

市町村通勤データを用いた都市圏設定

福本潤也, 北野翔太

Setting of Urban Metropolitan Areas Using Intermunicipal Commuting Data

Junya FUKUMOTO, Shota KITANO

Abstract: “Urban Metropolitan Areas” (UMA, in short) is important both for academic research and urban/regional policy. Existing UMA delineation methods have a limitation that they cannot determine the number of UMAs automatically without assuming some ad hoc rule. We propose a new UMA delineation method that employs the measure proposed by Girvan and Newman (2003) and known as “modularity”. Then we apply it to the Japanese inter-municipal commuting data. From the results, we can confirm that the proposed method has a potential as a new UMA delineation method.

Keywords: 都市圏 (urban metropolitan areas), モジュラリティ (modularity), 市町村間通勤データ (intermunicipal commuting data)

1. はじめに

多くの統計データは市町村単位で集計されている。しかし、一般に社会経済活動は市町村の境界を越えて広がっている。社会経済活動が広がる地理的な範囲を都市圏と定義できる。都市圏を設定することで、学術的には次のような実証分析を行うことが可能になる。

- ・都市規模の空間的・時間的な比較
 - ・都市集積が顕著な資源の把握
 - ・都市内の資源の空間分布の把握
- また、以下のような都市政策を立案・実施していく上でも都市圏の設定が必要になる。
- ・雇用政策の立案と実施
 - ・住宅政策の立案と実施
 - ・都市生活の質の計測

多くのOECD諸国は公式な都市圏を設定して都市政策の立案や実施に活用している(OECD, 2002)。わが国には公式な都市圏は存在しない。しかし、様々な研究者が都市圏の設定に関する研究・提案を行っている(金本・徳岡, 2002)。

都市圏の設定手法は、クラスタリング・アプローチ、階層的アプローチ、ルールベース・アプローチの3つに大きく分かれる(Coombes, 2000)。クラスタリング・アプローチでは、最終的に設定される都市圏数を外生的に定めた上でクラスタリング手法が用いられる。階層的アプローチでは、階層的なクラスタリングによって都市圏が求められる。最終的に設定される都市圏数についてはデンドログラム等の分析結果を参照しながら、分析者の判断によって決める必要がある。ルールベース・アプローチでは、分析者が事前に定めたルールに従って複数の地理単位を統合して都市圏が設定される。都市圏数こそ内生的に決まるものの、分析者が都市圏を設定

するためのルールを恣意的に決めなければならぬ。

3つのアプローチに共通する限界として、最終的に求める都市圏数の設定方法に分析者の判断が介在せざるを得ない点が挙げられる。都市圏を設定する目的が明確であり、都市圏の設定に利用可能なデータが限定されている状況下では、既存のアプローチでも問題は生じないと考えられる。しかし、異なるデータを用いて設定されたそれぞれの都市圏を比較する場合、あるいは複数のデータを組み合わせて新たに都市圏を設定する場合には、分析者の判断が極力介在しないアプローチが求められる。多種多様な空間データが GIS 上に整備されている現在、分析者の判断が介在せずに最終的に求まる都市圏数が内生的に決まる都市圏設定手法を新たに開発する意義は大きいと考えられる。

以上の問題意識のもと、本研究では、ネットワーク科学の分野で発展した“modularity”指標を用いたグラフ・クラスタリング手法に基づく都市圏設定手法を提案する。提案手法の特徴は、“modularity”指標の最大化基準に従って最終的に求まる都市圏数が内生的に決まる点にある。さらに、提案手法を宮城県内の市区町村間通勤データに適用し、有効性を実証的に検証する。

2. 関連研究

(1) Functional Region

地理学の分野では古くから“functional region”(以下、FR)の設定手法に関する研究が行われてきた。Brown and Holmes (1971)は、FRを、1) 複数の空間要素から構成される地域であり、2) 地域内の空間要素間には地域外の空間要素との間よりも強い結びつきがあり、3) 空間要素同士が機能的に補完しあう地域であるとしている。彼らの定義に従えば、都市圏は FR の特殊ケースに過ぎない。実際、FR の典型例である都市圏の設定に関する研究は、FR の設定手法に関する研究の中で進展してきた。

FR の設定手法は 1. でも述べた通り、3 つのアプ

ローチに大別される。このうち、実務で広く用いられてきたのは、ルールベース・アプローチである。代表例としては、イギリス等で用いられている TTWAs (Travel to Work Areas) が挙げられる。一方、学術的な研究が最も多く行われてきたのは階層的アプローチである。1970 年代にはクラスタリング手法の発展を踏まえて、様々な手法が提案されている (Brown and Holmes, 1971; Masser and Brown, 1975; Slater, 1976)。階層的アプローチの特徴は、空間要素間の OD データ (通勤 OD や人口移動 OD 等) をグラフとみなした上で、クラスタリング手法を適用している点にある。ただし、後述する近年のグラフ・クラスタリング手法と比べると、適用している手法はアドホックである。

(2) グラフ・クラスタリング

グラフ・クラスタリングの手法は、計算機科学の分野とネットワーク科学の分野で発展してきた。前者では、分散処理のシステム構築等を目的に手法の開発が行われてきた。最適なクラスタリングを、対象とするネットワークの諸条件について分析者が比較的多くの情報を有する状況下でいかに発見するかが解くべき問題である。一方、ネットワーク科学が対象とするのは、社会ネットワークや WWW のネットワーク、生態ネットワーク等である。ネットワークを特徴づけるクラスタリングを、ネットワークの諸条件について分析者が十分な情報を有さない状況下でいかに発見するかが解くべき問題になる。計算機の技術進歩と歩調を合わせて近年、急速に進展してきたグラフ・クラスタリング手法を援用することで、新たな都市圏設定手法を開発することが可能になるとを考えられる。

(3) Modularity

グラフ・クラスタリング手法の一つに、Girvan and Newman (2003)が提案した“modularity”指標に基づく手法がある。“modularity”とは、ネットワーク内のある一つのリンクを選ぶ場合にあるクラスター (ノードの部分集合) 内のリンクが選ばれる確率 (A) と、リンクをランダムに配置した等価な

ネットワーク内からある一つのリンクを選ぶ場合に当該クラスタ内のリンクが選ばれる確率の期待値 (B) の差と定義される (Nweman, 2006). 対称・無向グラフの場合, modularity は次式で表される.

$$Q_l = p_{ll} - q_l^2 \quad (1)$$

$$p_{ll} = \frac{1}{2M} \sum_{i \in V_l, j \in V_l} a_{ij} \quad (2)$$

$$q_l = \frac{1}{2M} \sum_{i \in V_l, j \in V} a_{ij} \quad (3)$$

ただし, V はノード集合, l はクラスターのサフィックス, V_l はクラスター (ノードの部分集合), a_{ij} はノード i とノード j を結ぶリンクのダミー変数, M はリンク総数 ($M = \sum_{i,j} a_{ij}$) である. p_{ll} が上述の A に, q_l^2 が上述の B に対応する. 生起確率の同時分布と周辺分布の積の差になっている. ネットワーク内に複数のクラスタが存在する場合, ネットワーク内のクラスタの modularity の合計として, ネットワークの modularity が計算される.

$$Q = \sum_l Q_l \quad (4)$$

各クラスタのモジュラリティは正の値をとる場合も負の値をとる場合もある. モジュラリティが正の値をとる場合, ノードが密に結合したクラスタの存在を示唆している. ネットワークの modularity を最大化するノード分割を求めることでネットワークを特徴づけるクラスタの発見を試みるのが, modularity に基づくグラフ・クラスタリング手法である.

modularity の定義式は, 当初は対称・無向グラフについて定義されたが, その後, 他のグラフにも拡張されている. また, ノード分割を求めるアルゴリズムについても様々な拡張がなされている. 具体的には, ノードの全体集合を小さな部分集合に徐々に分割していく分割法や個々のノードを徐々に大きな部分集合へと結合していく結合法, 複数のクラスタの重複 (同一のノードが複数のクラスタへの所属) を認めない非重複法や重複を認める重複法などである. 市町村間通勤データを非対称・有向・重み付きのグラフと捉えると, modularity に基づくグ

ラフ・クラスタリング手法を基本的にはそのまま援用できる. 具体的なアルゴリズムは 3.で定式化する.

3. アルゴリズム

(1) 定義

以下では, 本研究で提案する modularity 指標に基づく都市圏設定手法のアルゴリズムを具体的に定式化する. 本研究では, 結合法・非重複法に分類されるアルゴリズムを定式化する. また, 地理的な隣接関係を有さない市町村が都市圏を設定することのないアルゴリズムを定式化する. 本研究で取り上げるケース以外のアルゴリズムについては別稿で報告したい.

市町村間の OD 行列をグラフとして捉え, 行列 A で表す. また, 市町村間の物理的な隣接関係を行列 Δ で表す. 行列 A は非対称・有向・重み付きグラフを表し, modularity は次のように定義される (Leicht and Newman, 2008).

$$Q = \sum_l Q_l = \sum_l (p_{ll} - q_l^{in} q_l^{out}) \quad (5)$$

$$p_{ll} = \frac{1}{M} \sum_{i \in V_l, j \in V_l} a_{ij} \quad (6)$$

$$q_l^{in} = \frac{1}{M} \sum_{i \in V_l, j \in V} a_{ji} \quad (7)$$

$$q_l^{out} = \frac{1}{M} \sum_{i \in V, j \in V_l} a_{ij} \quad (8)$$

(2) modularity の性質

クラスタ l とクラスタ m を結合して出来る新たなクラスタの modularity Q_{lm} について次の関係式が成立する.

$$Q_{lm} = Q_l + Q_m + p_{lm} + p_{ml} - q_m^{out} q_l^{in} - q_l^{out} q_m^{in} \quad (9)$$

ただし, $p_{lm} = \frac{1}{M} \sum_{i \in V_l, j \in V_m} a_{ij}$ である. 式(9)より, クラスタ l とクラスタ m を結合した場合にネットワークの modularity が

$$\begin{aligned} \Delta Q_{(lm)} &\equiv Q_{lm} - Q_l - Q_m \\ &= p_{lm} + p_{ml} - q_m^{out} q_l^{in} - q_l^{out} q_m^{in} \end{aligned} \quad (10)$$

増加することがわかる。また、クラスタ l, m を結合した新たなクラスタ n について、 $q_n^{out} = q_l^{out} + q_m^{out}$, $q_n^{in} = q_l^{in} + q_m^{in}$, $p_{nn} = p_{ll} + p_{lm} + p_{ml} + p_{mm}$ という加法性が成り立つ。さらに、クラスタ n とクラスタ l, m 以外の任意のクラスタ o のペアについて、 $p_{no} = p_{lo} + p_{mo}$, $p_{on} = p_{ol} + p_{om}$ という加法性が成り立つ。さらに、地理的な隣接関係についても $\delta_{no} = \max\{\delta_{lo}, \delta_{mo}\}$ という関係が成り立つ。以上の性質を利用することで、計算量を抑えたアルゴリズムの開発が可能になる。

(3) アルゴリズム

本研究で提案する都市圏設定手法のアルゴリズムは次の通りである。

[Step 0]

全ての市町村を個別のクラスタに分割する。任意のクラスタ l について q_l^{out} , q_l^{in} , p_{ll} , Q_l を計算する。また、任意の異なるクラスタのペアについて p_{lm} , Q_{lm} , $\Delta Q_{(lm)}$ を計算する。

[Step 2]

任意のクラスタのペアについて $\delta_{lm} \cdot \Delta Q_{(lm)}$ の最大値が負の場合、アルゴリズムを終了する。

[Step 3]

$\delta_{lm} \cdot \Delta Q_{(lm)}$ が最大となるクラスタ l とクラスタ m を結合する。新たなクラスタを n として、クラスタ分割を更新する。クラスタ n について q_n^{out} , q_n^{in} , p_{nn} , Q_n を計算する。クラスタ n とクラスタ l, m 以外の任意のクラスタ o のペアについて p_{no} , p_{on} , Q_{no} , $\Delta Q_{(no)}$, δ_{no} を計算する。Step2 に戻る。

4. ケース・スタディ

(1) 適用地域

3.で定式化したアルゴリズムを宮城県の市区町村間通勤データに用いる。データは2005年の国勢調査の調査結果であり、当時の市区町村数は48である。各市区町村の従業者数を図-1に、従業者と就業者の比を図-2に示す。平成の大合併の途中段階であり、図-1と図-2からは直ちに分かりにくいが、宮城県内の都市構造に関する一般的な感覚からは、仙

台市を中心とする仙台都市圏、石巻市を中心とする石巻都市圏、古川市を中心とする古川都市圏、気仙沼市を中心とする気仙沼都市圏の4つが存在すると考えられる。金本・徳岡（2002）による都市雇用圏でも宮城県内では以上の4つの都市圏が設定されている。

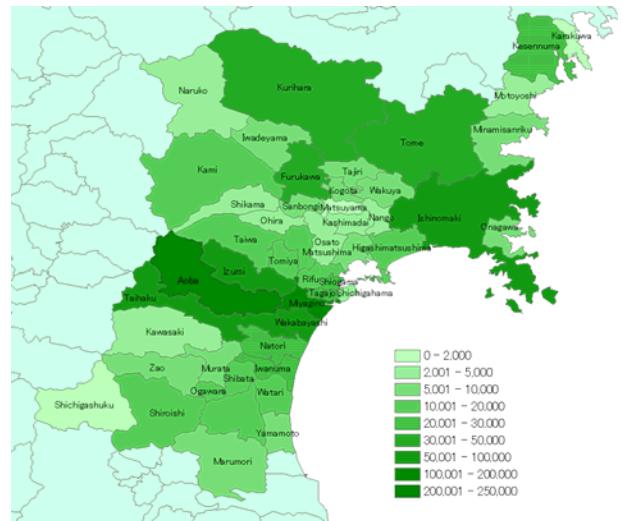


図-1 従業者数

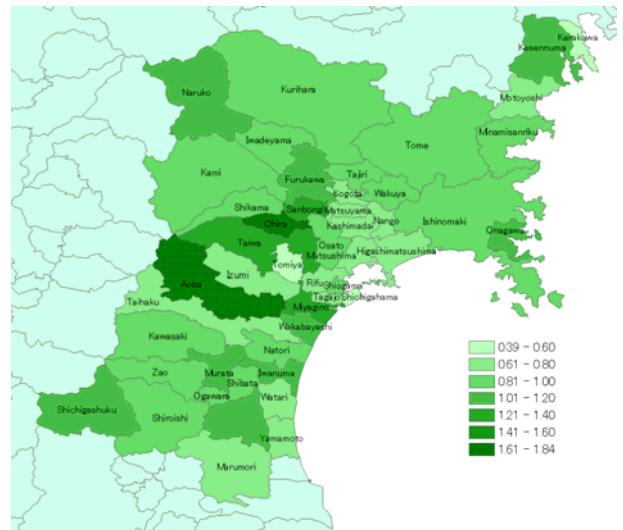


図-2 従業者数と就業者数の比

(2) 適用結果

適用結果を文末の図-3に示す。図-3には繰り返し計算の課程で市区町村がどのように結合されていくかも示している。また、提案手法の特徴を把握するため、Brown and Holmes(1971)の提案手法（以下、BH法）の適用結果を図-4に示す。BH法

では、まず、OD データをマルコフ理論の確率推移行列に変換し、市区町村間の非類似度をマルコフ理論の平均初到達時間として定義する。そして、非類似度行列に階層的クラスタリング手法を適用して市区町村を結合していく。

(3) 考察

図-3 と図-4 より、以下の三点を確認できる。第一に、提案手法では最終的に 4 つの都市圏（仙台都市圏、石巻都市圏、古川都市圏、気仙沼都市圏）が設定されていることが分かる。仙台都市圏が一般的な感覚と比べると大きすぎるように思われるものの、概ね直感と合致した結果であると言える。一方、BH 法では、市区町村を結合するプロセスのどの途中段階でも 4 つの都市圏を見出すことはできない。

第二に、提案手法では繰り返し計算の初期段階で仙台市周辺以外の市区町村が結合していることが分かる。提案手法のこうした性質が第一の帰結とも強く関係している。何故なら、仙台市の中心部のクラスタと郊外の市区町村のクラスタが結合したため、本来、仙台都市圏には入らないと考えられる市区町村が、郊外の市区町村のクラスタへの結合を通じて最終的に仙台都市圏に含まれているからである。ただし、この点は階層的な結合法の使用に起因する性質であり、分割法に分類されるアルゴリズムを利用したり、階層的なクラスタリング手法の分析結果を補正するための各種アルゴリズム（例えば、Kernighan-Lin アルゴリズム）を用いることで改良することは十分に可能である。

第三に、BH 法では初期段階では仙台市周辺の市区町村が結合して徐々に広がっていることがわかる。BH 法において仙台都市圏が徐々に拡大していく理由として、都市圏間の規模（通勤者数の絶対数）が大きく違うことが本質的な原因であると考えられる。裏返すと、提案手法は都市圏の規模の違いに対して比較的頑健な手法であると推測される。

以上の議論より、modularity 指標に基づく提案手法は、都市圏設定の手法として、一定のポテンシャルを有していると考えられる。無論、今回の分析結果は一事例に過ぎないので、今後、ケース・スタ

ディを蓄積して提案手法の有効性を示していく必要がある。

5. おわりに

本研究ではグラフ・クラスタリング手法として提案された modularity 指標に基づく都市圏設定手法を提案した。さらに、ケース・スタディを通じてポテンシャルを示した。ただし、実証分析が圧倒的に不足している。他のデータへの適用や、分析結果を踏まえたアルゴリズムの改良を行っていくことが今後の課題である。

参考文献

- Brown, L. A. and Holmes, J. (1971), The delineation of functional regions, nodal regions, and hierarchies by functional distance approaches, *J. Reg. Sci.*, 11, 57-72.
- Coombes, M. (2000), Defining locality boundaries with synthetic data, *Env. Plan. A*, 32, 1499-1518.
- Farmer, C. J. Q. (2009), Data driven functional regions, *Proc. 10th Int. Conf. GeoComp*.
- Girvan, M. and Newman, M. E. J. (2002), Community structure in social and biological networks, *Proc. Natl. Acad. Sci. U. S. A.*, 99, 7821-7826.
- 金本良嗣・徳岡一幸 (2002), 日本の都市圏設定基準, 応用地域学研究, 7, 1-15.
- Kernighan, B. W. and Lin, S. (1970), An efficient heuristic procedure for partitioning graphs, *Bell Syst. Tech. J.*, 49, 291-307.
- Leicht, E. A. and Newman, M. E. J. (2008), Community structure in directed networks, *Phys. Rev. Lett.*, 1000, 118703.
- Masser, I. and Brown, P. J. B. (1975), Hierarchical aggregation procedures for interaction data, *Env. Plan. A*, 7, 509-523.
- Newman, M. E. J. (2006), Modularity and community structure in networks, *Proc. Natl. Acad. Sci. U. S. A.*, 103, 8577-8582.
- OECD (2002), Redefining Territories: Functional Regions, 042002021P1.
- Schaeffer, S. E. (2007) Graph clustering, *Com. Sci. Rev.*, 1, 27-64.
- Slater, P. B. (1976), A hierarchical regionalization of Japanese prefectures using 1972 interprefectural migration flows, *Reg. Stud.*, 10, 123-132.

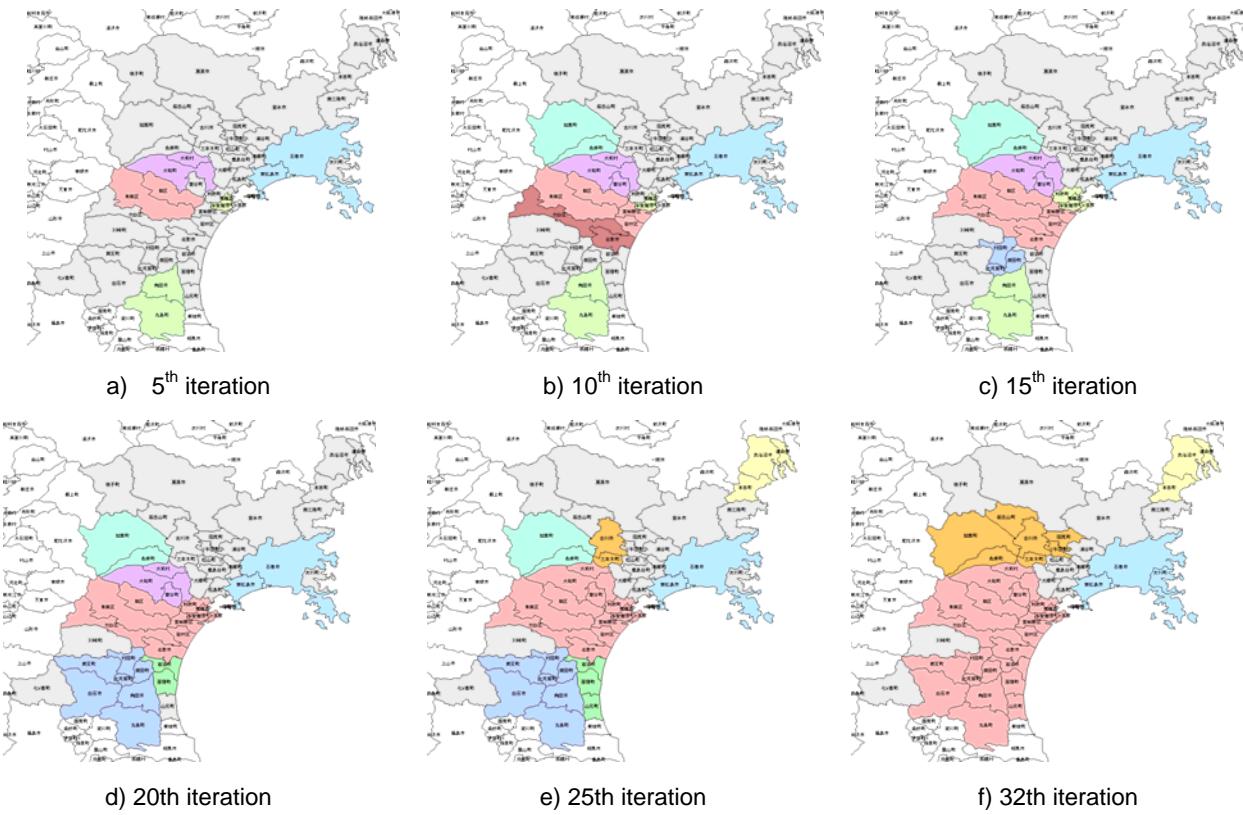


図-3 提案手法の適用結果

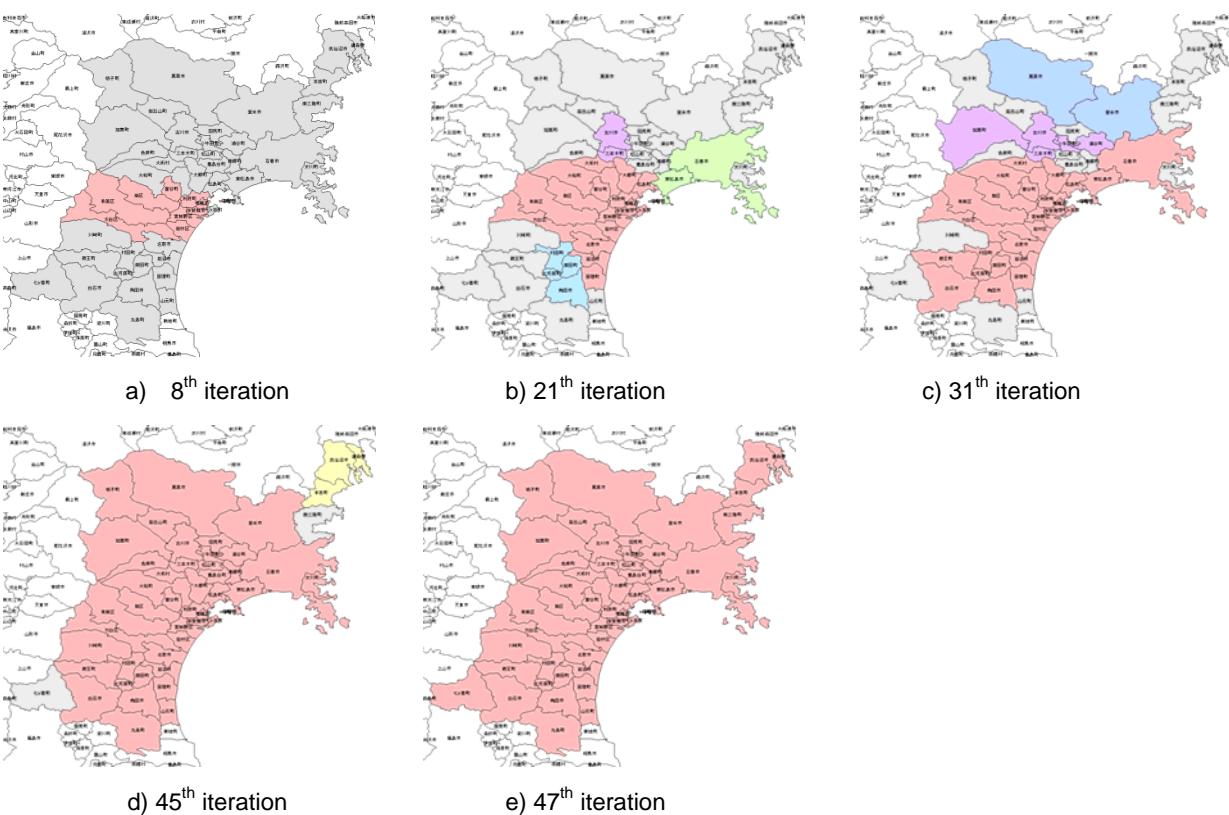


図-4 Brown and Holmes の提案手法の適用結果