

# フローグラフ・アプローチに基づく産業集積検出手法の開発

福本潤也・秋山智信・氏家晃仁

## Detection of Industrial Agglomerations based on Flow Graph Approach

Junya FUKUMOTO, Tomonobu AKIYAMA and Akihito UJIIE

**Abstract:** We develop a new detection method of industrial agglomeration based on a flow graph approach. In the proposed method, at first, a virtual network flow data is constructed by processing the number of and the geographical distance among business establishments. Then, industrial agglomeration areas are detected by applying a community detection method for flow graphs to the data above. The proposed method is simple and versatile. Through the case study, we show its effectiveness.

**Keywords:** フローグラフ (flow graph), 産業集積 (industrial agglomeration), コミュニティ抽出 (community detection)

### 1. はじめに

企業や産業の立地は地理空間上の狭い範囲に局所的に集積している。その要因を探り、分布決定メカニズムを明らかにすることは、空間経済学や地理学の最も重要な研究課題の一つである。

産業集積の性質や要因について分析する場合、上記の原因や類型の区別が困難という問題に直面する (McCann, 2001)。現実の都市には複数の産業が集積しており、仮にある産業について集積が観察されても、それが地域特化の経済によるものか、それとも都市化の経済によるものか、区別は容易でない。そもそも企業が集積する空間範囲を特定したり、集積の経済が働く産業グループを特定することすら容易ではない。産業集積の性質や要因を明らかにするには、個別事例を掘り下げて考察する事例分析、観測データから経験的事実を発見する実証分析、経験的事実の背後にあるメ

カニズムを理論的に解析する理論分析等、多角的にアプローチする必要がある。

産業集積の実態や特性を実証的に解明するため、これまでに様々な研究が行われてきたが、本研究では、地理空間上の事業所分布から“事業所が集積する空間範囲” (以下、集積領域) を検出する問題に着目する。

集積領域を検出する場合、最初に集積領域の定義を行う必要がある。定義方法として、事業所の空間密度に着目する方法と、事業所間のフロー密度に着目する方法の二種類が考えられる。前者では、事業所が相対的に多く立地する空間的にまとまった範囲を集積領域とみなす。空間集積事象の代表的検出手法である空間スキャン統計等を用いることで、空間密度に基づいて事業所の集積領域を検出することができる。一方、後者では、事業所間での資金や情報のフローに着目し、内部のフロー密度が相対的に高い空間領域を集積領域とみなす。ネットワーク内でリンクが高密度に張られたノード集合 (コミュニティ) を検出するコミュニティ抽出法を用いることで、フロー密度に

基づいて集積領域を検出することができる。

フロー密度に着目する手法では、事業所の空間集積の源泉となる事業所間の経済取引データ等を直接利用する。空間密度に着目する手法より集積領域の検出から産業集積の要因解明などの他の分析へと発展させやすい。しかし、事業所間のフローデータの利用可能性は限られており、データ環境が整っていないと利用できない。また、空間的に近接しない事業所間でも経済取引が行われているため、空間的な集積領域が検出される保証はない。

本研究では、フロー密度に着目する手法が抱える上記課題を克服する新たな手法を提案する。具体的には、事業所間の仮想的なフローを仮定してフローグラフとして表現した上で、フローグラフのコミュニティとして集積領域を検出する手法を提案する。フローグラフ・アプローチとは、ノードの次数やノード間最短距離といったネットワークの位相ではなく、ネットワーク上のフローに着目してネットワークの特性の理解を試みるアプローチである (Lambiotte et.al, 2011)。

## 2. 提案手法

本研究では、フローグラフ・アプローチの枠組みを用いた集積領域検出手法を提案する。提案手法の手順は以下の通りである。

### (1) 仮想フローの定義

最初に、事業所数データと事業所間距離データを用いて仮想的なネットワーク・フローを定義する。ポリゴン毎に事業所数が記録されている場合は各ポリゴンを、点事象として事業所が個別に記録されている場合は各事業所を、それぞれネットワークのノードとみなす。ノード間のフローはノード  $i, j$  間の距離  $d_{ij}$  を用いて次式のいずれかで定義する。

$$W_{ij} = 1 / \exp(\beta d_{ij}) \quad (1)$$

$$W_{ij} = 1 / d_{ij}^\beta \quad (2)$$

近接するノード間ほど大きなフローが発生する

と仮定している。

### (2) 遷移確率行列の定義

仮想ネットワーク上のランダムウォーク (RW) を考える。各ノードを個別の状態とみなし、ノード間の遷移を RW として表現する。状態遷移確率行列を次式で定義する。

$$M_{ij} = X_i W_{ij} X_j / \sum_k X_i W_{ik} X_k \quad (3)$$

ここで、 $X_i$  はノード  $i$  の事業所数である。式(3)は事業所数が多いノードやフローが大きいノードに状態が推移しやすい定式化となっている。

### (3) 定常分布の導出

次に、RW の定常分布を導出する。実空間の観測データを元にネットワークを作成すると、状態遷移確率行列  $M$  が既約にならない可能性がある。定常分布  $P$  を一意に求めるため、トランスポーション行列  $T$  を用いて状態遷移確率行列を既約化する。既約化済み遷移確率行列  $M'$  を式(4)で定義する。

$$M' = (1 - \alpha)M + \alpha T \quad (4)$$

$$T_{ij} = 1 / N \quad (5)$$

$N$  はネットワーク・データのノード数、 $\alpha$  はパラメータである。定常分布  $P$  は次式の通り求まる。

$$p_i = \sum_k p_k M'_{ki} \quad (6)$$

### (4) 集積領域検出問題の定義

ネットワーク上の RW に着目してコミュニティを抽出する手法の一つに安定性指標最大化法がある (Delvenne ら, 2010)。安定性指標は、ネットワークのコミュニティ分割を所与とした場合に、あるランダムウォーカーが一定ステップ後に出発ノードと同じコミュニティのノードに滞在する確率として定義される。同最大化法では、ランダムウォーカーの出発ノードの選択確率と指標計測のステップ数を事前に定め、指標値が最大化されるようにネットワークのコミュニティ分割を求める。

具体的には、安定性指標  $r(t)$  は次式で表される。

$$r(t) = \text{trace } H^T \{P(M')^t - p^T p\} H \quad (7)$$

ここで、 $H$  はノードとコミュニティの所属関係

を表す行列 ( $H_{ij}$  はノード  $i$  がコミュニティ  $j$  に所属する場合に 1, 所属しない場合に 0,  $\sum_j H_{ij} = 1$ ),  $P$  は  $P$  を対角成分とする対角行列である. 安定性指標  $r(t)$  は, 定常確率  $P$  に従って出発ノードを選択するランダムウォーカーが  $t$  ステップ  $t$  後に出発ノードと同じコミュニティ内のノードに滞在する確率と, 無限ステップ後に同じコミュニティ内のノードに滞在する確率の差を表している.

本研究では, 集積領域検出問題を,  $r(t)$  を目的関数,  $H$  を決定変数とする安定性指標最大化問題として定義する. 同問題は, モジュラリティ最大化法で利用される様々なヒューリスティクスを用いて近似的に解くことができる.

### (5) 事後処理

安定性指標最大化問題は, 各ノードが必ずいずれかの集積領域に所属する定式化となっている. そのため, あるノードが単独で集積領域を形成するなど, 分析目的と照らし合わせて意味のない領域までが検出される. そこで,  $H_j \cdot p < \varepsilon$  を満たす集積領域を間引くこととする (ただし,  $H_j$  は行列  $H$  の第  $j$  列ベクトル,  $\varepsilon$  は十分小さい正数).

### (6) 留意点

隣接行列  $W$  を式(1)および式(2)で定義する場合, バイアス付き  $RW$  の推移行列の  $(i, j)$  成分はそれぞれ次式で表される.

$$T_{ij}^{(\alpha)} = \frac{(x_j / d_j^\beta)}{\sum_k (x_k / d_k^\beta)} \quad (8)$$

$$T_{ij}^{(\alpha)} = \frac{x_j \exp(-\beta d_{ij}^2)}{\sum_k x_k \exp(-\beta d_{ik}^2)} \quad (9)$$

式(8)は消費者の商店選択行動等で広く用いられるハフモデルの選択確率と, 式(9)は発生制約型エントロピーモデルの選択確率と等しい. 提案手法では, 顧客・資金・情報等が事業所間を距離抵抗の影響下で循環していると考え, それらが一定期間経過後も内部に滞在する事業所の集合を集積領域として検出していると解釈できる. 式(1)と式(2)に含まれるパラメータ  $\beta$  は分析者が設定する必要があるが, 行動モデルの推定結果が利用可能

な場合にはそれらを利用できる.

### (7) 提案手法の意義

提案手法の位置として, 以下の二点を指摘できる. 第一に, 距離抵抗を考慮した仮想フローを定義することで, 事業所密度に基づく手法のように幾何形状制約を仮定せずに空間的な集積領域を検出することを可能としている. また, 行動論的解釈が容易で, 集積領域の検出結果を集積要因の解明等への発展が可能な定式化となっている. 第二に, 安定性指標のステップ数  $t$  を変化させることで感度分析が可能である. 一般に, 産業集積はマイクロ・メゾ・マクロと様々なスケールにおいて存在する. 特定のスケールの集積領域の検出に留まらず, 様々なスケールで集積領域を検出し, スケールが異なる分析結果の関係性を明らかにすることで, 集積現象についての理解が深まると期待される. 上記感度分析は, スケールが異なる集積領域を検出したり, それらの階層関係を明らかにする上で有益である.

## 3. ケーススタディ

### (1) 使用データ

平成 21 年経済センサス・基礎調査地域メッシュ統計の 1/2 地域メッシュデータを用いた. 以下では, 宮城県石巻市周辺の全産業事業所数データを用いた分析結果のみ紹介する. 図-1 より石巻市中心部に事業所の大きな集積が見られることが分かる.

### (2) 分析結果と考察

ステップ数  $t$  を 1 から 10,000 まで変化させながら集積領域を検出した結果を図-2 に示す. 図-2a) の横軸はステップ数  $t$  (常用対数), 縦軸は安定性指標  $r(t)$  である. 図よりステップ数の増加に伴い, 検出される集積領域の個数が減少していることが分かる. 図-2 の b) から e) の比較より, ステップ数が 1 から 10 の範囲では 4 個の集積領域に分割されていた東松島市矢本地区から石巻市渡波地区までの海沿いの地域が, ステップ数の増加に伴

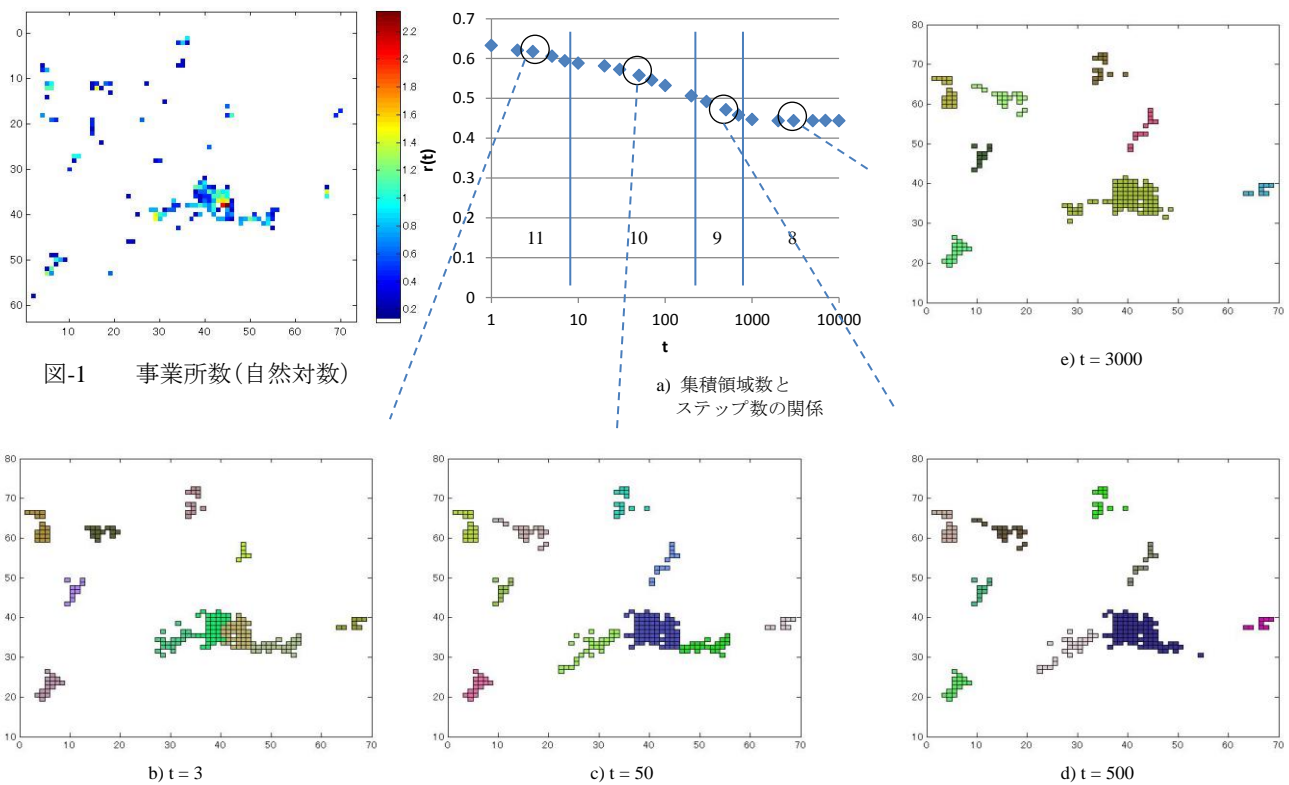


図-2 分析結果

い最終的に1つの集積領域に集約されていることが分かる。また、図-2のa)より、集積領域の個数が9個となるステップ数の範囲は他と比べると狭く不安定であることが示唆される。これより、東松島市矢本地区から石巻市渡波地区までの海沿いの地域は、渡波地区と石巻中心部、矢本地区に大きく分かれ、更に石巻中心部は石巻駅周辺と蛇田地区に細かく分かれることを確認できる。一方、図-2のb)からe)より、上記地域以外はステップ数の変化によらず検出される集積領域は変化しないことが分かる。

### 謝辞

本研究は東京大学空間情報科学研究センターの空間データ利用を伴う共同研究 (No. 582) の成果であります。平成 21 年経済センサス・基礎調査地域メッシュ統計のデータを利用しました。また、本研究は JSPS 科研費 26289196 の助成を受けたものです。ここに記して謝意を表します。

### 参考文献

- McCann, P., 2001. Urban and regional economics. New York: Oxford University Press.
- Lambiotte, R., Sinatra, R., Delvenne, J.-C., Evans, T. S., Barahona, M. and Latora, V., 2011. Flow graphs: Interweaving dynamics and structure. *PHYSICAL REVIEW E*, **84**, 017102.
- Delvenne, J.-C., Yaliraki, S. N. and Barahona, M., 2010. Stability of graph communities across time scales, *PNAS*, **107**, 29, 12755-12760.
- Delvenne, J.-C., Schaub, M. T., Yaliraki, S. N. and Barahona, M., 2013. The stability of a graph partition: A dynamics-based framework for community detection. In Mukherjee, A. et al., eds. *Dynamics on and of complex networks*, **2**. New York: Springer Science+Business Media.