

ホップフィールド型ニューラルネットによる 制約条件付きクラスタリング

館 俊太[†] 武藤佳恭^{††}

教師なしクラスタリングにおいて、各クラスタの標本数やクラスタ中心までの最大距離などといった条件や事前知識を反映させる手法は現実問題への応用が大きい。この分野ではファジィクラスタリングモデルや制約条件付き K-平均法などが広く研究されているが、本稿ではニューラルネットを用いた組合せ最適化問題として定式化する方法を述べる。提案モデルは代表点と標本点の組合せに対応したマトリックス状のニューロンからなり、入力値に制約条件を反映する制約項を設ける。これを最小化するようにニューロンを動作させることにより、ニューラルネットでの制約条件付きクラスタリングを実現できることを示した。また実験により他のニューラルネット手法に比べても振動解や局所解への収束が少ないことを示した。

Conditional Clustering Model by Hopfield Neural Network

SHUNTA TATE[†] and YOSHIYASU TAKEFUJI^{††}

In unsupervised clustering, the technique of using prior knowledge and setting conditions including restriction of the number of samples in each cluster and restriction of the maximum distance from each sample to a cluster center, has many applications in solving real-world problems. We have used a Hopfield neural network in order to formulate a clustering problem as optimized combination problem. The neural network consists of a neuron matrix corresponding to the combination of the clusters and each data sample. Each neuron's motion equation includes constraint terms for clustering. The experimental result showed the effectiveness of the proposed model with less degree of convergence to an oscillatory solution or local minimum.

1. はじめに

教師なしのクラスタリング手法は階層的な方法と非階層的な方法に大別され、本稿では後者を主に扱う。後者には K-平均法^{1),2)}、ファジィ c-平均法 (FCM)^{3)~5)}、自己組織化マップ⁶⁾などがある。またクラスタリング手法の多くは分類クラス数をあらかじめ決めておく必要があるが、分類クラス数を事前に与える必要のない手法として競合学習法⁷⁾も提案されている。データの分類基準に制約条件を設ける手法としてはファジィクラスタリングにおいては可能性クラスタリング^{3),4)}があり、この分野では広く数理的性質の研究や応用研究がなされている。K-平均法には制約条件付き K-平均法などが提案されているが^{1),2),8)}、初期値やデータセッ

トに最終結果が左右されやすく、局所解に収束しやすい問題がある。また制約付きの目的関数の最適化にラグランジュ未定乗数法を用いて解く方法⁹⁾などもあるが、複数のクラスタでは解空間が非線形非凸関数になり、局所解としてしか求めることができないなどの問題がある。本稿ではホップフィールド型ニューラルネット¹⁰⁾を用いて、制約条件付きクラスタリングを解く数理モデルを提案する。ホップフィールド型ニューラルネットを用いて組合せ最適化問題を解く研究は古くから多数あり^{10)~12)}、広い応用性を持っている。クラスタリング問題も、分類される標本データとクラスタとの最適組合せを求める最適化問題として定式化すると、ニューラルネットによって解くことができる。これの先行手法としては MNM (Maximum Neuron Model) を用いたクラスタリングと Kohonen の自己組織化的手法などがあげられる^{6),13)}。これらは数理的には K-平均法にも類似性がある。ニューラルネットを用いた最適化手法は簡便で低計算量の利点があるが、やはり局所解への収束の頻度や初期値の依存性が問題

[†] 慶應義塾大学大学院政策・メディア研究科
Graduate School of Media and Governance, Keio University

^{††} 慶應義塾大学環境情報学部
Faculty of Environmental Information, Keio University

になる。この問題に対して Oka らは 2 つのニューラルネット手法を組み合わせることなどによって収束を早め、かつクラスタリングの大域解への収束率を向上させることが可能であるとしている¹³⁾。このような工夫があるものの、確率モデルなどの分類手法がデータの分布の統計的性質などを基準にするために、結果に統計的妥当性を与えることができるのに対して、ニューラルネット手法は局所的に距離を最小化して解を発見的に求めるため、結果が初期値や局所的なデータの偏りに強く依存する問題があった。本稿ではこのような問題を受けて、ニューラルネット手法でもデータの分布やクラスタの規模などの先験知識を基準として利用できる数理モデルを提案する。また現実世界の制約条件——たとえば、資源配置の問題で配置する資源の上限や下限を設定したいなど——を定式化するための方法と実際の応用事例についても述べる。

2. アルゴリズム

今、 p 次元実空間 R^p 内に n 個の特徴点 $X = \{x_1, \dots, x_n\}$ が分布しているとき、これを c 個のクラスタ $W = \{w_1, \dots, w_c\}$ に分類する問題を考える。提案アルゴリズムとの比較のために、ニューラルネットワークを用いた 2 つの手法についてまず述べる。

2.1 MNM 法

Maximum Neuron Model (MNM)^{11),13)}を用いたクラスタリング手法について述べる。評価関数(エネルギーコストの関数)としてクラスタの代表点から各特徴点への距離の総和

$$E = \sum_i \sum_j d_{ij} v_{ij}, \quad v_{ij} \in \{0, 1\} \quad (1)$$

を定める。ただし d_{ij} は代表点 w_i から x_j へのユークリッドノルム

$$d_{ij} = \sum_p \|w_i^p - x_j^p\|^2 \quad (2)$$

である。また d を要素とする行列を D とする。 V はどの点がどのクラスタに属するかを表現する 2 値を要素とする行列である。するとクラスタリングは上式(1)を最小化する V を求める問題となる。このとき、MNM 法では v_{ij} をニューロン ij の出力状態と見なし、静止と発火の状態を 0 と 1 の 2 値に対応させる。ここで i と j はそれぞれクラスタ i に属する特徴点 j に対応する。ニューロンの出力 V を決定する膜電位として実数値 u_{ij} を要素とする U を考える。そして式

$$v_{kj}(t+1) = \begin{cases} 1 & (u_{kj}(t) = \max[u_{ij}(t); \forall i]) \\ 0 & (\text{otherwise}) \end{cases} \quad (3)$$

によって V を発火させるとする。 t は計算のステップ数である。このニューロンの電位の値 U はファジィクラスタリングにおけるメンバシップ値との類似性がある。 V を仮に連続値と見なしたとき式(1)を v_{ij} で偏微分した値は

$$\frac{\partial E}{\partial v_{ij}} = d_{ij} \quad (4)$$

となるが、これに負の係数 $-\delta$ をかけた値を用いて U の t についての微分方程式

$$\frac{du_{ij}}{dt} = -\delta d_{ij} \quad (0 < \delta < 1) \quad (5)$$

を定めることにより評価関数 E を局所最小化することが期待できる。式(5)は一次のオイラー法を用いて

$$u_{ij}(t+1) = u_{ij}(t) - \delta d_{ij} \quad (6)$$

としておく。なお、クラスタ W は以下の代表点を重心として求める式、

$$w_i(t+1) = \frac{\sum_l x_l v_{il}(t)}{\sum_l v_{il}(t)} \quad (7)$$

によって決定する。アルゴリズムは下のステップよりなる。

- ステップ 1 U を一様乱数で初期化する。
- ステップ 2 V を式(3)によって計算する。
- ステップ 3 W を式(7)によって計算する。
- ステップ 5 D を式(2)によって計算する。
- ステップ 4 U を式(6)によって更新する。
- ステップ 6 式(1)の値が収束するまでステップ 2 に戻って反復する。

2.2 Kohonen の自己組織化法

Kohonen の方法のアルゴリズムが上述の MNM 法と異なる点は W の更新時に、任意の x_j とそれが属する代表点 w_i の 1 組を選び、係数 α を掛けて W を徐々に自己組織的に更新する点である。

$$w_i(t+1) = w_i(t) + \alpha(x_j - w_i) \quad (0 < \alpha < 1) \quad (8)$$

また MNM 法と異なってここでは内部状態 U を設けずに V を D より直接求める。

$$v_{kj}(t+1) = \begin{cases} 1 & (d_{kj}(t) = \min[d_{ij}(t); \forall i]) \\ 0 & (\text{otherwise}) \end{cases} \quad (9)$$

アルゴリズムは下のステップよりなる。

ステップ 1 W を一様乱数で初期化する .

ステップ 2 D を式 (2) によって計算する .

ステップ 3 V を式 (9) より更新する .

ステップ 4 任意の $x_j(t)$ を選び, $w_i(t)$ を式 (8) に従って更新する .

ステップ 5 式 (1) の値が収束するまでステップ 2 に戻って反復する . また, α は 0.8 程度から始めて反復とともに 0 に近づくようにする .

なお, この Kohonen の方法は自己組織化の部分が高計算コストを要する . ところでこれをせすにいっせいに $w_i(t)$ を $w_i(t+1)$ に変更するようになると簡便な方法である K-平均法と等しくなる .

2.3 提案アルゴリズム

まず, 制約条件付き評価関数として以下のような式を考える .

$$E = \sum_m \lambda_m |\text{constr.eq.}(m)| \left(\sum_m \lambda_m = 1 \right) . \quad (10)$$

評価関数 $\text{constr.eq.}(m)$ には以下のようなものが考えられる .

$$\left| \sum_i \sum_j d_{ij} v_{ij} \right| , \quad (11)$$

$$\left| \frac{1}{2} \sum_i \left(ns_i - \sum_j v_{ij} \right)^2 \right| , \quad (12)$$

$$\left| \frac{1}{2} \sum_i \sum_j (d_{avg} - d_{ij})^2 v_{ij} \right| , \quad (13)$$

$$\left| \sum_i \sum_j \hat{d}_{ij} \right| , \quad \hat{d}_{ij} = \begin{cases} d_{ij} - d_{max} & (v_{ij} = 1 \\ & \text{and } d_{ij} > d_{max}) \\ 0 & (\text{otherwise}) \end{cases} . \quad (14)$$

式は上からそれぞれ, 総距離の最小化, クラスタのサイズ (すなわち, クラスタに含まれる点の数) の条件, 代表点への距離の分散の最小化, 代表点までの最大距離の制約を表現する . ただし s_i は $\sum_i s_i = 1$ となるクラスタのサイズを制約する係数, d_{avg} は代表点までの距離の全平均, d_{max} は事前に与えられる代表点までの距離の上限値である . なお, λ はそれぞれの

評価関数の項を重視するかの係数である . たとえば距離の最小化項とクラスタサイズの制約は解空間上で部分的に競合するが, λ を調整することで制約の効果の割合を決めることができる .

ここで, 上式 (10) からなる評価関数 E を v_{ij} で偏微分して MNM 法と同様に $u_{ij}(t)$ の更新式を求める . 本稿においては制約条件として式 (11) および式 (12) のみを扱う . すると,

$$\frac{\partial E}{\partial v_{ij}} = \lambda_1 d_{ij} + \lambda_2 \left(\sum_j v_{ij} - ns_i \right) , \quad (15)$$

ここから u_{ij} の式は

$$\frac{du_{ij}}{dt} = -\alpha_1 \lambda_1 d_{ij} - \alpha_2 \lambda_2 \left(\sum_j v_{ij} - ns_i \right) \quad (16)$$

となる . α はニューロンの内部電位と出力の間の特性を示す係数である . 評価関数のうち, 距離の最小化項は最適解においても 0 にはならないので, U が発散しないように減衰項 $-\eta u_{ij}(t)$ を加える . 正の係数 η は減衰の時定数である . $u_{ij}(t)$ の更新式は最終的に

$$u_{ij}(t+1) = u_{ij}(t) - \tau(\alpha_1 \lambda_1 d_{ij} + \alpha_2 \lambda_2 \left(\sum_j v_{ij}(t) - ns_i \right) + \alpha_3 \eta u_{ij}(t)) \quad (17)$$

となる . ただし τ は更新の変化分を決める時定数である . また, $W(t)$ の更新式は下式のように係数 β を用いてある程度連続的に $W(t)$ の値が変化するようにして振動解に陥るのを避ける .

$$w_i = \frac{\sum_l x_l v_{il}(t)}{\sum_l v_{il}(t)} , \quad (18)$$

$$w_i(t+1) = w_i(t) + \beta(w_i - w_i(t)) \quad (0 < \beta < 1) . \quad (19)$$

下にアルゴリズムを示す .

ステップ 1 U と W を一様乱数で初期化する .

ステップ 2 V を式 (3) によって計算する .

ステップ 3 W を式 (18), (19) によって計算する .

ステップ 4 D を式 (2) によって計算する .

ステップ 5 U を式 (17) によって更新する .

ステップ 6 式 (10) の値が収束するまでステップ 2 に戻って反復する .

ここで, パラメータセット λ を調整することによって評価関数がクラスタリング結果に与える影響を調節することができる . なお, クラスタサイズの制約が満たされたときに式 (17) 右辺の α_2 の項は 0 になるが, クラスタ数と特徴点数が同一でない限り距離の最小化

項である α_1 の項は 0 になることはないので、特徴点の分布次第によっては収束値の近くで振動状態を起こす原因になる。また、初期の繰返し計算の段階で α_2 の項が効きすぎると距離の最小化が進まず、クラスタサイズの制約は満たしていても、総距離について大域的に最適化されていない解に収束しやすい。このような局所解や振動解に陥ることを少なくさせるためには、反復計算にあわせて α を変化させる必要がある。1つの方法としては解が収束するまで一方の制約項のみを有効にして反復計算し、解が収束したら他方の制約項のもとで反復計算をする方法である。もう1つの方法としては以下のようにパラメータ α をアニーリング

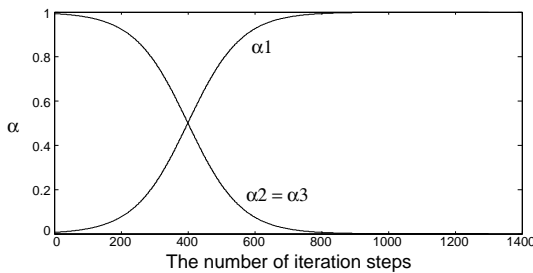


図1 アニーリングに用いるシグモイド関数

Fig. 1 Sigmoid function which was employed for annealing.

スケジュールによって変化させる方法である。本稿ではこちらを用いた。 α のチューニングには図1のシグモイド関数

$$\alpha_1 = \frac{1}{1 + e^{-\rho(t-t_0)}}, \quad (20)$$

$$\alpha_2 = \alpha_3 = 1 - \alpha_1 \quad (21)$$

を用いた ($\rho = 0.0125$, $t_0 = 400$)。これによって十分な反復計算の回数をとれば動作が振動せずに収束することが確かめられた。

3. テストデータ

図2のように、ガウス関数の混合分布よりなるテストデータを用意した。それぞれクラスタの数は3, 5, 7の3種類あり、1つのクラスタがそれぞれ16個, 64個, 256個の点よりなる計9種類のテストデータである。これを用いて3つの手法の比較テストを行った。

4. 実験結果

パラメータはそれぞれ $\lambda_1 = \lambda_2 = 1$, $\beta = 0.5$, $\eta = 0.1$, $\tau = 0.2$ としてシミュレーションを行った。提案手法に関しては各クラスタに含まれる特徴点の数がすべて等しいとする制約条件を与えた。3つの手法で各9データにつき200試行の実験を行ったが、提案

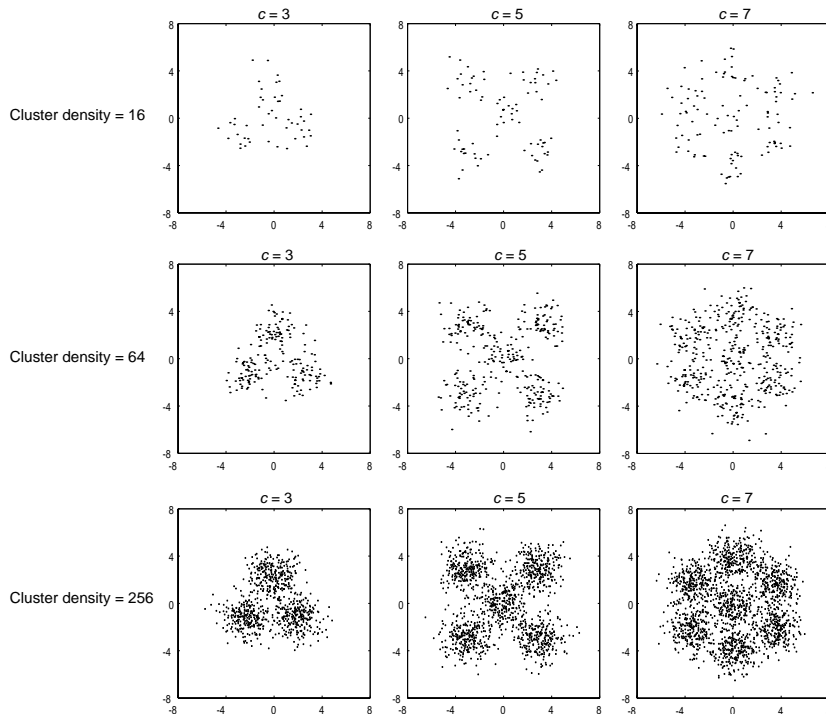


図2 テストデータのセット

Fig. 2 Data set for experiment.

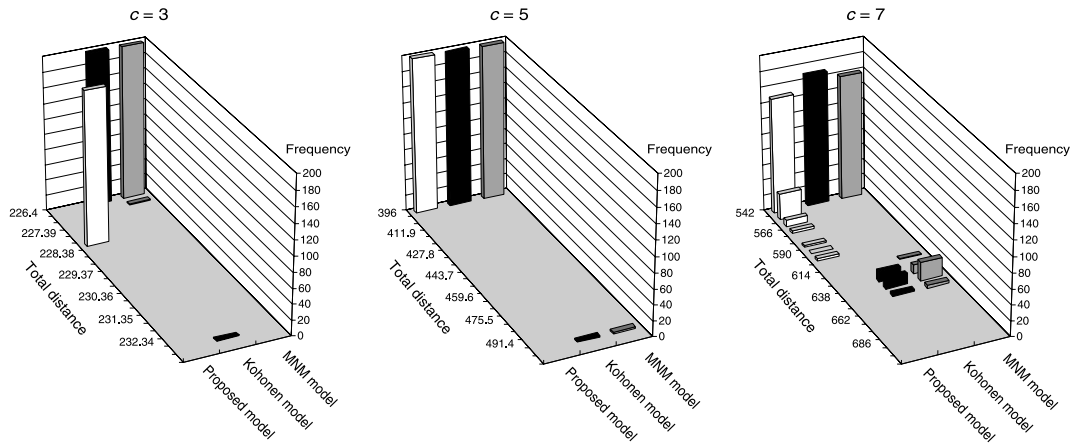


図3 クラスタリング結果における総距離の度数分布

Fig. 3 Frequency distribution of the total distances of the clustering results.

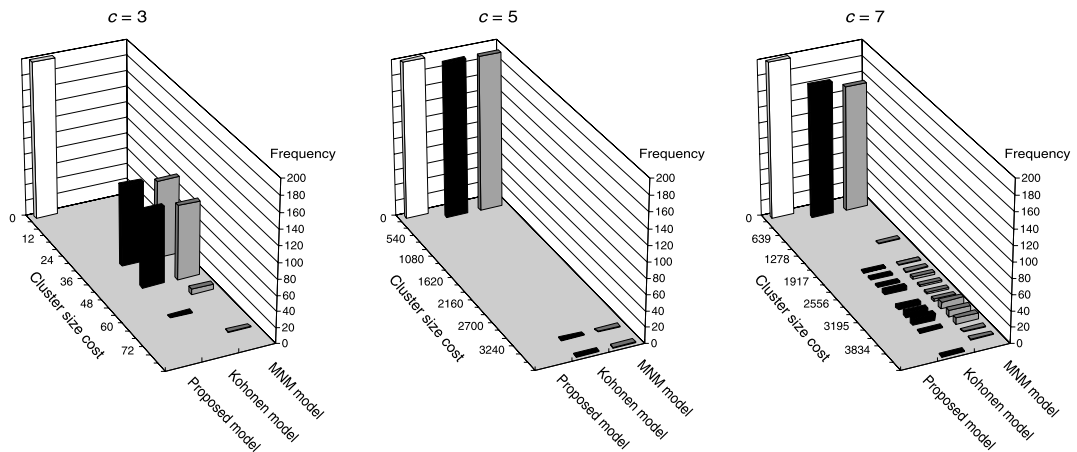


図4 クラスタリング結果におけるクラスタサイズの制約のコストの度数分布

Fig. 4 Frequency distribution of the deviation of the cluster-size condition.

手法は9つのすべてのデータとすべての試行において制約条件を満たす結果を得られた。特徴点から代表点までの距離の総和の度数分布と、制約条件の評価値である各クラスタの点の数と実際の数との乖離の度数分布をそれぞれ図3, 図4に示す。図3, 図4は左よりそれぞれクラスタの数が3, 5, 7のデータでの結果である。クラスタ1つあたりの点の数は3つのデータとも64個である。提案手法が初期値に依存せず、度数内の最頻値への収束率が良いことがわかる。また、提案手法はクラスタサイズの制約条件を満たすために、制約のない他の2手法よりも距離の総和が少し大きくなることも分かる。

また距離の総和の200試行における平均値と標準偏差を表1にまとめた。数値は特徴点1つあたりで割った値である。また参考のためにテストデータのガウス

分布の中心に各代表点を固定して算出した数値もあわせて示す。表より、9つのデータのうち7つにおいて提案手法の解の収束率が他の2手法より優れていることが分かる。クラスタの数が7かつサイズが64と256の場合のみ標準偏差が大きくなっているが、これは特徴点が増えて特徴点の密度が上がったため、代表点の移動量が小さくなり、距離の最小化項によって各代表点の位置が十分最適化される前にサイズの制約条件の緩和計算が行われてしまったためと思われる。また、9つのデータのうち6つの場合においてガウス分布の中心点からの総距離よりも提案手法の値の方が小さい。これは与えられたデータについてはガウス分布の中心を多少外れた位置に総距離を最小化する代表点があることを示している。

表 1 200 試行での距離の総和の平均値 (括弧内は標準偏差)

Table 1 The average cost of the total distances (and the standard deviation in the bracket) by 200 simulation runs.

Cluster density	Method	c = 3		c = 5		c = 7	
16	Proposed model	1.2641	(0.000)	1.0755	(0.269)	1.2483	(13.73)
	Kohonen model	1.2658	(3.845)	1.0816	(9.348)	1.2432	(13.94)
	MNM model	1.2679	(3.723)	1.0780	(8.923)	1.2614	(14.36)
	Fixation on centers	1.3357	—	1.0974	—	1.1509	—
64	Proposed model	1.1883	(0.016)	1.2388	(0.000)	1.2691	(43.98)
	Kohonen model	1.1798	(0.457)	1.2409	(10.29)	1.2462	(34.03)
	MNM model	1.1797	(0.093)	1.2409	(10.30)	1.2574	(38.64)
	Fixation on centers	1.1983	—	1.2456	—	1.2187	—
256	Proposed model	1.2394	(0.007)	1.2753	(0.073)	1.2446	(153.17)
	Kohonen model	1.2388	(0.000)	1.2739	(0.109)	1.2279	(121.74)
	MNM model	1.2389	(0.043)	1.2739	(0.106)	1.2192	(98.20)
	Fixation on centers	1.2405	—	1.2758	—	1.2060	—

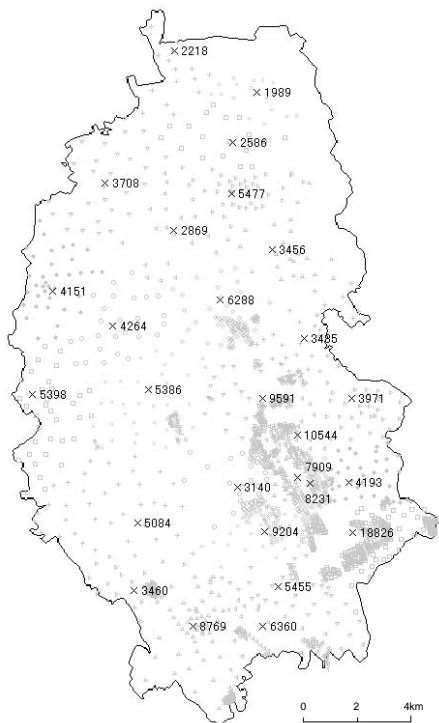


図 5 つくば市の郵便局の配置図

Fig. 5 Locations of the post offices in Tsukuba city.

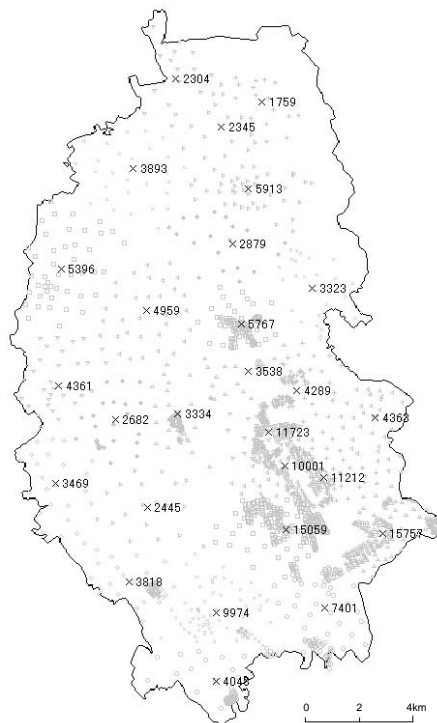


図 6 Kohonen の方法による結果の例

Fig. 6 Experimental result by Kohonen model.

5. 制約条件付きクラスタリングの資源配置問題への応用

実問題への応用として、つくば市の人口データに基づく郵便局の配置を行った。ここでは実際の道路の条件は考慮せずにユークリッド距離のみについてクラスタリングを行った。図 5 に人口分布と現実の 27 の郵便局の位置を示す。データは平成 7 年度国勢調査区別の人口データである。このデータでは人口 156,012

人が 2,144 個の座標点として平面直角座標系 (9 系) 上に示されている。なお図中の数字はそれぞれの郵便局 (クラスタ) に属する人口の数を示す。図 6, 図 7 にそれぞれ Kohonen の方法, MNM 法によるクラスタリング結果の例をあげる。表 2 より 2 手法の結果が安定していないことが分かる。特に図 6 の結果例では上方に小規模な郵便局が多数生じてしまっている。このように、ある程度以下の規模の施設を多数設置することは固定費用が高つくことが考えられて望まし

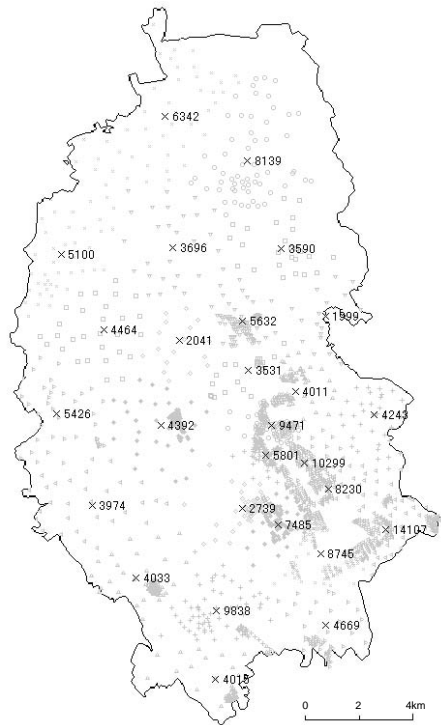


図7 MNM法による結果の例

Fig. 7 Experimental result by MNM model.

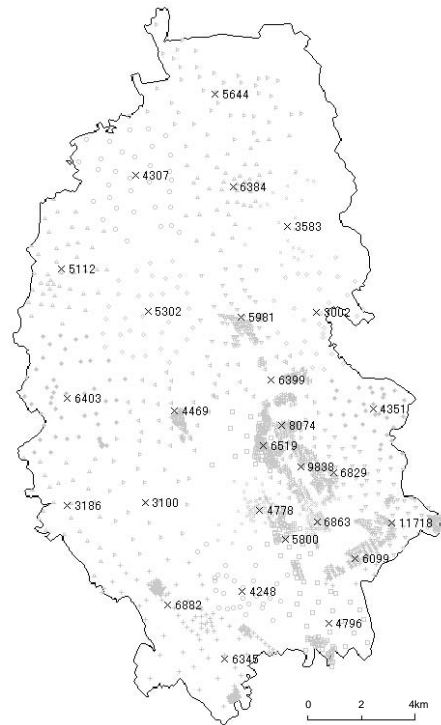


図8 提案手法による結果の例

Fig. 8 Experimental result by the proposed model.

表2 1人あたり平均距離の50試行の平均値および標準偏差
Table 2 The average of the average distances and the standard deviation by 50 simulation runs.

Method	Average distance (m)	(Standard deviation)
Proposed model	952.48	(18.501)
Kohonen model	965.64	(22.143)
MNM model	957.58	(20.967)
Actual locations	1148.75	—

くない。そこで次の実験として提案手法を用いて制約条件付きのクラスタリングを行う。ここでは制約として郵便局の規模の最小値を設定する。具体的には、あるクラスタに属する人口が決められた下限以下のときには、それ以外のクラスタに属するニューロンの評価関数にコストが課せられて発火が抑えられるような制約項を設けて配置を行った。結果の例が図8である。ここでは制約人数として3,000人とした。図9はこの実験で各クラスタごとの人口を多い順から並べたものの50試行平均である。提案手法では制約条件(3,000人以上)が満たされていることが分かる。

6. 考察とまとめ

実験によって、与えられた制約条件を満たし、かつ、MNM法やKohonenの方法と比較しても十分総コス

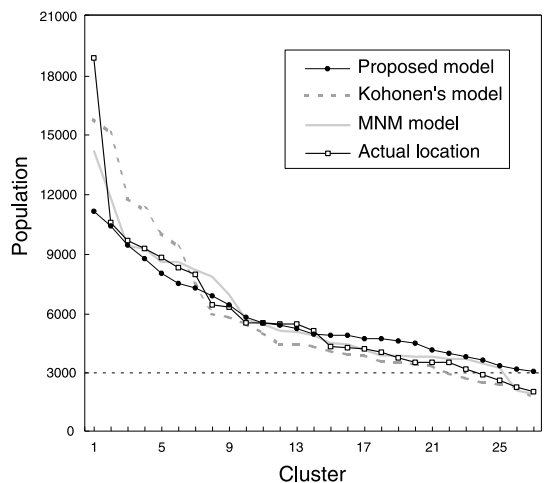


図9 クラスタあたりの人口の数

Fig. 9 Population of each cluster.

トが小さな解が得られることが分かった。提案手法には現実の問題での様々な制約条件を応用させることができる。今後は確率分布推定との関連での評価や、またクラスタ数と特徴点が増えたときに最適解への収束率が落ちることから、最適化が十分に進むための動作式の性質の良いパラメータとその境界条件について評価する必要があると考える。

謝辞 地理情報データの使用に関してご助言およびご協力をいただいた大日本印刷株式会社 C&I 事業部の米村征洋氏に深く感謝いたします。

参 考 文 献

- 1) Tung, A.K.H., Han, J., Lakshmanan, L.V.S. and Ng, R.T.: Constraint-Based Clustering in Large Databases, *Proc. 2001 Int. Conf. on Database Theory (ICDT'01)*, Vanden Bussche, J. and Vianu, V. (Eds.), London, U.K. (2001).
- 2) Lefkovich, L.P.: Conditional Clustering, *Biometrics*, Vol.36, pp.43-58 (1980).
- 3) Krishnapuram, R. and Keller, J.M.: A Possibilistic Approach to Clustering, *IEEE Trans. Fuzzy Syst.*, Vol.1, No.2, pp.98-110 (1993).
- 4) 宮本定明, 高田 治, 渋谷和宏, 水津大輔: 正則化ファジィc-平均におけるいくつかの目的関数と2つの制約, *FAN2001*, 大阪, pp.313-316 (2001).
- 5) 宮本定明: クラスタ分析入門: ファジィクラスタリングの理論と応用, 森北出版, 東京 (1999).
- 6) Kohonen, T.: *Self-Organizing Maps*, Springer, Berlin (1997).
- 7) Liu, Z.-Q. and Zhang, Y.: Compensation Competitive Learning, *Int. J. of Computational Intelligence and Applications*, Vol.1, No.3, pp.303-322 (2001).
- 8) MacQueen, J.: Some Methods for Classification and Analysis of Multivariate Observations, *Proc. 5th Berkeley Symposium on Math. Stat. and Prob.*, Vol.1, pp.281-297 (1967).
- 9) 岡部篤行, 鈴木敦夫: 最適配置の数理, 朝倉書店, 東京 (1992).
- 10) Hopfield, J.J. and Tank, D.W.: Neural Computation of Decisions in Optimization Problems, *Biological Cybernetics*, Vol.52, pp.141-152 (1985).
- 11) Takefuji, Y.: *Neural Network Parallel Computing*, Kluwer Academic Publishers (1992).
- 12) Takefuji, Y. and Oka, S.: Neural Computing for Solving Intractable Problems, *The Journal of the Institute of Electronics, Information, and Communication Engineers*, Vol.79, No.9, pp.943-946 (1996).
- 13) Oka, S., Ogawa, T., Oda, T. and Takefuji, Y.: A New Self-Organization Classification Algorithm for Remote-Sensing Images, *IEICE Trans. Information and Systems*, Vol.E81-D, No.1, pp.132-136 (1998).

(平成 14 年 11 月 1 日受付)

(平成 14 年 12 月 20 日再受付)

(平成 15 年 1 月 3 日採録)



館 俊太 (学生会員)

昭和 49 年生。平成 9 年慶應義塾大学環境情報学部卒業。平成 11 年同大学大学院政策・メディア研究科修士課程修了。現在、同大学院博士後期課程在学中。画像処理、ニューラルコンピューティング、視覚の脳機能モデルの研究に従事。



武藤 佳恭

昭和 30 年生。昭和 58 年慶應義塾大学大学院博士課程電気工学専攻修了。工学博士。現在、ケースウェスタンリザーブ大学電気工学準教授、慶應義塾大学環境情報学部教授。ニューラルコンピューティング、ハイパースペクトラルコンピューティング等の研究に従事。情報処理学会 20 周年記念論文賞 (1980), IEEE Trans. on Neural Networks 功労賞 (1992) 受賞。