차분진화에 기초한 클러스터링

함서현[O**], 이현창[*], 신성윤[**]
원광대학교 디지털콘텐츠공학과[*]
군산대학교, 컴퓨터정보통신공학부[O]
군산대학교, 컴퓨터정보통신공학부[**]
e-mail: hamsh0210@naver.com[O**], hclglory@wku.ac.kr[*], s3397220@kunsan.ac.kr[**]

Differential Evolution Based Clustering

Seo-Hyun Ham[O**], Hyun-Chang Lee[*], Seong-Yoon Shin[**]
Dept. of Digital Contents Eng, Wonkwang University[*]
School of Computer Information & Communication Eng. Kunsan National University[O]
School of Computer Information & Communication Eng. Kunsan National University[**]

● 요 약 ●

Tensor decomposition, proven to be an efficient data processing method, can be used to provide data-driven services. we propose a novel datadriven mutation strategy for parent individuals selection, namely tensor-based DE with parapatric and cross-generation(TPCDE).

키워드: Tensor decomposition, tensor-based DE with parapatric and cross-generation(TPCDE)

I. Introduction

BIg data, which are digital data with the exponential growth and wide availability, are recognized with 4Vs characteristics (volume, velocity, variety, and veracity)[1].

II. TENSOR-BASED DE ALGORITHM

In this section, the parapatric and cross-generational selection scheme are proposed. The first scheme is to select the parapatric individuals from the current generations, the other is to select the cross-generational elite individuals. The algorithm mainly improves the parents selection in the mutation strategy, and enhances the diversity of the population. The most important components of the algorithm will be discussed as follows as follows.

2.1 Individual Distribution in DE

In DE, the population is initialized by the Eqs. (1). The individual is uniformly distributed within the constrained space. During evolution, the individuals will gradually be concentrated around certain solutions. In order to observe the distribution of individuals the population during the evolutionary process,

$$x_{i,j,G} = L_i + rand_{i,j}(0,1) \cdot (U_i - L_i), \quad (1)$$

2.2 Proposed Method

1) Parapatric Selection Scheme

As we all know, the performance of DE depends on the variation process which can enhance the exploration abilities of DE. The variation process is to obtain a perturbation in the search space to increase the population diversity. However, in the evolution process, the difference between individuals gradually decrease, resulting in the loss of population diversity. Based on the above considerations, the parapatric selection scheme (PSS) is proposed in this paper, the primary idea of which is partition it in different groups according to individual components, and then select different individuals from the other groups to perform mutation.

Algorithm The Parapatric Selection Scheme
Input: Population size: NP, Groups: C, Threshold: .
Output: r is the chosen indexes of the population.
1: for i ← 1 to NP do
2: m ← find(C==i)
3: Ai ← select individuals =∈ C(m)
4: if rand < then
5: Select r1/= r2/= r3/= i from the set Ai.
6: else
7: Select uniform randomly r1/= r2/= r3/= i.

8: end if
9: r(i; :) ← r1; r2; r3.
10: end for
11: Return the chosen indexes r.

2) Cross-Generation Selection Scheme

Adaptive Cross-generation differential evolution(ACGDE), proposed by Qiu [32], is one of novel mutation strategies which employs information across generations to help guide the searching directions. Unlike classic DE, ACGDE mutation strategy utilize the 6 information from current and previous generation to generate the donor vector. On that basis, we propose a novel cross-generational selection scheme (CGSS), which is mainly different in two aspects. First, the historical population of previous three generations has been stored into an archive. Second, the individuals of cluster centers are selected from the archive to generate the mutant vector.

3) TPCDE

According to the above analysis, the pseudo-code of TPCDE is combines the parapatric-based selection and the cross-generational selection to enhance the performance of DE. In TPCDE, the population tensor is firstly constructed, then the individuals closest to the optimal solution are stored in the elite collection, and the population was clustered into group C by combing tensor and AP algorithm, finally the parent individuals is selected through a adaptive mechanism.

III. Experiments

Table 1. Average ranking based on the Friedman test

Algorithm	DE	OXDE	CoDE	jDE	TPCDE
Ranking	4.85	3.23	3.08	2.85	2.08

IV. Conclusions

In this paper, we have proposed a tensor-based differential evolution with parapatric and cross-generation selection (TPCDE). which provides proper strategies for the population evolution through data-driven methods.

REFERENCES

[1] X. Wang, L. T. Yang, H. Liu, and M. J. Deen, “A tensor computation and optimization model for cyber-physical-social big data,” IEEE Transactions on Sustainable Computing, DOI: 10.1109/TSUSC. 2017.2777503, 2017.