

# Complementary Control Particle Swarm Optimization の提案と 遺伝子ネットワーク推定問題への応用

村田 裕章<sup>†</sup> 米山 尊通<sup>‡</sup> 三田村 公智<sup>†</sup> 越野 亮<sup>‡</sup> 木村 春彦<sup>†</sup>  
金沢大学<sup>†</sup> 石川工業高等専門学校<sup>‡</sup>

## 1. はじめに

Particle Swarm Optimization (PSO : 粒子群最適化) は鳥の群れや魚の群泳など、群れを成して移動する生物の行動パターンから着想を得た最適化手法である[1]. 自然界において、群れ内の個体は、移動速度や位置などの自身の情報だけでなく、群れ全体の移動方向などの群れの情報も用いて行動している. PSO ではこのような生物の群れを問題空間上に再現することで、最適解を求める手法であり、群れを制御することで最適化を行う. これまで様々な特性を持つ群れの制御方法が提案されている. そこで、本研究では相補関係にある2種類の制御方法を用いる Complementary Control PSO (CCPSO) を提案する. また近年、注目されている遺伝子ネットワーク推定問題[2,3]に CCPSO を適用した結果、従来手法である GA や GP, PSO に比べて、十分な精度の解を短時間で得られたことを示す.

## 2. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO : 粒子群最適化) は Kennedy らによって提案された最適化手法であり[1], 鳥や魚などの群れを成して行動する生物を問題空間上に再現することで最適化を行う手法である. 各個体は粒子 (Particle) と呼ばれ、与えられた問題の解を表す位置 ( $x_i^k$ ) と探索方向を表す速度 ( $v_i^k$ ) を持っている. PSO は各粒子の速度を制御することで最適化を行う. 基本的な制御方法として式(1)がある.  $p_i$  は粒子  $i$  が過去に発見した最良解を示し,  $p_g$  は  $p_i$  の中の最良解を示す.  $rand_1, rand_2$  は [0,1] の一様性乱数を示し,  $c_1$  と  $c_2$  はパラメータである.  $w^k$  は慣性重みと呼ばれ, 式(2)で定義される.  $k$  は現在の移動回数,  $k_{max}$  は最大移動回数を示し,  $w_{max}, w_{min}$  はパラメ

### Complementary control particle swarm optimization and its application in inference of genetic network

Hiroaki MURATA<sup>†</sup>, Takamichi YONEYAMA<sup>‡</sup>,  
Masatomo MITAMURA<sup>†</sup>, Makoto KOSHINO<sup>‡</sup>,  
Haruhiko KIMURA<sup>†</sup>

<sup>†</sup> Kanazawa University,

Kanazawa City, Ishikawa, 920-1192, Japan

<sup>‡</sup> Ishikawa National College of Technology,

Kahoku-gun, Ishikawa, 929-0392, Japan

hiro@blitz.ec.t.kanazawa-u.ac.jp

ータである. なお、この制御方法を用いた PSO を Standard PSO (SPSO) と呼ぶ.

$$v_i^{k+1} = w v_i^k + c_1 rand_1 (p_i - x_i^k) + c_2 rand_2 (p_g - x_i^k) \quad (1)$$

$$w^k = w_{max} - \frac{w_{max} - w_{min}}{k_{max}} k \quad (2)$$

この他に、CPSO[1], PSO with Gaussian[4], PSOPC[5], HPSO-TVAC[6], Tribe-PSO[7], MCPSO[8]などの手法が提案されている.

## 3. Complementary Control Particle Swarm Optimization (CCPSO)

従来の PSO では1種類の制御方法のみを用いており、すべての粒子は同じ動作原理によって移動している. しかしながら、どのような状況においても、有効に機能する“最適な制御方法”は未だに発見されていない.

そこで、本研究では相補的な2種類の制御方法を用いる Complementary Control PSO (CCPSO) を提案する. 相補的な制御方法とは、一方の制御方法が不得意とする状況を、他方は得意とする制御方法の組であり、少なくともどちらか一方の制御方法が有効に機能し、効率的に最適化を行えると考えられる. CCPSO において、粒子は式(3)のように、どちらか一方の制御方法を用いる. ここで、相補的な2種類の制御方法を  $f_A(v)$  と  $f_B(v)$  で示した.

$$v_i^{k+1} = \begin{cases} f_A(v_i^k) & (rand < \alpha) \\ f_B(v_i^k) & else \end{cases} \quad (3)$$

$rand$  は [0,1] の一様性乱数を示し,  $\alpha$  はパラメータを示す. これにより、異なる制御方法を用いた粒子が混在する群れを実現し,  $\alpha$  を変化させることで状況に応じて制御方法を変更できる.

## 4. 遺伝子ネットワーク推定問題

生物や細胞の中では、様々な遺伝子が互いに制御しあい発現量を調節しながら生命活動を維持しており、それらの制御関係を明らかにすることで、新薬の開発や病気の治療に応用できる. 遺伝子ネットワーク推定問題とは、遺伝子の発現量の時系列データから、遺伝子の制御関係 (遺伝子ネットワーク) を推定することであり、与えられた遺伝子発現量の時系列データを再現する数理モデルを求めることが目的となる. 遺

伝子ネットワークを表す数理モデルとして、ベキ乗則に基づく S-system モデル[9]が用いられることが多く、進化的計算手法を用いて S-system モデルを推定する研究が行われている[2,3]. N 個の遺伝子から成る遺伝子ネットワークは  $2N(N+1)$  個のパラメータを含む S-system モデルで表現され、与えられた時系列データと、推定したモデルが生成した時系列データとの相対誤差の総和を目的関数とする最適化問題を解くことで、遺伝子ネットワークの推定を行う[2,3]. しかし、問題空間が  $2N(N+1)$ 次元となり、大規模な遺伝子ネットワークの推定は難しい. そのため、N 個の  $2(N+1)$ 次元の下位問題（各遺伝子に相当する）に分割する方法が提案されており[2]、本研究でも下位問題に分割し、推定を行う.

## 5. 評価実験

提案手法の有効性を確認するために、文献[3]と同様に、人工的に作成した5つの遺伝子から成る遺伝子ネットワークモデルを用いて評価実験を行った. 生成する際に用いたパラメータを目標パラメータと呼ぶ. なお、事前調査により下位問題に分割した場合、目標パラメータの評価値は 0 ではなく、0.026 になることが分かっている. これは評価値の計算時に、スプライン補間を用いて他の遺伝子の発現量を推定するために生じる誤差によるものと考えられ、本研究ではこの値を最適値と呼ぶ. CCPSO の設定として、制御方法  $f_A(v)$  を CPSO,  $f_B(v)$  を HPSO-TVAC とし、 $\alpha$  は 0.2 から 0.4 へと単調増加する値とした. 粒子数を 36, 最大評価回数を  $3.6 \times 10^4$  とし、40 回試行した結果を表 1 に示す. 表 1 には、得られたパラメータと目標パラメータとの誤差（絶対誤差の絶対値の総和）、最終評価値の平均値、評価値の平均値が最適値を下回った時の評価回数を示す. なお、GA と GP に関しては、文献[2]と[3]に記載されているパラメータから算出した.

実験結果より、パラメータの誤差に関しては、GA が最小であり、PSO 内で比較したとしても、CCPSO が最小ではない. しかしながら、平均評価値で比較すると、CCPSO が最小であり、この値は、GA や GP で得られたパラメータの評価値（最良評価値を示し、それぞれ 0.024 と 9.026 である）より良い値となっている. これは、下位問題に分割する際に生じる誤差も考慮したパラメータを CCPSO が求めたためだと考えられ、CCPSO の評価値が最適値より低いことから、そのような状況が伺える.

次に、評価値の平均値が最適値を下回った時の評価回数を見ると、CCPSO が最も少なく、CCPSO が目標パラメータと同等の評価値のパラ

表 1 : 遺伝子ネットワーク推定問題の実験結果

	誤差	平均 評価値	評価回数
SPSO[1]	21.8212	0.081	-
CPSO[1]	9.7974	0.032	-
PSO with Gaussian[4]	46.8058	7.144	-
PSOPC [5]	45.4033	0.738	-
HPSO-TVAC[6]	8.0423	0.019	33840
Tribe-PSO [7]	22.7007	0.093	-
MCPSO[8]	14.8116	0.069	-
CCPSO	11.6997	<b>0.010</b>	<b>24120</b>
GA[2]	<b>2.4630</b>	-	-
GP[3]	9.7319	-	-

メータを短時間で求めていることが分かる.

## 6. おわりに

本研究では、相補的な2種類の制御方法を用いた Complementary Control PSO を提案し、遺伝子ネットワーク推定問題に適用した. その結果、正確なパラメータの推定に関しては、従来手法より劣るという結果が得られた. しかしながら、最良の評価値のパラメータを求めており、さらに短時間で目標パラメータと同等の評価値のパラメータを得ており、実用的には有効であると考えられる. しかしながら、実際の時系列データには誤差が含まれており、今後は、誤差が含まれる時系列データによる遺伝子ネットワーク推定問題や、より大規模な遺伝子ネットワーク推定問題への適用を検討している.

## 参考文献

- [1] Kennedy. J., Eberhart. R. C. and Y. Shi., "Swarm intelligence," Morgan Kaufmann Publishers. San Francisco (2001)
- [2] Kimura. S., Hatakeyama. M. and Konagaya. A., "Inference of S-system Models of Genetic Networks from Noisy Time-series Data," Chem-Bio Informatics Journal, Vol. 4, No. 1, pp.1-14. (2004)
- [3] 松村幸輝, 種田洋, 木村周平, "遺伝的プログラミングを用いた関数最適化手法による遺伝子ネットワークの S-system モデルの推定," 情報処理学会論文誌, Vol.46, No.11, pp.2814-2830. (2005)
- [4] Higashi. N. and Iba. H., "Particle Swarm Optimization with Gaussian Mutation," Proc. of the 2003 IEEE Swarm Intelligence Symposium, pp.72-79. (2003)
- [5] He. S., Wu. C. Q., Saunders. R. J. and Paton. C. R., "A particle swarm optimizer with passive congregation," BioSystems, Vol. 78, pp. 135-147. (2004)
- [6] Ratnaweera. A., Halgamuge. S. K. and Watson. C. H., "Self-Organizing Hierarchical Particle Swarm Optimizer With Time-Varying Acceleration Coefficients," IEEE Trans. Evolutionary Computation, Vol.8, No.3, pp.240-255. (2004)
- [7] Chen. K., Li. T. and Cao. T., "Tribe-PSO: A novel global optimization algorithm and its application in molecular docking," Chemometrics and Intelligent Laboratory Systems, Vol.82, No.1-2, pp.248-259. (2006)
- [8] Nin. B., Zhu. Y., He. X. and Wu. H., "MCPSO: A multi-swarm cooperative particle swarm optimizer," Applied Mathematics and Computation. (掲載予定)
- [9] Savageau MA, "Biochemical systems analysis. A study of function and design in molecular biology," Addison-Wesley. (1976)