Adaptive Crossover Memetic Differential Harmony Search for Optimizing Document Clustering

Ibraheem Al-Jadir[1, 2], Kok Wai Wong[1], Chun Che Fung[1], and Hong Xie[1]

[1]School of Engineering and Information Technology, Murdoch University, Perth, Australia 6163
[2]College of Science, Baghdad University, Baghdad, Iraq
{I.Al-Jadir, K.Wong, L.Fung, H.Xie} @murdoch.edu.au

Abstract. An Adaptive Crossover Memetic Differential Harmony Search (ACMDHS) method was developed for optimizing document clustering in this paper. Due to the complexity of the documents available today, the allocation of the centroid of the document clusters and finding the optimum clusters in the search space are more complex to deal with. One of the possible enhancements on the document clustering is the use of Harmony Search (HS) algorithm to optimize the search. As HS is highly dependent on its control parameters, a differential version of HS was introduced. In the modified version of HS, the Band Width parameter (BW) has been replaced by another pitch adjustment technique due to the sensitivity of the BW parameter. Thus, the Differential Evolution (DE) mutation was used instead. In this paper the DE crossover was also used with the Differential HS for further search space exploitation, the produced global search is named Crossover DHS (CDHS). Moreover, DE crossover (Cr) and mutation (F) probabilities are dynamically tuned through generations. The Memetic optimization was used to enhance the local search capability of CDHS. The proposed ACMDHS was compared to other document clustering techniques using HS, DHS, and K-means methods. It was also compared to its other two variants which are the Memetic DHS (MDHS) and the Crossover Memetic Differential Harmony Search (CMDHS). Moreover, two state-of-the-art clustering methods were also considered in comparisons, the Chaotic Gradient Artificial Bee Colony (CGABC) and the Differential Evolution Memetic Clustering (DEMC). From the experimental results, it was shown that CMDHS variant (the non-adaptive version of ACMDHS) and ACMDHS were highly competitive while both CMDHS and ACMDHS were superior to all other methods.

Keywords: Clustering, Memetic, Optimization, and Harmony Search

1 Introduction

Document clustering is a necessary process for efficient document management such as retrieval, archival, topic extraction, and summarization [1]. Due to the drawbacks of the traditional methods of clustering, optimization methods have been used [2]. Evolutionary Algorithms (EA) were introduced to optimize the selection of the centroids in document clustering [3]. In the last few years, a large number of optimization-based methods have been proposed [4]. As an improvement to the evolutionary algorithms, Memetic algorithm (MA) [5] have been proposed to combine the advantages of the evolutionary algorithms with problem-specific optimization methods. The powerful global search ability is combined with some local improver methods in the memetic algorithms. Such a combination have been successfully applied to global optimization of numerical functions [6] and have been utilized to solve many real-world optimization problems [7]. MA has been introduced in many other applications but has not been investigated in detail for optimizing document clustering.

For making MA more powerful in terms of the exploration of the search space, modifications to MA's global search could also be productive. For instance, combining two or three global search methods might be more efficient than the old methods. For example, the use of Differential Evolution (DE) Mutation with Harmony Search instead of the pitch adjustment step can produce better optimizing performance. Furthermore, the use of the adaptive parameter settings can also enhance the performance in comparison to the performance of the static parameter setting [8, 9]. These improvements can all be combined to produce an efficient method for document clustering.

In this paper, an intelligent document clustering method is proposed that uses the Differential HS as an optimization method to better search centroids of the document clusters. DHS is used as a modification to the original HS that uses bandwidth technique for the harmonies' pitch adjustment. In this paper, DHS is used to eliminate the need for manually setting the Bandwidth (BW) parameter as it might affect document clustering performance negatively. Also in this paper, DHS was further enhanced by incorporating the DE crossover along with the mutation; the modified version was named Crossover DHS (CDHS). The mutation scaling factor (F) and the crossover probability rate (Cr) are adaptively set. In that case, whenever an optimal solution is generated, the best control parameters contributed to the best solution are admitted to the next generations. Finally, the memetic optimization is used with CDHS to reduce the chances of entrapment in local optima. The resulted method is an Adaptive Crossover Memetic Differential Harmony Search (ACMDHS). Moreover, the other two variant of the ACMDHS resulted in this paper are the Memetic DHS (MDHS) and the non-adaptive version of the ACMDHS named CMDHS, which both have slight differences in the way they navigate through the search space using different set of parameters.

2 Related Work

Genetic Algorithm (GA) was first used as an efficient optimization method in different optimization problems. However, due to the drawbacks of GA, many other optimization methods have been devised in the last two decades. For example, the swarm intelligence methods were deployed to enhance the traditional document clustering. Two swarm methods have been adopted as document clustering algorithms which are the Artificial Bee Colony (ABC) [10] and the Bee Colony Optimization (BCO) [11]. The ABC is recently combined with two local search methods used for document clustering. The two local search methods are the Gradient and the Chaotic methods, which are used to enhance the performance of local searches [12]. The BCO is another swarm-based method which is modified in [13] by introducing two new heuristic operators which are *fairness* and *cloning*. This method is also combined with a local search method. According to their proposal, the BCO is combined with k-means, and the k-means is used to perform the local search instead of clustering. In both ABC and BCO methods, it can be noted that using the global search is insufficient in the exploitative aspect to find the best solution for document clustering.

Besides the above mentioned examples, recently the Differential Evolution (DE) has been successfully combined with the HS. The resulted method is named Differential HS (DHS) which is a combination of HS and DE. DHS has been empirically experimented and it outperformed both DE and HS [14]. In the DHS, the DE mutation operator was used instead of the original pitch adjustment process of the HS [15].

Before DHS, HS was used as an efficient document clustering along with the k-means in [16]. The k-means was combined in three different positions, namely, interleaved, sequential, and one-step. The position of using the k-means differs from one to another. The tests revealed that the hybrid approaches outperformed the HS, GA, and k-means as the search algorithms. Furthermore, a Global-Best HS combined with the k-means was introduced in [17] to cluster Web documents. The hybrid approach outperformed both methods separately. The same combination described previously can also be seen in [18], but in this case the GA was integrated with k-means. The solutions produced from the GA-k-means were used to construct the HS memory (population). The authors claimed that with the combination of HS and GA performed better than using just the GA or the HS with the k-means separately.

Although HS has been successful, it suffers from slow convergence and stagnation problems [15]. DE can be considered another variant of the HS [19]. In contrast to the HS, DE has not been utilized extensively in document clustering. Owing to its promising results seen in many other combinatorial problems, DE was used in [20] to enhance the k-means clustering and to prevent the production of weak solutions. Later, the same combination was improved so that the K cluster number could be set automatically [21].

One of the attempts to integrate the prospects of both HS and DE was proposed in [15] where the tests showed that using the DE operators instead of the pitch adjustment step can outperform both the HS and the DE. However, the proposed method still needs an accurate tuning of the DE parameters as the DE is sensitive to the setting of its control parameters [8, 9] in the same way as many other global search methods. Therefore in [22], an adaptive method named Self-adaptive Differential Evolution algorithm with Improved Mutation Mode (IMMSADE) was proposed. In IMMSADE, the control parameters were dynamically modified on the basis of the diversity of the population. As a result, it can be concluded that the combination of the adaptive way of parameter tuning with the DHS can lead to further improvement to the DHS.

The majority of the above-mentioned optimization methods have a drawback in the local search. Therefore, the Memetic optimization was proposed. Moreover, due to the sensitivity of the Bandwidth (BW) parameter in the pitch adjustment of the Harmony Memory (HM) in the traditional HS, the DHS is used instead with the document clustering problem. In this way, the need to manually set the BW parameters can be eliminated. In the DHS, the use of the DE operators replaces the need to use the traditional Pitch Adjustment technique that uses the bandwidth operator. Moreover, the static settings of both DE crossover and mutation operators represented by the *Cr* and *F* parameters were also adaptively set. Due to the gaps noted earlier, the aim of this work was intended to overcome the limitations of both HS and DE using the DHS first as global search in MA. Later, the DHS was enhanced to modify its parameters adaptively.

4 Document Clustering using ACMDHS Optimization

The proposed method is based on the combination of the global search using DHS once and CDHS once again with the k-means local search. The main steps of the proposed methods are listed below:

1. The first step begins by transforming the text documents into a numeric format using the same techniques explained in [23]. These techniques involve the text tokenization, stop words removal, stemming and weighting keywords.
2. The harmony memory initialization, the initial population of the harmony memory contains a random assignment of documents to cluster centroids. Each row of the harmony memory is solution whereas the length of each solution is fixed to the number of documents. In this step the evaluation of each random solution will be calculated using the fitness function explained in step 5.
3. The harmony memory update using mutation, in the differential harmony search DHS method, the DE mutation used instead of the traditional pitch adjustment step in HM as illustrated in equation (1).

$$v_{j,i,G+1} = x_{j,r1,G} + F.(x_{j,r2,G} - x_{j,r3,G}), \qquad i=1,2,..,N;j=1,2,..,D \qquad (1)$$

   where, $v_{j,i,G+1}$ is the newly generated trial vector, $x_{j,r1,G+1}$, $x_{j,r2,G+1}$, and $x_{j,r3,G+1}$ are three randomly selected vectors, j is the solution inedex and i is the bit index with the jth solution. G is the generation number, and F is a scaling factor [0,1] (F is adaptively tuned. If the new solution has a higher fitness, the current F value will be retained. Otherwise; it will be discarded and substituted by a newly generated value).

4. The harmony memory update using crossover, this step is only applied with the CDHS. Unlike DHS, that only makes use of the mutation for the pitch adjustment, a modified DHS version used in this paper that uses DE crossover after applying the mutation. The new solution is created based on equation (2).

$$U_i^g = \begin{cases} v_i^g & rand(0,1) < Cr \\ x_i & otherwise \end{cases} \qquad (2)$$

   where U_i^g is the newly generated solution and x_i is the target solution from the HMS and v_i^g is the mutant vector obtained from equation (1) and Cr is the crossover probability.

5. The fitness function used to evaluate the viability of solutions is the Average Distance of Document to Centroid (ADDC). Equation (3) shows how the ADDC is calculated.

$$ADDC = \left[\sum_{i=1}^{k} \frac{1}{n_i} \sum_{j=1}^{m_i} D(c_i, d_j) \right] \Big/ k \qquad (3)$$

where k clusters and m represent the total number of documents, and the D is the cosine similarity measure calculated in equation (4), d is a particular document and c is a centroid calculated in equation 10.

$$s(d_1, d_2) = \frac{d_1.d_2}{\|d_1\|.\|d_2\|} = \frac{\sum_{i=0}^{n-1} d_{1i} d_{2i}}{\sqrt{\sum_{i=0}^{n-1} (d_{1i})^2} \times \sqrt{\sum_{i=0}^{n-1} (d_{2i})^2}} \qquad (4)$$

where d_1 and d_2 are a pair of documents and n is the entire number of documents in the corpus.

6. After improvising the best solutions in steps 4 and 5, the local search using k-means is applied. The k-means can adjust the current best solution by modifying the current centroids positions. A comparison between the solution resulted from the local search and its original version is then conducted. The solution that achieves the best ADDC value will replace the inferior.
7. Substitution: after every fitness evaluation process, the solutions resulted in the current generation are compared with those of the older generation. If their fitness is better, they will replace them.

In Table 1, the datasets used in this research are explained in terms of the documents number and the number of classes in each of the datasets mentioned above is reported.

Table 1. Datasets

Dataset	D#	#Classes	Instances	Features
6 event crimes	D1	6	223	3864
10 types crime	D2	10	2422	15601
Reuters	D3	10	2277	13310
Pair 20news Groups	D4	2	1071	9497
20 news Groups	D5	20	1489	6738

5 Initial Parameters Setting and Test Results

In this section, an explanation of the parameters used and their setting is give. In addition, the test results of the proposed method and other methods are also provided.

5.1 Parameter Settings

Although an adaptive mechanism of parameter tuning is being followed in this research, still an initial tuning of the DE parameters *F* and *Cr* should be performed in prior. The tuning is based on three values for each single parameter. For the *Cr*, the three selected values were 0.2, 0.5 and 0.9 whereas for the *F* the values tested were 0.8, 0.1 and 0.5. These numbers are selected randomly. These two parameter values should be selected between zero and one. Therefore, the tested value for each parameter can show at which value the performance increases with all used datasets. Even though the performance might not be similar in all cases, a studied selection of these parameters will be preferable than a random one. For instance, if the parameter generated the highest performance with the majority of datasets, it will be used for the later tests. The parameter tests were conducted using these values with all datasets. After the selection of the best value of the *Cr* parameter, the *F* is then tested. As shown in Table 2, the value that helped to obtain the highest F-measure and the lowest ADDC was 0.5. Therefore, the other two values were discarded. In Table 2, it becomes evident that the use of Cr=0.5 has four out of five highest F-measure scores. When it comes to ADDC measure, the values of all the runs are almost consistent with only minimal differences. Thus, using the F-measure values to determine the best parameter value seems to be more appropriate in that case.

When it comes to the HS parameters, the Harmony Memory Size (HMS) is set to twice the number of centroids. According to [16], it was discovered that this is the best value of this parameter. The PAR and the HMCR were set based on [15], and their values were PAR =0.9 while HMCR=0.99. It is important to note that other candidate values of the Cr and F can be used, but at this stage using three experimental values seem to be sufficient.

5.2 Test Results

In this section, the comparison results of the proposed method with other methods are given. The methods used for the comparison are the Harmony Search (HS) method [16], Differential Harmony Search (DHS) [15], the Memetic HS [24] and the k-means. Two state-of-the-art methods were also used which are CGABC [25] and DEMC [26]. Besides two variants of the ACMDHS were also tested which are the MDHS and the CMDHS.

5.2.1 The Internal and External Evaluation

In order to evaluate the performance of each one of the competent algorithms, the internal and external evaluation measures are used. The ADDC internal measure serves as a fitness function with the aim of minimizing the distance among documents within one cluster. As mentioned before, it is used as a clustering compactness measure. The smaller the ADDC, the more compact the clusters are. The F-measure is utilized to express an expert view of the resulted clusters. In other words, this measure uses a truth table of the original representation of documents and it compares it with the newly resulted clusters. Thus, the F-measure needs the original class labels to be used in evaluations. The ADDC values needs to be minimized while the F-measure values need to be maximized. Both measures have their values listed in Table 3 and Table 4. Table 3 depicts the external measure values using F-measure while Table 4 shows the internal measure values using ADDC.

It is apparent from Table 3 that the performance of HS method was relatively better than the performance of DHS with all datasets while in Table 4 their ADDC values were almost the same. In that case the ADDC values will be ignored and depending only on F-measure values. The performance of the modified version of harmony search was supposed to be better than the native version due to the elimination of the need to use the Bandwidth parameter[15]. This comparison indicates that the modification of DHS was insufficient to empower it to outperform its ancestor HS. Therefore, two Memetic versions were based on the HS and the differential HS were tested; (i) Memetic HS (MHS) and (ii) Memetic Differential HS (MDHS). For the MHS, the performance is still similar to native HS whereas an improvement can be observed after using the MDHS in comparison to the DHS. The single most striking observation to emerge from the comparison of both DHS and MDHS was the local optima problem could have been solved using the local search.

However, the MDHS method performance has a different attitude when it is compared to other two state-of-the-art-methods which are the chaotic gradient-based artificial bee colony method CGABC and the differential evolution memetic clustering

DEMC. The results reveal that MDHS underperformed the CGABC method with all. When it comes to the DEMC, another scenario can be observed. The performance of MDHS was better than DEMC with almost all datasets. Based on that comparison, the use of DHS as a global search is better than the differential evolution in DEMC. On the other hand, it is highly suspected that the use of more than one local search in CGABC has a positive effect on the performance. However, using more than one local search will add more complexity to the DHS. Thus, modifying the global search further could be worthy for another test. Thus, the Crossover-added version of DHS was tested in two modes one in an adaptively-based parameter tuned mode and another is based on statically-based parameter tuned mode after using the best of them in Table 2.

As Table 3 shows, the F-measure values of both the crossover memetic MDHS (CMDHS) and the Adaptive CMDHS (ACMDHS) outperformed all of the two states-of-the-art methods and other versions of the harmony search. What is interesting in this comparison is that the enhancement of the global search has a remarkable effect on the performance even when it is compared with another Memetic scheme that incorporates two local search methods as is seen with CGABC. Now, by restricting the comparison between CMDHS and ACMDHS, it becomes clear that the statically-based parameter tuned version (CMDHS) outperformed the dynamically-based one (ACMDHS). The single most striking observation to emerge from that comparison is the tuned parameters i.e. the *F* for the mutation and the Cr for the crossover have only minor effect on the performance of the centroids allocation.

Table 2. Cr and F parameter Tuning Table

	Cr=0.2		**Cr=0.5**		**Cr=0.9**	
Dataset	***F-Measure***	***ADDC***	***F-Measure***	***ADDC***	***F-Measure***	***ADDC***
	83.909	0.71	**85.811**	**0.71**	79.866	0.71
D1	45	8938	**85**	**4983**	677	4957
	94.540	0.74	**98.429**	**0.74**	96.452	0.74
D2	028	8749	**393**	**538**	157	342
	88.052	0.71	89.205	0.71	**92.794**	**0.70**
D3	35	1795	819	1147	**232**	**7603**
	94.953	0.84	**96.090**	**0.84**	94.857	0.83
D4	637	2767	**954**	**3154**	401	9821
	66.480	0.84	**99.547**	**0.84**	66.666	0.84
D5	562	7179	**975**	**6073**	667	8331
	F=0.1		**F=0.5**		**F=0.8**	
Dataset	***F-Measure***	***ADDC***	***F-Measure***	***ADDC***	***F-Measure***	***ADDC***
	89.732	**0.72**	77.772	0.71	85.811	0.71
D1	**655**	**1584**	432	9301	85	4983
	97.834	0.74	96.460	0.74	**98.429**	**0.74**
D2	008	9086	589	7813	**393**	**538**
	92.757	**0.71**	90.263	0.70	89.205	0.71
D3	**559**	**1581**	957	8986	819	1147
	98.230	**0.84**	90.055	0.82	94.953	0.84
D4	**616**	**2119**	755	3152	637	2767
	99.646	**0.84**	99.176	0.84	99.547	0.84
D5	**309**	**6778**	776	6605	975	6073

Table 3. F-measure Values

Runs	HS	DHS	MHS	MDHS	KM	CGABC	DEMC	CMDHS	ACMDHS
D1	76.98	50.06	78.62	63.84	63.62	82.65	86.00	**88.50**	85.97
D2	90.50	80.85	90.75	80.31	35.52	94.28	94.00	**98.69**	96.03
D3	91.94	83.84	88.01	88.75	16.37	91.95	80.30	96.94	**97.56**
D4	96.72	90.76	95.17	89.53	67.72	95.48	88.75	97.56	**98.84**
D5	98.16	93.86	97.23	95.69	0.60	98.96	53.88	**99.91**	98.92

Table 4. ADDC Values

Runs	HS	DHS	MHS	MDHS	CGABC	DEMC	CMDHS	ACMDHS
D1	0.72	0.71	0.71	0.72	0.72	0.72	0.72	0.72
D2	0.71	0.71	0.68	0.71	0.86	0.86	0.74	0.71
D3	0.84	0.82	0.82	0.82	0.72	0.67	0.73	0.71
D4	0.74	0.74	0.74	0.74	0.83	0.82	0.82	0.75
D5	0.73	0.73	0.73	0.73	0.84	0.83	0.72	0.73

Finally, in Table 4 the ADDC values of the K-means were not listed because the K-means is not an optimization-based method. However, in the ACMDHS the k-means is used as a local search. Through Table 4, it can be noted that the general trends in all results are compatible. The stability of the ADDC in comparison to the F-measure did not mean all methods performing equally. That is because the F-measure values were changing when ADDC values were almost steady.

6 Conclusions

In this paper a Memetic-based clustering method named Adaptive Crossover Memetic Differential Harmony Search (ACMDHS) is proposed. Other two variants of ACMDHS resulted in this paper are the Memetic DHS (MDHS) and the non-adaptive version of the ACMDHS named CMDHS, which both have slight differences in the way they navigate through the search space using different set of parameters. The experimental results showed that the proposed ACMDHS provided the best F-measure results in comparison to other document clustering methods: Harmony Search (HS) method, Differential Harmony Search (DHS) [15], the Memetic HS, Memetic DHS and the k-means. Moreover, it was also compared to its other two variants: MDHS and CMDHS. The test results showed that the CMDHS performed the same or slightly better than the Adaptive CMDHS indicating to the minor effect of the differential evolution parameters on the performance. Finally among the other two state-of-the-art methods which are CGABC and DEMC, the ACMDHS has achieved the highest F-measure.

Acknowledgements

Ibraheem would like to express his gratitude to the Higher Committee of Education Development in Iraq (HECD) for the scholarship he has received to fund his PhD study.

References

1. Saiyad, N.Y., H.B. Prajapati, and V.K. Dabhi.: A survey of document clustering using semantic approach. in Electrical, Electronics, and Optimization Techniques (ICEEOT), International Conference on. IEEE (2016).
2. Feng, A.: Document clustering: an optimization problem. in Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. ACM (2007).
3. Hruschka, E.R., et al..: A survey of evolutionary algorithms for clustering. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, **39**(2): p. 133-155(2009).
4. Yang, Y., M. Kamel, and F. Jin.: A model of document clustering using ant colony algorithm and validity index. in Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on IEEE (2005).
5. Smith, J.E.: Coevolving memetic algorithms: a review and progress report. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, **37**(1): p. 6-17 (2007).
6. Nguyen, Q.H., Y.-S. Ong, and M.H. Lim.: A probabilistic memetic framework. Evolutionary Computation, IEEE Transactions on, **13**(3): p. 604-623 (2009).
7. Neri, F. and E. Mininno.: Memetic compact differential evolution for cartesian robot control. Computational Intelligence Magazine, IEEE, 2010. **5**(2): p. 54-65.
8. Reynoso-Meza, G., et al.: Hybrid DE algorithm with adaptive crossover operator for solving real-world numerical optimization problems. in Evolutionary Computation (CEC), Congress on IEEE (2011).
9. Zhang, J. and A.C. Sanderson.: *JADE:* adaptive differential evolution with optional external archive. IEEE Transactions on evolutionary computation,. **13**(5): p. 945-958 (2009).
10. Karaboga, D. and C. Ozturk.: A novel clustering approach: Artificial Bee Colony (ABC) algorithm. Applied soft computing, **11**(1): p. 652-657 (2011).
11. Lučić, P. and D. Teodorović.: Computing with bees: attacking complex transportation engineering problems. International Journal on Artificial Intelligence Tools, **12**(03): p. 375-394(2003).
12. Bharti, K.K. and P.K. Singh.: Chaotic gradient artificial bee colony for text clustering, Soft Computing, **20**(3): p. 1113-1126 (2016).
13. Forsati, R., A. Keikha, and M. Shamsfard.: An improved bee colony optimization algorithm with an application to document clustering. Neurocomputing, **159**: p. 9-26(2015).
14. Chakraborty, P., et al.: An improved harmony search algorithm with differential mutation operator. Fundamenta Informaticae, **95**(4): p. 401-426 (2009).
15. Abedinpourshotorban, H., et al.: A differential-based harmony search algorithm for the optimization of continuous problems. Expert Systems with Applications, **62**: p. 317-332(2016).
16. Forsati, R., et al.: Efficient stochastic algorithms for document clustering, Information Sciences, **220**: p. 269-291 (2013).
17. Cobos, C., et al.: Web document clustering based on global-best harmony search, K-means, frequent term sets and Bayesian information criterion. in Evolutionary Computation (CEC), 2010 IEEE Congress on IEEE (2010).
18. Rafi, M., et al.: Towards A Soft Computing Approach to Document Clustering, in Proceedings of the 2017 International Conference on Machine Learning and Soft Computing, ACM: Ho Chi Minh City, Vietnam. p. 74-81(2017).
19. Vakil-Baghmisheh, M.-T. and M.A. Ahandani.: A differential memetic algorithm. Artificial Intelligence Review, **41**(1): p. 129-146 (2014).
20. Abraham, A., S. Das, and A. Konar. Document clustering using differential evolution. in Evolutionary Computation, CEC Congress on IEEE (2006).

21. Das, S., A. Abraham, and A. Konar.: Automatic clustering using an improved differential evolution algorithm. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, **38**(1): p. 218-237 (2008).
22. Wang, S., Y. Li, and H. Yang.: Self-adaptive differential evolution algorithm with improved mutation mode. Applied Intelligence, p. 1-15(2017).
23. Uysal, A.K. and S. Gunal.: The impact of preprocessing on text classification. Information Processing & Management, **50**(1): p. 104-112(2014).
24. Gao, X.Z., X. Wang, and K. Zenger.: A memetic-inspired harmony search method in optimal wind generator design. International Journal of Machine Learning and Cybernetics, **6**(1): p. 43-58(2015).
25. Bharti, K.K. and P.K. Singh.: Opposition chaotic fitness mutation based adaptive inertia weight BPSO for feature selection in text clustering. Applied Soft Computing,. **43**: p. 20-34(2016).
26. Al-Jadir, I., et al.: Differential Evolution Memetic Document Clustering Using Chaotic Logistic Local Search, in Neural Information Processing: 24th International Conference, ICONIP. Guangzhou, China, November 14-18, 2017, Proceedings, Part I, D. Liu, et al., Editors. Springer International Publishing: Cham. p. 213-221 (2017).