## An anomaly detection approach based on hybrid differential evolution and K-means clustering in crowd intelligence

Jianran Liu

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China and Fujian Key Laboratory of Granular Computing and Application, Minnan Normal University, Zhangzhou, China

**Bing Liang** 

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China and University of Chinese Academy of Sciences, Beijing, China, and

### Wen Ji

Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

#### Abstract

**Purpose** – Artificial intelligence is gradually penetrating into human society. In the network era, the interaction between human and artificial intelligence, even between artificial intelligence, becomes more and more complex. Therefore, it is necessary to describe and intervene the evolution of crowd intelligence network dynamically. This paper aims to detect the abnormal agents at the early stage of intelligent evolution.

**Design/methodology/approach** – In this paper, differential evolution (DE) and K-means clustering are used to detect the crowd intelligence with abnormal evolutionary trend.

Findings – This study abstracts the evolution process of crowd intelligence into the solution process of DE and use K-means clustering to identify individuals who are not conducive to evolution in the early stage of intelligent evolution.

**Practical implications** – Experiments show that the method we proposed are able to find out individual intelligence without evolutionary trend as early as possible, even in the complex crowd intelligent interactive environment of practical application. As a result, it can avoid the waste of time and computing resources.

**Originality/value** – In this paper, DE and K-means clustering are combined to analyze the evolution of crowd intelligent interaction.

**Keywords** Differential evolution, K-means, Anomaly detection, Crowd intelligence, Intelligence evolution

Paper type Research paper

© Jianran Liu, Bing Liang and Wen Ji. Published in *International Journal of Crowd Science*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/ licences/by/4.0/legalcode

This work is supported by the National Key R&D Program of China (2017YFB1400100), and the Beijing Natural Science Foundation (4202072).

K-means clustering in crowd intelligence

129

Received 17 July 2020 Revised 15 September 2020 Accepted 22 September 2020



International Journal of Crowd Science Vol. 5 No. 2, 2021 pp. 129-142 Emerald Publishing Limited 2398-7294 DOI 10.1108/IJCS-07-2020-0013

#### IJCS 1. Introduction

5.2

130

Before the concept of crowd intelligence came into being, most of the work is based on the measurement of single agent intelligence. With the rapid development of artificial intelligence, the cooperation and collaboration between individuals of artificial intelligence make the advantages of crowd intelligence more and more prominent. The measurement of monomer intelligence can no longer meet the needs of development. Crowd intelligence with different structure obviously has different intelligence degree. In the previous work, the research on human crowd intelligence (Anita *et al.*, 2010) has become the cornerstone of the measurement method of crowd intelligence. IQ testing (David and José, 2012) has also been improved to measure agents. Based on the quality time complexity system, Ji *et al.* (2018) implement the measurement method of crowd intelligence. This transcendental knowledge has inspired us to further explore crowd intelligence.

With the passage of time, the amount of intelligence of crowd network should present a changing state. This state may evolve with the cooperation between individuals, or degenerate with the competition or even confrontation between individuals. For example, when quantifying two competing individual intelligence, the quantized value should be equal to the individual intelligence with large amount of intelligence, which is obviously not conducive to the evolution of crowd intelligence. In particular, it is a challenging problem to predict and find out the individual intelligence with abnormal cooperative mode without knowing the cooperative relationship between them.

In crowd intelligence network, the first task is to establish the evolution model among individuals. Because over time, individual intelligence evolve (or degenerate) to obtain the optimal value of the entire evolutionary model. However, in the process of evolution (or degradation), individual intelligence does not necessarily follow the monotonous greedy strategy to change the amount of intelligence. Moreover, intelligence needs to add perceptual factors to the crowd intelligence. The final task is to cluster the intelligent population in the early process of evolution, which can intervene the individual intelligence in the unfavorable cooperative relationship early and avoid the loss of time and computing power.

In this paper, we propose a series of methods to describe the evolution and detect abnormalities in crowd intelligence. First of all, taking crowd intelligence with three kinds of individual intelligence as an example, we define the mathematical model corresponding to the interaction mode of crowd intelligence. This method can be derived to an intelligent crowd with multiple agents. Then, we abstract the evolution of crowd intelligence into a polynomial approach to the optimal solution. In particular, we integrate the perceptual factors of individual intelligence into the polynomial approximation, use the differential evolution (DE) method to get the optimal solution of the polynomial. Finally, we use K-means clustering to distinguish individual intelligence in the early stage of evolution. Experimental results show that our proposal can effectively describe the dynamic evolution process of crowd intelligence and effectively distinguish the individual intelligence with interaction problems in the early evolution.

The rest of the paper is organized as follows. Section 2 describes the details of DE and Kmeans clustering. In Section 3, we describe the evolutionary method of crowd intelligence based on DE and K-means clustering. Section 4 shows the simulation experiment and its results. Finally, the conclusion and future work is shown in Section 5.

## 2. Differential evolution and K-means clustering for the evolution of crowd intelligence

To better build and improve the evolution model of crowd intelligence on the basis of existing knowledge, in this section, we present the basic characteristics of DE model and K-means clustering.

#### 2.1 Differential evolution model

DE algorithm is first proposed by Storn *et al.* in 1995 (Storn and Price, 1997). The original idea is to use the de method to find the optimal solution of Chebyshev polynomials. Then it is found that the DE method is also very effective for complex optimization problems. DE algorithm is one of the most widely used stochastic real parameter optimization algorithms. DE algorithm is operated by similar calculation steps as standard evolution algorithms (Zhou *et al.*,2011). However, different from the standard evolutionary algorithm, the differential evolutionary algorithm incorporates the possibility of mutation, which interfere with the current generation of population members and randomly select different population members. Instead of using a separate probability distribution to produce offspring. DE method based on population random search technology (Bergstra and Bengio, 2012), which is used to solve optimization problems in continuous space, has been widely used in many scientific and engineering fields, such as in financial prediction by Ardia *et al.* (2011), Chauhan *et al.* (2011) and in automatic driving by Fu *et al.* (2012).

For a DE model, the following basic elements and behaviors are required:

- *Optimization problem:* Generally, the optimization problem is abstract from the actual problem and designed as a bounded function. The purpose of DE is to get a set of solutions and find the upper (lower) bound of the optimization problem.
- *Scope*: To prevent the function of the designed optimization problem from having no maximum (minimum) value in the wireless domain, at the same time, the optimal solution of the objective optimization problem needs to be found in a limited range in the actual problem.
- *Generation*: Finding the optimal solution of optimization problem is a process of gradual approximation, which needs to be completed by iteration. A generation represents a set of values (variables) for optimization problems. Generally, compared with the previous generation, each new generation can make the optimization problem get better solution. However, in the evolution and iteration process of practical problems, each agent may have some "perceptual" factors, which makes the iteration cannot always develop in an ideal state incremental situation.
- *Population:* The concept of population abstraction from biology, which originally refers to all individuals of the same species occupying a certain space in a certain period of time. In the field of crowd intelligence, the purpose of population design is to simulate the interaction and mutual learning of agents or another kind of population, so as to achieve the purpose of simulation evolution. Mutation and crossover within the population and selection based on optimization problem drive the evolution of crowd intelligence. Generally, population *P* at iteration *g* in DE is described as:

$$P^{g} = \left[X_{1}^{g}, X_{2}^{g}, \dots, X_{N_{p}}^{g}\right],\tag{1}$$

the *i*th individual is:

$$X_{i}^{g} = \left[X_{i,1}^{g}, X_{i,2}^{g}, \dots, X_{i,D}^{g}\right], \quad i = 1, 2, \cdots N_{p},$$
(2)

K-means clustering in crowd intelligence where  $X_{ij}^g$  represents the *j*th dimension of the *i*th individual in the *g* generation. *D* represents the dimensions of individual decision variables. Before the start of DE, individuals in the population are actually generated randomly within the definition domain, as:

$$X_{i,j} = \operatorname{rand}\left(X_{j}^{\min}, X_{j}^{\max}\right), \quad i = 1, 2, \cdots, N_{p}; j = 1, 2, \cdots, D,$$
 (3)

where rand (*a*,*b*) represents the uniform distributed value between *a* and *b*.

• Mutation: Mutation is an important sign different from traditional genetic algorithm (Deb *et al.*, 2002). In the population, the DE model realizes individual variation through the differential strategy. The mutant individual V is generated by the current individual X through:

$$V_{ij}^{g} = X_{r_{1}j}^{g} + \mu \left( X_{r_{2}j}^{g} - X_{r_{3}j}^{g} \right), \quad i = 1, 2, \cdots, N_{p}; j = 1, 2, \cdots, D,$$
(4)

where  $\mu$  is the mutation factor and  $r_1 \neq r_2 \neq r_3 \in [1, N_p]$ . According to the progress of generations and algorithm search, the speed of evolution should also be adjusted adaptively. The idea is to diversity of the population in the early stage of evolution and avoid getting the "optimal" value considered by generations at that time too early. In the later stage of evolution, the mutation factor  $\mu$  should also gradually converge to keep the good information away from damage.

• *Crossover*: The above mutation is not carried out between each individual in the population. After mutation, not every population has cross behavior with its variants. In other words, in the new generation, the proportion of variation part and the proportion of complete genetic part need to be fixed, this behavior becomes cross operation. The temporary vector  $U_{i,j}$  formed by the crossover operation is defined as:

$$U_{ij}^{g} = \begin{cases} V_{ij}^{g}, \operatorname{rand}(0, 1) \leq Cr\\ X_{ij}^{g}, \text{ otherwise} \end{cases},$$
(5)

where Cr is the crossover probability parameter.

Selection: DE algorithm substitute the generation after mutation crossover into the
optimization equation. Theoretically, greedy algorithm is used. The value with the
largest value is selected as the real next generation, as:

$$X_i^{g+1} = \begin{cases} U_i^g, f(U_i^g) \le f(X_i^g) \\ X_i^g, \text{ otherwise} \end{cases},$$
(6)

where  $f(s^2)$  means the optimization problem. However, in practical problems, individual intelligence or populations may have "perceptual" psychology, which leads to not adopting the greedy strategy when choosing the optimal generation, resulting in not reaching the ideal maximum value. The entire selection strategy evolved into:

132

**IICS** 

5,2

$$X_i^{g+1} = \begin{cases} U_i^g, \operatorname{rand}(0,1) \le W(f(U_i^g) - f(X_i^g)) \\ X_i^g, \text{ otherwise} \end{cases}, \tag{7} \qquad \text{K-means} \\ \text{clustering in} \end{cases}$$

where  $W(\cdot)$  is a perceptual measurement function with a range of 0 to 1, which makes the selection of the optimal generation not follow the greedy algorithm.

The basic elements of DE can be defined and adjusted in terms of evolution objectives, evolution conditions, iteration times and selection strategies according to the actual application requirements. In the next section, we state the DE model for crowd intelligence evolution and the adjustment details for this application.

#### 2.2 K-means clustering

Clustering data (Jain et al., 1999) according to rules is one of the most basic patterns of machine learning and machine understanding. Clustering analysis is the study of the methods and algorithms of grouping or clustering according to the measured or classified objects. The internal characteristics or similarity of perception. In particular, clustering analysis do not use prior identification and category label, its essence is to explore the potential mapping and structure in the data.

Among the clustering methods, K-means clustering is one of the most popular and simple clustering algorithms. Although this algorithm is born in 1955 and is derived a variety of clustering algorithms (Selim and Ismail, 1984; Huang et al., 2005; Krishna and Murty, 1999), K-means clustering is still widely used because of its crowdity. At first, Kmeans clustering originated from the field of signal processing, which is a vector quantization method. Now, it is more popular in the field of data mining as a data analysis method. Generally, the n-attributes of data are mapped to the n-dimensional space and their K centroids are initialized randomly.

For the evaluation of K-means clustering results, a method originated from information theory is widely used: mutual information (McKay et al., 2008). Mutual information is a measure of information, defined as J(X; Y), which means that random variable X contains information about another random variable Y. Let the joint probability distribution of two random variables (x, y) be p(x, y), the marginal probability distribution be p(x) and p(y). The mutual information J(X; Y) is the relative entropy of the product p(x, y) of the joint distribution p(x, y) and the marginal distribution p(x) p(y), defined as:

$$J(X;Y) = \sum_{x} \sum_{y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}.$$
(8)

In order to solve the problem that K-means clustering between different distributions is difficult to compare, the mutual information J(X; Y) is normalized to obtain the clustering evaluation index: normalized mutual information (NMI) Estevez et al. (2009), defined as:

$$NMI(X;Y) = 2\frac{J(X;Y)}{H(X) + H(Y)},$$
(9)

where H(X) is the information entropy of the random variable X.

#### 3. Improved differential evolution model

The optimization function in the DE model should be adjusted according to the different definition of the interaction mode of different agents to simulate the intelligent interaction 133

intelligence

in crowd

behavior (including cooperation, competition, confrontation and no behavior) in the evolution process of intelligent crowds. Different intelligence measurement means that the measurement formula of crowd intelligence is also different. Figure 1 shows the reference calculation formula of crowd intelligence when there are only two individuals in the intelligent crowd, where *a* represents the amount of intelligence of individual intelligence. Note that this measurement method is not unique and can be adjusted according to the actual measurement.

According to the rule of Figure 1, the crowd intelligence calculation formula is extended to the intelligent crowd with three agents. Figure 2 traverses all possible crowd intelligence interactions with three intelligent individuals. The tree traversal is a display of the calculation of combination with repetition, because we default that each individual intelligence is equivalent. For each possible crowd intelligence interactions, we demonstrate the corresponding evolutionary formula. In particular, the evolution formula marked with "\*" are the target of K-means clustering, for the following reasons:

- In all interactions, cooperative and non-cooperative relationships must exist simultaneously.
- The number of non-cooperative relationships must be two, so there must be an abnormal population.

The variation and crossover mechanism in the framework of DE is the characteristic of DE. In our improvement, the crossover and variation mechanisms remain almost unchanged. We replace the invariable greedy strategy in the selection mechanism with a perceptual strategy. For this new strategy, we consider that Fermi-Dirac function (Dirac, 1926) is suitable for simulate the evolution rule of sensibility. Fermi-Dirac function is derived from physics, which represents the probability of electron occupying the eigenstate of a certain energy, is applied to statistical mechanics (Sho *et al.*, 2016). The Fermi-Dirac function used as the rule of perceptual evolution is defined as:

$$W(\Delta u) = \frac{1}{1 + e^{-(\Delta u)/\kappa}},$$
(10)

where,  $\Delta u$  represents the intelligence difference between the current generation of agents and the next generation of agents and  $\kappa$  represents noise, which is used to describe a perceptual behavior. In the evolution of intelligence, Fermi-Dirac function means that the intelligence population and the intelligence population of the next generation are compared in terms of revenue. The strategy of transforming one probability into the next generation is adopted. Figure 3 shows the effect of  $\kappa$  on the strategy conversion rate in the range of  $\Delta u \in (-10, 10)$ . When the noise tends to 0, the whole evolutionary selection strategy tends

Figure 1. Graph of the population with only two individuals and the calculation formula of intelligence quantity corresponding to them

		Cooperation	Noninteractive	Competition	Aggressive
, .	Illustration	<b>⊶⊙⊸</b>	<b>0-</b> ()-0	<b>-⊕-</b> 0	-⊗0
_	Formula	$a_1 \times a_2$	0	$  \max(a_1, a_2) - a_1 - a_2$	$  \operatorname{abs}(a_1 - a_2) - a_1 - a_2$

134

**IICS** 

5.2





Figure 3. Fermi-Dirac function curve under different noise influence

to greedy algorithm. In the new era, the optimal evolutionary strategy is almost decisively adopted. When the noise tends to be positive infinity, the whole evolutionary selection strategy tends to be random. The evolutionary strategy is selected by an "arbitrary" chaos method, regardless of whether this strategy is the best choice.

Taking  $\kappa = 1$  as an example, when  $\Delta u = -2$ , it means that even if the intelligence of the new generation is not as good as that of the old generation, there is still about 10% possibility to choose the intelligence strategy of the new generation. When  $\Delta u = 2$ , it means that even if the intelligence of the new generation is greater than that of the old generation, only about 90% of the new generation is likely to choose the intelligence strategy of the new generation. The larger the noise is, the more disadvantageous it is to detect the degenerated individuals in the early stage of evolution. However, the K-means clustering method can avoid the noise to a great extent and achieve the purpose of detecting the abnormal interactive individuals, which is proved experimentally in the following sections.

#### 4. Details and analysis of simulation experiments

#### 4.1 Design of simulation experiment

In order to verify the effectiveness of the improved DE model, we design the DE model according to the crowd intelligence calculation formula listed in Figure 2 and the following parameters:

- The maximum number of iterations is set to 1000;
- The variation rate is set as 0.5 (random variation);
- The crossing probability is set to 0.9;
- According to the formula in Figure 2, there are three individual intelligence, so the dimension of the problem is three dimensions;
- The intelligent range of the initial individual intelligence is 0 to 10;
- The noise parameter  $\kappa$  is 60. Moreover, the noise parameter is adjusted in the next parameter sensitivity analysis experiment;
- The crowd intelligence calculation formula with "\*" label in Figure 2 is selected for optimization problem.

Figure 4 illustrates the evolution of intelligent crowds with different interaction modes in the evolution process of 1000 generations. It also shows the intelligent evolution trend of each individual intelligence. After some fluctuations in the early evolution of crowd intelligence, the trend of change is gradually stable. Most of crowd intelligence stop evolve after 600 evolutions. In the early stage of evolution, the change of individual intelligence is chaotic. It is difficult to detect the individual intelligence with interaction problems. This kind of interaction problem is obviously exposed at the end of evolution, but in reality, it is too late to find the problem at the end of evolution. Moreover, the number of individual intelligence may be far more than three in the actual crowd intelligent interaction process and the definition of the model is more complex, so it is difficult to directly analyze the model.

K-means clustering can help to mining the individual intelligence that hinder the evolution of crowd intelligence in the early stage of evolution and avoid the time and energy costs in the later stage of evolution. In the actual evolution of intelligent crowds with more individual intelligence, K-means clustering can find abnormal individual intelligence without knowing the evolution framework of crowd intelligence.

To verify the effectiveness of K-means clustering in abnormal individual intelligence detection, the following experiments are designed:

IICS

5.2



# • In order to avoid the evolutionary failure caused by small probability events in the evolution process of crowd intelligence, according to the evolution experiment of crowd intelligence in Figure 4 as the blueprint, 200 experiments are repeated for each evolution framework;

- In each group of experiments, five generations of individual intelligence are randomly selected from the first 100 generations as samples. In the same crowd intelligence framework, each individual intelligence samples 1000 generations in total, each 5 generations are individual intelligence samples under the same crowd intelligence framework, which are listed as a 5-Dimensional feature tensor, 200 such feature tensors in total and 600 feature tensors in total for three kinds of individual intelligence;
- 600 feature tensors are clustered into two categories by K-means. NMI is used as the evaluation index to calculate the clustering accuracy;
- Analyze the feasibility of K-means clustering to detect the anomaly of individual intelligence in crowd intelligence.

#### 4.2 Experimental results and analysis

For the crowd intelligence with "\*" label in Figure 2, five simulation experiments of DE and K-means clustering are conducted with the calculation formula to calculate the evaluation index NMI of each clustering. Figure 5 shows the results of K-means clustering. The results show that the effect of K-means clustering in the early stage of DE is different for different intelligent calculation formulas. The clustering results of NMI above 0.6 are generally good, which can accurately distinguish abnormal individual intelligence from normal individual intelligence, while the clustering results of NMI below 0.3 show that K-means clustering cannot detect abnormal crowd intelligence of this kind of framework very well. However, in our proposed method, there are many factors affecting the result of k-mean clustering. For example, the clustering result is not only affected by the swarm intelligence evolution framework and its calculation formula, but also by the noise parameter  $\kappa$ . The experimental results shown in Figure 5 consider  $\kappa = 60$  only, which may have a significant impact on some noise-sensitive frameworks.

In order to verify the influence of noise parameters on the evolution of crowd intelligence and the detection of abnormal intelligence individuals, we performed parameter sensitivity analysis experiment. Figure 6 shows the results of the noise parameter sensitivity analysis. Among them, the crowd intelligence evolution model with "\*" in Figure 2 is still selected to analyze the impact of noise parameters on the detection of abnormal intelligence individuals. The value range of noise parameter  $\kappa$ is 0.1 to 99. The results show that the evolution performance of Model (1) is almost unaffected by noise. Because there is a population that do not participate in the calculation of objective function in the model. For Model (5), the influence of noise on the detection of abnormal individuals could not be detected because the abnormal individuals could not be excluded accurately. The abnormal clustering performance of other models is affected by noise. Among them, when the noise parameter  $\kappa$ increases, the degree to which the Models (2), (4) and (6) are affected by noise are all different. Model (2) when  $\kappa = 10$ , the evaluation parameter NMI of anomaly detection begins to show a downward trend. Model (4) when  $\kappa = 20$ , the value of NMI gradually decreases as  $\kappa$  increases. Model (6) has strong robustness. When  $\kappa = 30$ , the accuracy of anomaly detection begins to show a downward trend, indicating that it is very little

**IICS** 

5.2



affected by noise. When the noise parameter  $\kappa > 0$ , it starts to affect the anomaly detection accuracy of Model (3), and the highest detection accuracy of Model (3) is only 0.8. Subsequently, when  $\kappa > 50$ , the clustering accuracy performance of Model (3) began to stabilize, and was no longer significantly affected by noise.



K

#### 5. Conclusion and future work

For the crowd intelligence with interactive behavior, its total intelligence and the intelligence of each individual intelligence should be in a dynamic state. Different from the previous static description of crowd intelligence, this paper proposes a differential dynamic intelligent evolution model. In particular, our method can cluster the individuals by K-means with abnormal evolution in the early stage of evolution. The effectiveness and robustness of our method are verified by experiments and noise sensitivity analysis. In the future, we will continue to explore the evolution of swarm intelligence interaction based on the existing work of crowd intelligence evolution and modeling more complex application scenarios.

K-means clustering in crowd intelligence

#### References

- Anita, W.W., Christopher, F.C., Alex, P., Nada, H. and Thomas, W.M. (2010), "Evidence for a collective intelligence factor in the performance of human groups", *Science*, Vol. 330 No. 6004, pp. 686-688.
- Ardia, D., Boudt, K. and Carl, P. (2011), "Differential evolution with DEoptim: an application to nonconvex portfolio optimization", *The R Journal*, Vol. 3 No. 1, pp. 27-34.
- Bergstra, J. and Bengio, Y. (2012), "Random search for hyper-parameter optimization", Journal of Machine Learning Research, Vol. 13 No. 1, pp. 281-305.
- Cai, H.R., Chung, C.Y. and Wong, K.P. (2008), "Application of differential evolution algorithm for transient stability constrained optimal power flow", *IEEE Transactions on Power Systems*, Vol. 23 No. 2, pp. 719-728.
- Chauhan, N., Ravi, V. and Chandra, D.K. (2010), "Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks", *Expert Systems with Applications*, Vol. 36 No. 4, pp. 7659-7665.
- David, L.D. and José, H. (2012), "IQ tests are not for machines", Yet. Intelligence, Vol. 40 No. 2, pp. 77-81.
- Deb, K., Pratap, A. and Agarwal, S. (2002), "A fast and elitist multi-objective genetic algorithm: NSGA-II", *IEEE Transactions on Evolutionary Computation*, Vol. 6 No. 2, pp. 182-197.
- Dirac, P.A.M. (1926), "On the theory of quantum mechanics", Proceedings Mathematical Physical and Engineering Sciences, Vol. 112 No. 762, pp. 661-667.
- Estevez, P.A., Tesmer, M. and Perez, C.A. (2009), "Normalized mutual information feature selection", *IEEE Transactions on Neural Networks*, Vol. 20 No. 2, pp. 189-201.
- Fu, Y., Ding, M. and Zhou, C. (2012), "Phase angle-encoded and quantum-behaved particle swarm optimization applied to three-dimensional route planning for UAV", *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, Vol. 42 No. 2, pp. 511-526.
- Huang, J.Z., Ng, M.K. and Rong, H. (2005), "Automated variable weighting in K-means type clustering", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27 No. 5, pp. 657-668.
- Jain, A., Murty, M. and Flynn, P. (1999), "Data clustering: a review", ACM Computing Surveys, Vol. 31 No. 3, pp. 264-323.
- Ji, W., Liu, J., Pan, Z., Xu, J., Liang, B. and Chen, Y. (2018), "Quality-time-complexity crowd intelligence measurement", *International Journal of Crowd Science*, Vol. 2 No. 1, pp. 18-26.
- Kitayama, S., Arakawa, M. and Yamazaki, K. (2011), "Differential evolution as the global optimization technique and its application to structural optimization", *Applied Soft Computing*, Vol. 11 No. 4, pp. 3792-3803.
- Krishna, K. and Murty, M.N. (1999), "Genetic K-means algorithm", *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, Vol. 29 No. 3, pp. 433-439.
- McKay, M.R., Smith, P.J. and Suraweera, H.A. (2008), "On the mutual information distribution of DMbased spatial multiplexing: exact variance and outage approximation", *IEEE Transactions on Information Theory*, Vol. 54 No. 7, pp. 3260-3278.

IJCS 5,2	Selim, S.Z. and Ismail, M.A. (1984), "K-Means type algorithms: a generalized convergence theorem and characterization of local optimality", <i>IEEE Transactions on Pattern Analysis and Machine</i> <i>Intelligence</i> , Vol. 6 No. 1, pp. 81-87.
	Sho, S., Odanaka, S. and Hiroki, A. (2016), "A simulation study of short channel effects with a QET model based on Fermi-Dirac statistics and non-parabolicity for high-mobility MOSFETs", <i>Journal of Computational Electronics</i> , Vol. 15 No. 1, pp. 76-83.
142	Storn, R. and Price, K. (1997), "Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces", <i>Journal of Global Optimization</i> , Vol. 11 No. 4, pp. 341-359.
	Zhou, A., Qu, B., Li, H., Zhao, S. and Zhao, Q. (2011), "Multiobjective evolutionary algorithms: a survey of the state of the art", <i>Swarm and Evolutionary Computation</i> , Vol. 1 No. 1, pp. 32-49.

#### Corresponding author

Wen Ji can be contacted at: jiwen@ict.ac.cn

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm Or contact us for further details: permissions@emeraldinsight.com