K-means Algorithm Using an Improved Firefly Algorithm Applied Simple Toy Model

Masaki Takeuchi
Dept. Electrical
and Electronic Eng.,
Tokushima University
masaki@ee.tokushima-u.ac.jp

Thomas Ott
Zurich University
of Applied Sciences
thomas.ott@zhaw.ch

Haruna Matsushita
Dept. Electronics
and Info. Eng.,
Kagawa University
haruna@eng.kagawa-u.ac.jp

Yoko Uwate and Yoshifumi Nishio
Dept. Electrical
and Electronic Eng.,
Tokushima University
{uwate, nishio}@ee.tokushima-u.ac.jp

I. Introduction

Senthilnath et. al. [1] proposed an algorithm that used the firefly algorithm for K-means clustering (KMFA). Numerical experiments have indicated that this algorithm is more efficient algorithm than the standard algorithm or other optimization heuristics. In this paper, we introduce a new clustering algorithm that combines K-means clustering with improved firefly algorithm (KMIFA). In our algorithm, one parameter is changed when the assignment does not change. We compare the conventional K-means algorithm, KMFA and KMIFA using a toy data model. [1] These experiments indicate that our algorithm is more efficient than the other algorithms.

II. KMIFA

The K-means algorithm sometimes converges to a local minimum. Therefore, the purpose of this study is to remove this disadvantage. In our algorithm, each firefly has its own value of $\alpha(0)$ and we set all values of $\alpha(0)$ to 1.0 when initializing the population of fireflies. In the case of a firefly i, if the assignment of all objects does not change, the value of $\alpha(0)$ of firefly i decreases. We define the minimum value of $\alpha(0)$ is 0.0.

$$\alpha(0)_i^{new} = \begin{cases} \alpha(0)_i & , \textit{the assignment changes} \\ \alpha(0)_i - V & , \textit{otherwise} \end{cases} \quad (1)$$

The parameter V is a predefined value. In the case of $\alpha(0) = 1.0$, a firefly moves with a relatively strong random influence. This makes the firefly easier to escape from a local minimum. In the case of $\alpha(0) = 0.0$, a firefly does not move randomly, which leads to a faster convergence.

III. Numerical Experiments

We compare the conventional K-means algorithm, KMFA and KMIFA using a simple data toy model. The number of dimensions is 2, the range of each dimension is [0, 100], the number of clusters is 6 and the number of objects is 120 (20 objects per cluster). For the data generation, the ideal cluster centers were placed at (30, 30), (70, 70), (10, 80), (90, 20), (50, 90) and (70, 70). Then, the data objects were generated randomly around the ideal centers within a ball of radius 10. Each numerical experiment was run 300 times and we compared the success rate of each algorithm, where the success rate is defined as the fraction of objects that are assigned to the correct cluster.

[1]The extended version of this study is being reviewed for NCSP2017 [2].

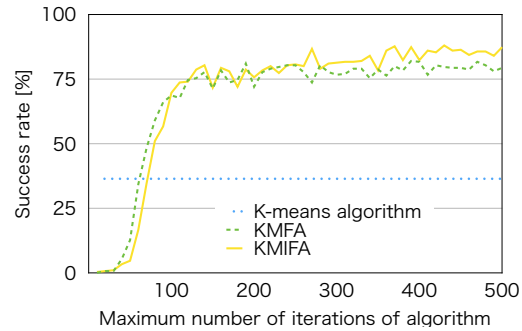

Fig. 1. Comparison of the three algorithms

Figure 1 shows the numerical results for the parameter $V = 0.1$ and 20 fireflies. We assume that our algorithm is more efficient algorithm than the other two algorithms.

IV. Conclusion

In this study, we have introduced a new clustering algorithm that utilizes an improved firefly algorithm for K-means clustering. Our algorithm is based on the idea that the randomization parameter is changed when the assignment does not change. In our algorithm, at the beginning of the search, all fireflies move with a relatively strong random influence. Hence they can more easily escape from a local minimum. As the number of iterations increases, the firefly tend to converge. Numerical experiments have indicated that our algorithm is more efficient than K-means algorithm and KMFA.

References

[1] J. Senthilnath, S.N. Omkar and V. Mani, *"Clustering using Firefly Algorithm: Performance Study"*, Swarm and Evolutionary Computation 1, pp. 164–171, 2011.

[2] M. Takeuchi, T. Ott, H. Matsushita, Y. Uwate and Y. Nishio, *"K-means Algorithm using Improved Firefly Algorithm"*, RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing, 2017. (being submitted)