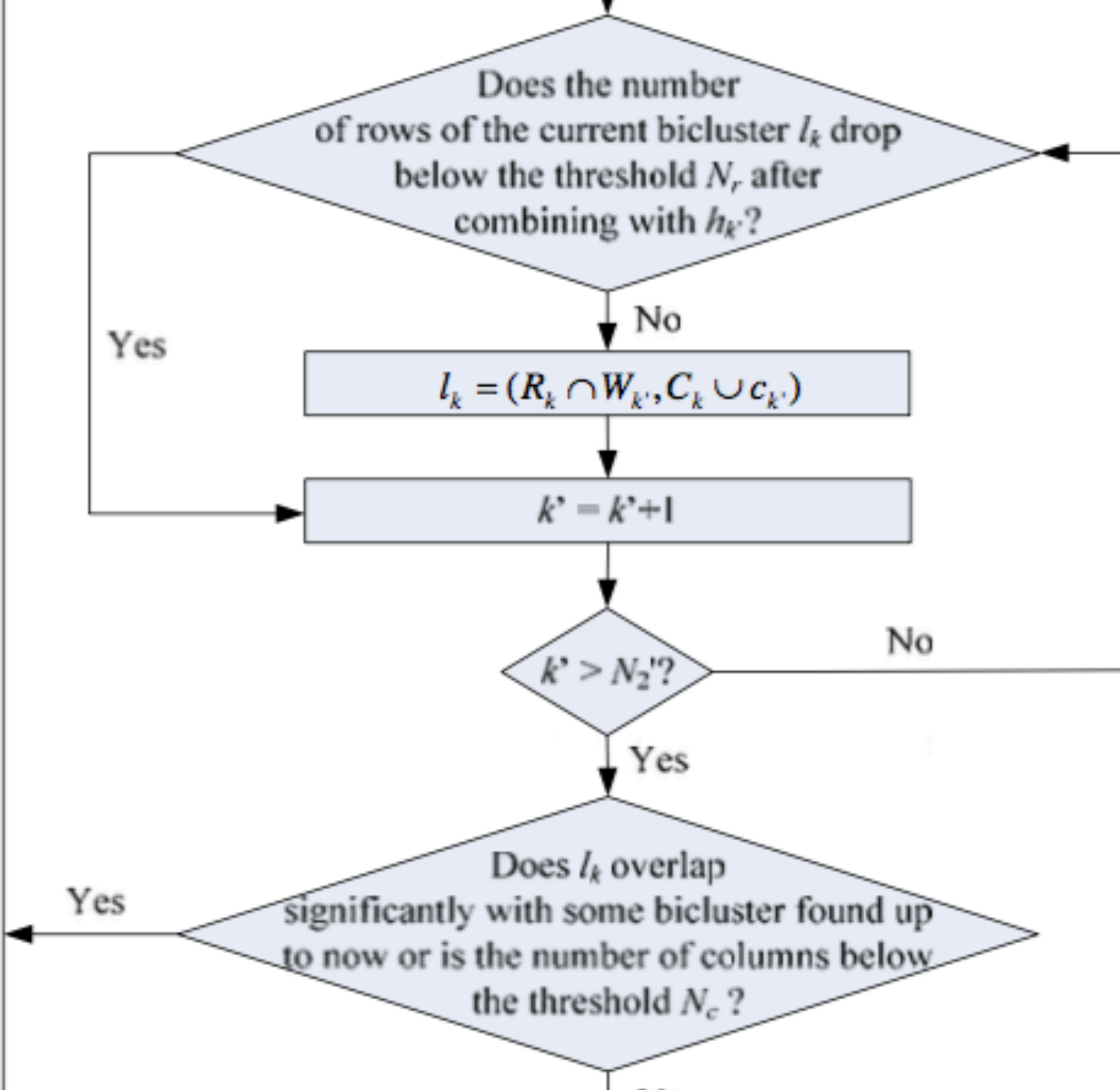
■ ACV (Average Correlation Value) : 高いほど良い

$$ACV = \max \left\{ \frac{\sum_{i,j \in R} |CorR_{ij}| - m}{n(n-1)}, \frac{\sum_{k,l \in C} |CorC_{kl}| - n}{m(m-1)} \right\}$$

$CorR_{ij}$ $CorC_{kl}$: 列 ij 間, 行 kl 間の相関係数

3. Methodology





3. Methodology

2. Biclustering techniqueの適用

■ Biclustering Algorithm

Algorithm 0 (Brute-Force Deletion and Addition).

Input: A , a matrix of real numbers, and $\delta \geq 0$, the maximum acceptable mean squared residue score.

Output: A_{IJ} , a δ -bicluster that is a submatrix of A with row set I and column set J , with a score no larger than δ .

Initialization: I and J are initialized to the gene and condition sets in the data and $A_{IJ} = A$.

Iteration:

1. Compute the score H for each possible row/column addition/deletion and choose the action that decreases H the most. If no action will decrease H , or if $H \leq \delta$, return A_{IJ} .

Algorithm 1 (Single Node Deletion).

Input: A , a matrix of real numbers, and $\delta \geq 0$, the maximum acceptable mean squared residue score.

Output: A_{IJ} , a δ -bicluster that is a submatrix of A with row set I and column set J , with a score no larger than δ .

Initialization: I and J are initialized to the gene and condition sets in the data and $A_{IJ} = A$.

Iteration:

1. Compute a_{iJ} for all $i \in I$, a_{Ij} for all $j \in J$, a_{IJ} , and $H(I, J)$. If $H(I, J) \leq \delta$, return A_{IJ} .
2. Find the row $i \in I$ with the largest

$$d(i) = \frac{1}{|J|} \sum_{j \in J} (a_{ij} - a_{iJ} - a_{Ij} + a_{IJ})^2$$

and the column $j \in J$ with the largest

$$d(j) = \frac{1}{|I|} \sum_{i \in I} (a_{ij} - a_{iJ} - a_{Ij} + a_{IJ})^2$$

remove the row or column whichever with the larger d value by updating either I or J .

3. Methodology

2. Biclustering techniqueの適用

■ Biclustering Algorithm

Algorithm 2 (Multiple Node Deletion).

Input: A , a matrix of real numbers, $\delta \geq 0$, the maximum acceptable mean squared residue score, and $\alpha > 1$, a threshold for multiple node deletion.

Output: A_{IJ} , a δ -bicluster that is a submatrix of A with row set I and column set J , with a score no larger than δ .

Initialization: I and J are initialized to the gene and condition sets in the data and $A_{IJ} = A$.

Iteration:

1. Compute a_{iJ} for all $i \in I$, a_{Ij} for all $j \in J$, a_{IJ} , and $H(I, J)$. If $H(I, J) \leq \delta$, return A_{IJ} .

2. Remove the rows $i \in I$ with

$$\frac{1}{|J|} \sum_{j \in J} (a_{ij} - a_{iJ} - a_{Ij} + a_{IJ})^2 > \alpha H(I, J)$$

3. Recompute a_{Ij} , a_{IJ} , and $H(I, J)$.

4. Remove the columns $j \in J$ with

$$\frac{1}{|I|} \sum_{i \in I} (a_{ij} - a_{iJ} - a_{Ij} + a_{IJ})^2 > \alpha H(I, J)$$

5. If nothing has been removed in the iterate, switch to Algorithm 1.

Algorithm 3 (Node Addition).

Input: A , a matrix of real numbers, I and J signifying a δ -bicluster.

Output: I' and J' such that $I \subset I'$ and $J \subset J'$ with the property that $H(I', J') \leq H(I, J)$.

Iteration:

1. Compute a_{iJ} for all i , a_{Ij} for all j , a_{IJ} , and $H(I, J)$.

2. Add the columns $j \notin J$ with

$$\frac{1}{|I|} \sum_{i \in I} (a_{ij} - a_{iJ} - a_{Ij} + a_{IJ})^2 \leq H(I, J)$$

3. Recompute a_{iJ} , a_{IJ} , and $H(I, J)$.

4. Add the rows $i \notin I$ with

$$\frac{1}{|J|} \sum_{j \in J} (a_{ij} - a_{iJ} - a_{Ij} + a_{IJ})^2 \leq H(I, J)$$

5. For each row i still not in I , add its inverse if

$$\frac{1}{|J|} \sum_{j \in J} (-a_{ij} + a_{iJ} - a_{Ij} + a_{IJ})^2 \leq H(I, J)$$

6. If nothing is added in the iterate, return the final I and J as I' and J' .

Algorithm 4 (Finding a Given Number of Biclusters).

Input: A , a matrix of real numbers with possible missing elements, $\alpha \geq 1$, a parameter for multiple node deletion, $\delta \geq 0$, the maximum acceptable mean squared residue score, and n , the number of δ -biclusters to be found.

Output: n δ -biclusters in A .

Initialization: Missing elements in A are replaced with random numbers from a range covering the range of non-null values. A' is a copy of A .

Iteration for n times:

1. Apply Algorithm 2 on A' , δ , and α . If the row (column) size is small (less than 100), do not perform multiple node deletion on rows (columns). The matrix after multiple node deletion is B .

2. (Step 5 of Algorithm 2) Apply Algorithm 1 on B and δ and the matrix after single node deletion is C .

3. Apply Algorithm 3 on A and C and the result is the bicluster D .

4. Report D , and replace the elements in A' that are also in D with random numbers.

3. Methodology

2. Biclustering techniqueの適用

- あと重回帰分析もやりました.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

$$R^2 = \frac{SSR}{\sum (y_i - y)^2}$$

3. Methodology

3. リンク寸断時のgeodesic distanceの算出

- 従来の評価は旅行時間増加やトリップ長増加が多い.
- 本研究ではgeodesic distanceを用いる.

$$\overline{G}_l = \frac{G_{in} G_{jn}}{SP_{ij}}$$

\overline{G}_l : ノード ij で表されるリンク l の平均最短経路距離
 G_{in} : ノード i 発, 任意ノード着の最短経路距離の合計
 G_{jn} : ノード j 発, 任意ノード着の最短経路距離の合計
 SP_{ij} : ノード ij 間の最短経路数

3. Methodology

3. リンク寸断時のgeodesic distanceの算出

$$\overline{G}_o = \frac{\sum_i \sum_j \overline{G_{in} G_{jn}}}{\sum_i \sum_j SP_{ij}}$$

original scenarioでの平均測地距離

$$\overline{G}_a = \frac{\sum_i \sum_j \overline{G_{ina} G_{jna}}}{\sum_i \sum_j SP_{ij}}$$

link aが切れたときの平均測地距離

$$\overline{L}_a = \overline{G}_a - \overline{L}_0$$

link aがないときの全体の距離増加分

3. Methodology

3. リンク寸断時のgeodesic distanceの算出

■ BiclusterとMean geodesic distanceを比較.

$$x = a + by$$

x : biclusterごとの平均距離増加分
 by : biclusterのregression result

$$x = a + bbp$$

bbp : biclusterごとの平均Bonacich Power

$$x = a + \alpha$$

α : biclusterごとの平均alpha index

$$x = a + d$$

d : biclusterごとの平均次数

当てはまりを決定係数から判断する.

4. Case study

- Coimbra in Portugal
- 総道路延長：509km
- 様々な道路で構成される都市圏.



Fig. 1. Case study location.

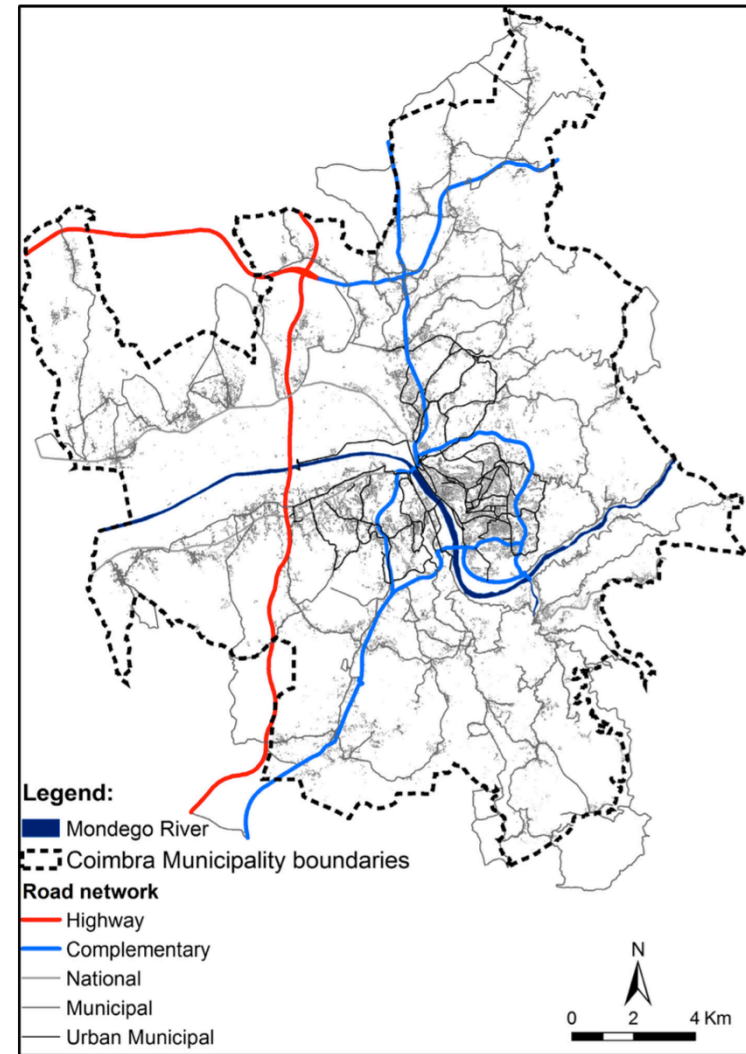


Fig. 4. Road hierarchy.

Table 2
Road hierarchy in percentage.

Designation	Length (km)	%
Highway road network	33	6
Complementary road network	55	9
National road network	36	6
Municipal road network	225	60
Urban municipal road network	180	19

5. Results

- 結果. 24%のリンクがクラス分けされずに残った. 似たような特徴を持っていないため.
- 重回帰から, alpha index:100%, Bonacich Power : 44%, Degree : 30%と出た.

Table 3

Biclusters in terms of their volume (road \times attributes), MSRS and ACV, length (km) and correspondent percentage.

Designation	Composition (roads \times attributes)	Quality criteria		Road's length (km)	Road's length (%)
		MSRS	ACV		
Bicluster 1	184 \times 7	0.29	0.98	179	30
Bicluster 2	85 \times 7	0.26	0.96	117	20
Bicluster 3	30 \times 7	0.18	0.92	44	8
Bicluster 4	35 \times 7	0.23	0.98	26	5
Bicluster 5	23 \times 7	0.17	0.92	15	3
Bicluster 6	30 \times 5	0.29	0.96	56	10
Unclassified	-	-	-	137	24
Total	-	-	-	574	100

5. Results

■ bicluster4に最重要のリンクが含まれている。平均値が全行で最も高い。全リンク長の5%ほどしかなく、ネットワークが集中している。

- | | |
|--|--------------------|
| 1. Noise threshold | $\varepsilon = -2$ |
| 2. Minimum percentage of rows | $N_r = -5$ |
| 3. Minimum numbers of columns | $N_c = 5$ |
| 4. Maximum bicluster overlap in percentage | $P_o = -10$ |

Table 5
Biclusters' results.

Designation	Bonacich Power		Alpha index		Degree		Regression analysis results	
	Mean	Std. deviation	Mean	Std. deviation	Mean	Std. deviation	Mean	Std. deviation
Bicluster 1	3.8	0.7	8.4	1.3	4.2	0.7	6.5	1.0
Bicluster 2	3.8	0.8	5.6	1.1	4.4	0.6	4.4	0.8
Bicluster 3	2.1	0.7	4.1	0.8	2.2	0.8	3.2	0.7
Bicluster 4	4.4	0.7	10.2	1.1	5.5	0.6	7.8	0.8
Bicluster 5	2.8	0.6	2.9	0.5	3.5	0.6	2.3	0.4
Bicluster 6	2.8	0.9	5.2	1.1	4.1	1.9	4.1	0.8
Unclassified	3.3	1.8	7.6	4.3	4.2	2.2	5.7	3.3

The blue cells represent the lowest values in each column and the red cells represent the highest values in each column.

5. Results

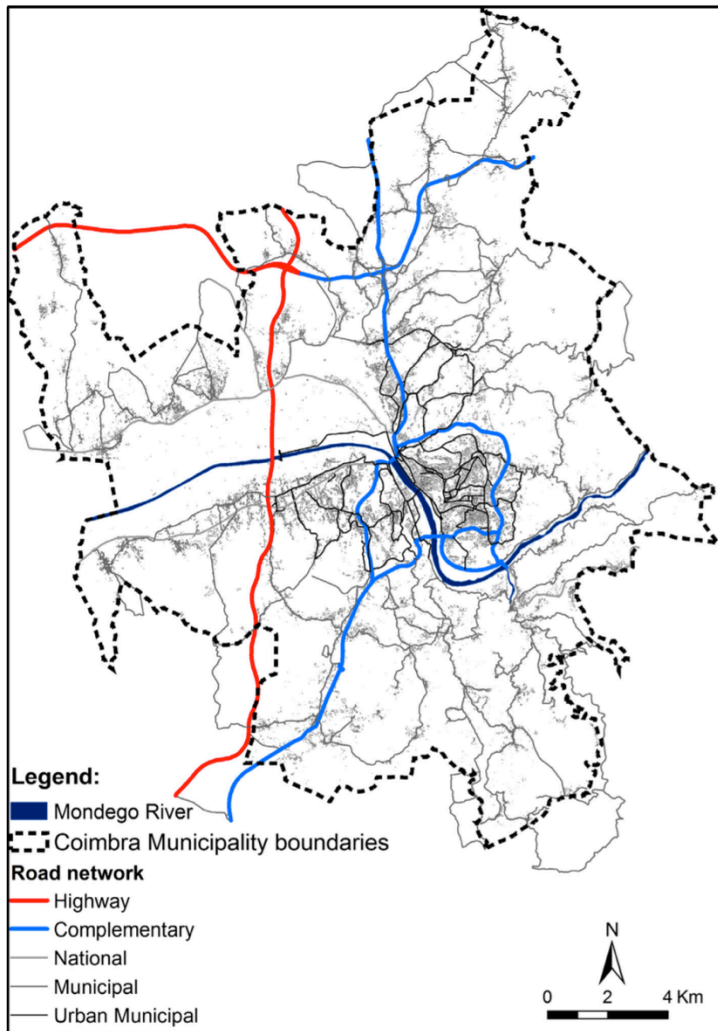


Fig. 4. Road hierarchy.

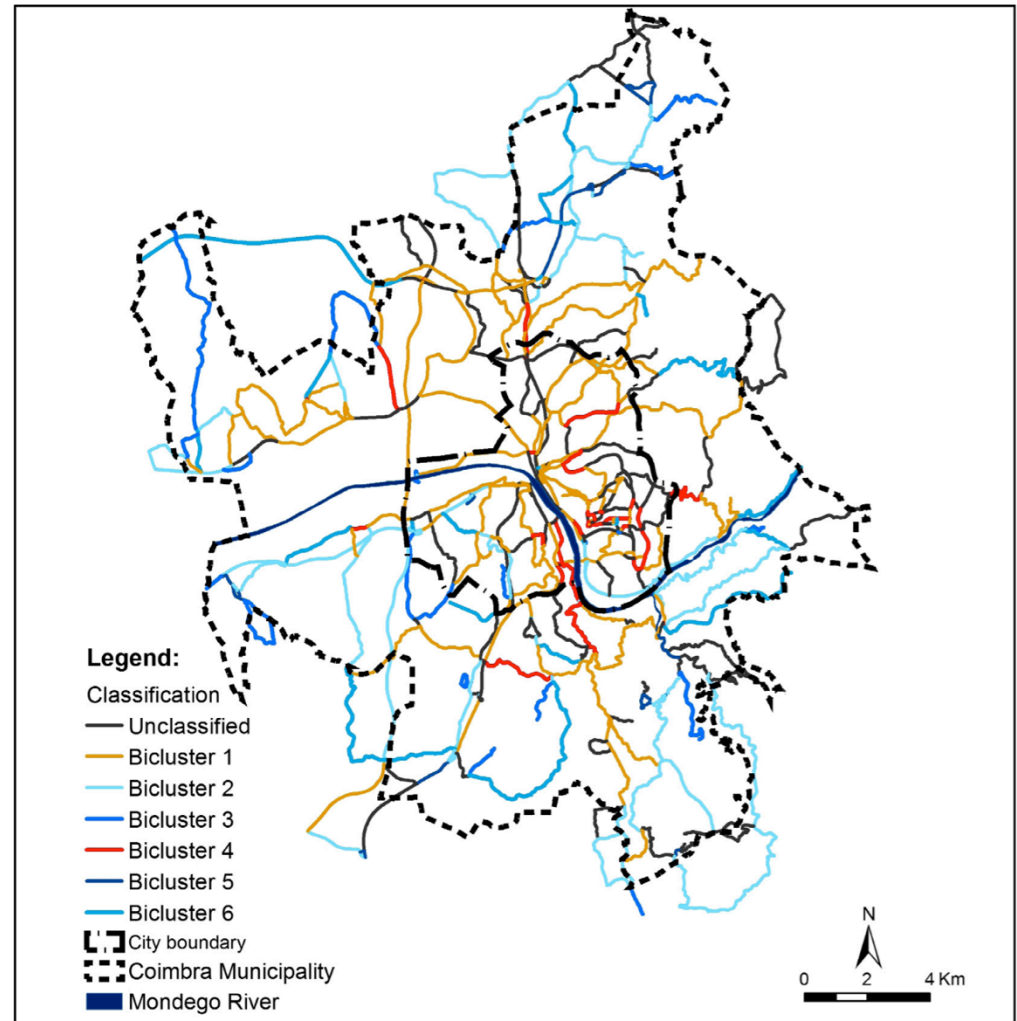


Fig. 5. Cartographic representation of the biclusters results.

5. Results

- bicluster3の道路は付随的な道路。自治体の分割で生じることが多い。
- 高速道路はヒエラルキーの頂点にあるが、重要なbiclusterには含まれていない。

Table 6
Average near geodesic distance among the links.

Designation	Biclusters' connectivity index	Average geodesic distance (km)
Bicluster 1	6.5	7.5
Bicluster 2	4.4	9.5
Bicluster 3	3.2	9.6
Bicluster 4	7.8	7.1
Bicluster 5	2.3	11.4
Bicluster 6	4.1	9
Unclassified	5.7	8.2
Total (average)	5.5	8.3

The blue cells represent the lowest values in each column and the red cells represent the highest values in each column.

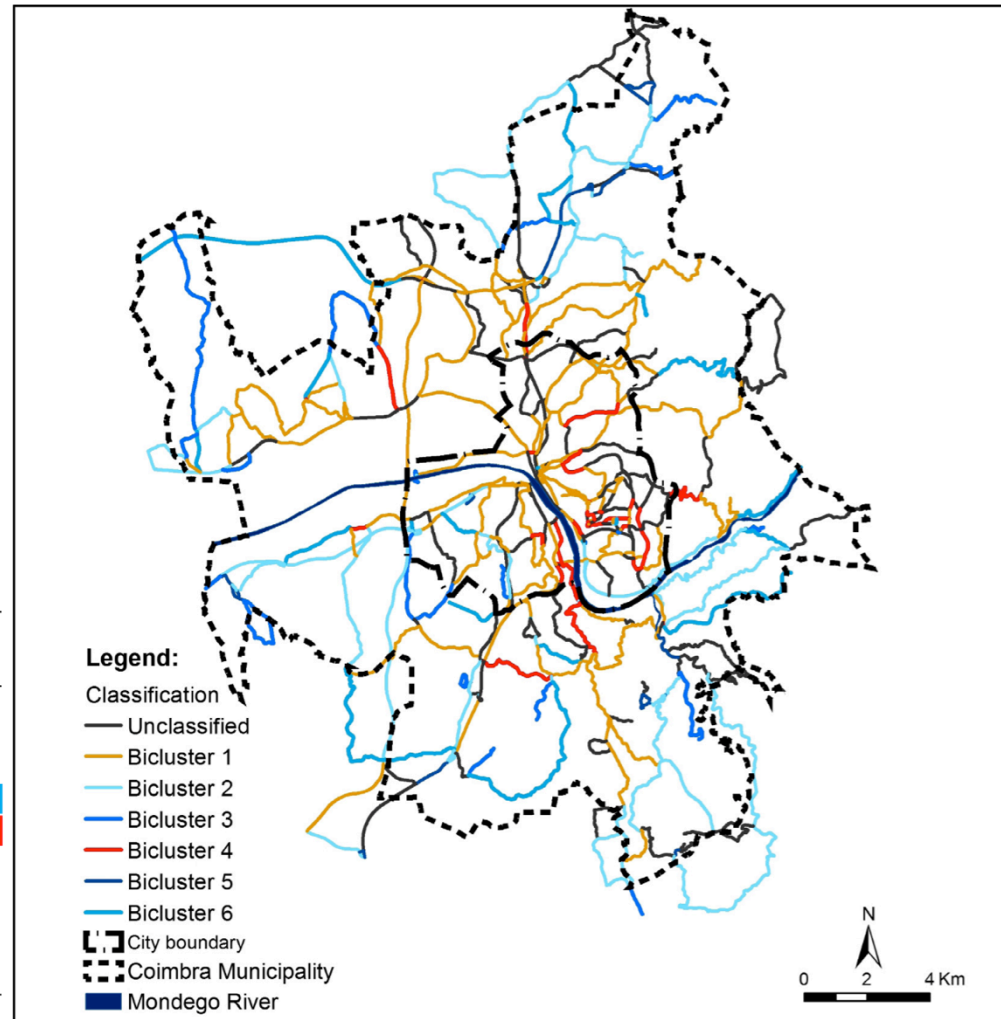


Fig. 5. Cartographic representation of the biclusters results.

5. Results

- bicluster4のリンクを切ると57m増える.
- bicluster3のリンクを切るとむしろスムーズになる。
(孤立した道が多い?)

Table 7

Mean geodesic distance variation by road network interruptions.

Designation	Biclusters' connectivity index	Average geodesic distance increase (m)	Maximum accumulated geodesic distance increase (m)
Bicluster 1	6.5	34	17,252
Bicluster 2	4.4	33	16,573
Bicluster 3	3.2	-15	-7795
Bicluster 4	7.8	57	28,755
Bicluster 5	2.3	4	2020
Bicluster 6	4.1	28	14,144

The blue cells represent the lowest values in each column and the red cells represent the highest values in each column.

6. Conclusion

- ネットワークの構造的特徴量を重ねて評価する手法を提案.
- Biclustering 手法を適用.
- その成分比と, geodesic distance との関係から切った方が良い道もある.