8th MUCET 2014, Date: 10-11 November 2014, Melaka, Malaysia

K-Harmonic Means Data Clustering with Firefly Search Approach

Ong Pauline
Faculty of Mechanical and Manufacturing Engineering
Universiti Tun Hussein Onn Malaysia (UTHM)
86400 Parit Raja, Batu Pahat, Malaysia
ongp@uthm.edu.my

Zarita Zainuddin
School of Mathematical Sciences
Universiti Sains Malaysia (USM)
11800 USM, Penang, Malaysia
zarita@cs.usm.my

***Abstract*—K-harmonic means (KHM), has been introduced as one of the vital solutions to the classical K-means, which alleviates the sensitive dependence on the initial clusters conditions for the latter. However, it does retain the same deficiency as K-means: execution of KHM has a propensity to converge to local optima easily. In response to circumventing this problem, a new variant of KHM based on a recent nature inspired firefly search approach – specifically, the Firefly K-harmonic means (FAKHM) algorithm – is resorted to. Assessment analysis on several artificial and real life datasets demonstrates the superiority of the proposed FAKHM algorithm.**

Keywords—clustering; firefly algorithm; K-harmonic means; K-means; unsupervised learning

I. Introduction

Attempting to gain insight from a large scale of dataset in a succinct and human-comprehendible manner can be posed as data mining problem, where data clustering is one of the facets of this interesting field which has roots in variety of domains; including bioinformatics, pattern recognition and artificial intelligence [1, 2]. Clustering aims to discover and describe the natural groupings of observations being located in the input space. Observations are partitioned into several clusters, depending to a predefined similarity criterion, in which those observations belong to the same cluster, will have higher similarity than the observations in other clusters. Generally speaking, the goal of a clustering is to assign a set N of n observations in d input space to a set K of k points, denoted as cluster centers (centroids), based on optimizing a performance criterion. A typical used performance criterion is the total within-cluster variance [3].

The fascinating features of simplicity, easy interpretation and computationally easy-to-use of the K-means (KM) algorithm, has vulgarized its widespread implementation in data clustering [4]. Nonetheless, this cluster-seeking approach was reported to be dependent substantially on the initial designation of centroids and trapped in local optima easily [5]. In this regard, efforts to resolve this limitation fruit in an improved variant of KM, specifically, the K-harmonic means (KHM) algorithm [6]. KHM remedies the sensitive dependence of the initial clusters representation as in KM, which attributed to the fact that KHM minimizes the objective function based on the weighted harmonic average from all observations in the input space to all centroids [7]. Although such significant improvement based on KHM is certainly noteworthy, it is still exposed to the problem of converging to local optima easily [4]. As such, a wide range of attempts are directed to solve this concern, especially by integrating the metaheuristic approaches [3-5, 7-10].

Simulated annealing was proposed by Güngör and Ünler as a means to solve local optima constraint in KHM [3]. Apart from that, Güngör and Ünler have extended their work by merging the Tabu search algorithm which possesses the neighborhood search capability with KHM [8]. Both of the resulted approaches with metaheuristic algorithms were shown to be just as effective or even outperformed the KM and KHM algorithms. Alguwaizani et al. enhanced the local search of KHM by embedding a variable neighborhood search algorithm [9]. Comparison with the results obtained by Güngör and Ünler [3, 8] indicated the superiority of their proposed method. A more recent hybrid KHM with gravitational search algorithm (GSA) approach was developed by Yin et al. [4]. By borrowing the strength from GSA, in addition to overcoming the local optima limitation in KHM, the hybrid GSAKHM showed relatively high speed of convergence than the GSA as well. Particle swarm optimization (PSO) was integrated by Yang et al. in KHM in order to escape from local minima [5]. The efficiency of the proposed PSOKHM was corroborated through empirical approaches, with simulated as well as real-world datasets. Jiang et al. used the ant clustering algorithm (ACA) to solve the local optimal problem in KHM [7]. The proposed ACAKHM algorithm gave encouraging results in terms of clustering effectiveness, as compared with the ACA and KHM. Another bio-inspired swarm computing – cat swarm optimization (CSO) algorithm, was coupled with KHM by Liu and Shen [10]. Improvement in clustering performance was observed, when combination of CSO and KHM was adopted. The newly developed cuckoo search was combined with KHM by Song et al. [11] .

A recent swarm intelligence-based technique, specifically, the firefly algorithm (FA), which investigates and exploits the foraging behavior of fireflies based on their flashing characteristic, was constructed by Yang to solve the multiobjective optimization problem [12]. In his preliminary studies, the beneficial potential of FA in finding the global optima for various classical benchmark functions was validated, and it was superior to both PSO and genetic algorithm [12, 13]. For clustering purposes, it has been presented that in most cases, the global search ability of FA surpassed the other techniques; including artificial bee colony, PSO, Bayes net, multilayer perceptrons, radial basis function neural networks, and Naïve Bayes Tree [14]. On the other hand, the implementation of FA in various domains shows promising results [15-17].

In this present study, the applicability of FA with global search ability for solving the local optima problem in KHM is explored. A hybrid clustering algorithm, namely, the firefly K-harmonic means (FAKHM) algorithm, which takes merits from both FA and KHM into account, is proposed. The validity and superiority of the FAKHM are assessed in clustering several simulated and real life datasets. The remainder of this paper is organized as follows. An introduction of KHM is given in Section 2, followed by a brief discussion on FA in Section 3. Section 4 presents the proposed FAKHM algorithm in detail. In Section 5, the experimental simulations of the proposed FAKHM, as compared with the FA and KHM algorithms are provided, and finally, conclusions are drawn in Section 6.

II. K-Harmonic Means Clustering

On account of its simplicity and computationally efficient, KM is one of the early invented data clustering methods, which has been studied for decades. By applying the KM, a set of n observations x_j, $j = 1,\ldots,n$ are partitioned into k groups C_i, $i = 1,\ldots,k$, where the observations are assigned to the cluster the center of which is the nearest. In other words, the cluster-seeking of KM is based on minimization of the sum of squares of the distance between each observation, x_j, to its nearest cluster center, c_i. The progressive computation of KM begins with random initialization a set of k cluster centers. However, if the initial cluster centers are created in a densely packed manner, a centroid might have difficulty moving out from the locally dense area, which eventually affects the clustering solutions drastically [3, 5]. In place of allocating the observation to the cluster whose center has the shortest distance to the observation, the KHM assigns the observations to the clusters by minimizing the weighted harmonic means of the distance from each observation to all cluster centers. As such, if multiple centroids are seeded within the same dense area, the KHM will shift one or more centroids to the region of observations with no close centroids reside in [8].

Before presenting the KHM algorithm in details, the nomenclature for the formulation of KHM is defined as follows:

$X = \{x_1, \ldots, x_n\}$: A set of n observations to be clustered.

$C = \{c_1, \ldots, c_k\}$: A set of k cluster centers.

KHM(X,C): The objective function of the KHM algorithm to be minimized.

$m(c_i|x_j)$: The membership function which defines the degree of belongingness of each observation x_j to the cluster center c_i.

$w(x_j)$: The weight function which defines the weighting influence of observation x_j in calculating the new position for cluster center c_i in the next iteration.

Hence, presenting a set of observations x_j, KHM clustering algorithm determines the cluster centers c_i iteratively using the following steps:

1. Initialize the cluster centers c_i, $i = 1,\ldots, k$, by selecting k observations randomly from all the available observations.
2. Determine the objective function, according to:

$$KHM(X,C) = \sum_{j=1}^{n} \frac{k}{\sum_{i=1}^{k} \| x_j - c_i \|^{-p}}, \tag{1}$$

where parameter p is associated with distance calculation and typically p is chosen as equal to or greater than 2.

3. Calculate the membership function for each observation x_j with respect to each cluster center c_i, according to:

$$m(c_i \mid x_j) = \frac{\| x_j - c_i \|^{-p-2}}{\sum_{i=1}^{k} \| x_j - c_i \|^{-p-2}}, \tag{2}$$

where $m(c_i \mid x_j) \in [0,1]$. This allows an inherent built-in fuzziness in KHM.

4. Calculate the weight function for each observation x_j, according to:

$$w(x_j) = \frac{\sum_{i=1}^{k} \| x_j - c_i \|^{-p-2}}{(\sum_{i=1}^{k} \| x_j - c_i \|^{-p})^2}. \tag{3}$$

5. Update the new position for each cluster center c_i according to the membership and weight functions from all observations obtained in Step 3 and 4 by using:

$$c_i = \frac{\sum_{j=1}^{n} m(c_i \mid x_j) w(x_j) x_j}{\sum_{j=1}^{n} m(c_i \mid x_j) w(x_j)}. \tag{4}$$

6. If the improvement of the objective function over previous iteration falls below a certain threshold or the iterations reach a predefined maximum value, then stop. Otherwise, go to Step 2.
7. Assign the observation x_j to the k-th cluster with the highest value of $m(c_i|x_j)$.

It is worth mentioning that the weighting influence for each observation is updated dynamically based on a harmonic means. As observed from Equation (3), due to the reciprocal of the summation of the reciprocals of the distance between observation x_j and cluster center c_i, the observation which is far away from any cluster center will possess large weight. Meanwhile, the observation that is near to the cluster centers will be assigned with a small weight value. This is essentially important as by putting a larger weight on the observations that are not near to any of the cluster center, this action will attract those cluster centers away from the region with high local density of multiple cluster centers and thus, KHM is insensitive to the initialization of cluster centers than KM [8].

III. Firefly Algorithm

Fireflies, also recognized as glow worms, emit short and rhythmic flashes. By flashing the light off and on, through a process known as bioluminescence, the glow can be used as a decoy to lure its prey, to attract mates, and also to ward off its predators. FA, a nature-inspired metaheuristic optimization technique which finds the global optima of objective function based on this foraging behavior, was formulated by Yang [12]. In essence, the proposed FA is idealized with respect to the following assumptions: (i) All fireflies are gender-blindness. It means that one firefly can mate with other fireflies, regardless of their sex; (ii) Attractiveness in the eyes of the other is proportional to their brightness, in such a way that the firefly will be attracted and move toward its adjacent firefly which flashes brighter than the rest. If no such member from the swarm exists, it will move randomly; and (iii) The brightness of a firefly is considered as equivalent to the value of the objective function for a maximization/minimization problem [18]. In the implementation of FA, there are two important issues to be concerned with.

i. *Determination of the light intensity*: Biologically, each firefly lights up proportionally to the amount of light-emitting compound, called luciferin, found in its abdomen. In FA, in order to quantify this value relatively, the produced light intensity is associated with the encoded objective value [12]. Hence, if x_i is the solution for a firefly i and $f(x_i)$ represents its fitness value, the brightness I for the firefly i is denoted as:

$$I_i = f(x_i), \tag{5}$$

where $1 \le i \le n$ and n is number of fireflies.

ii. *Movement of the firefly*: As mentioned earlier, a firefly i at location x_i with lower light intensity will move toward an adjacent firefly j at location x_j, if the latter glows brighter than it. In FA, this movement is characterized as [12]:

$$x_i(t+1) = x_i(t) + \beta_0 \exp(-\gamma r^2)(x_j - x_i) + \alpha(rand - 1/2) \tag{6}$$

where r is the Euclidean distance between fireflies i and j.

The second term in Equation (6) characterizes the movement of firefly. As the light is absorbed in the medium, the light absorption coefficient γ is imposed (γ = 1 usually). β_0 represents the initial attractiveness at r = 0, and it is chosen as 1 typically. The last term randomizes the movement direction of firefly i, with $\alpha \in [0,1]$ denotes the randomization parameter, and *rand* generates random number uniformly in the interval of $\alpha \in [0,1]$.

IV. The Hybrid Firefly K-Harmonic Means Clustering Algorithm

KHM is favored, in the sense that fewer function evaluations are needed for convergence and thus, its short runtime is an added benefit in addition to its insensitiveness to the initialization of the cluster centers. Nonetheless, the issue of getting stuck easily in infeasible local optima has remained unresolved for KHM. On the other hand, the strong global searching ability of FA has been investigated and corroborated by Yang [18]; moreover, its computation time may be comparable to that of KHM, if the number of maximum iteration are not assigned to a large value.

In this present study, a hybrid clustering algorithm, specifically, the FAKHM algorithm is formed by combining the merits of both KHM and FA. By borrowing the strength of FA, this hybrid approach can overcome the local optima problem and reach the global optima in limited iterations. Meanwhile, a more appropriate position of the initial cluster centers can be determined from FA for KHM, from which it generates a better input to FA in return, so as to accelerate its convergence to the global optima.

The algorithm for the implementation of FAKHM is given as follows [12, 14]:

Step 1: Set the initial parameters; including the maximum iteration count *Itercount*, number of fireflies n, light absorption coefficient γ, and initial attractiveness β_0.

Step 2: Generate the initial population of n fireflies within d-dimensional search place randomly x_i, $i = 1,\ldots,n$ to carry out their forage activity.

Step 3: Set iterative *countGen_1* = 0.

Step 4: Set iterative *countGen_2* = 0 and *countGen_3* = 0.

Step 5: FA Approach

- Step 5.1: Determine light intensity I_i at x_i which is directly proportional to $f(x_i)$.
- Step 5.2: Compare the light intensity for each firefly. If the light intensity for firefly i is greater than firefly j, move firefly i to firefly j in d-dimension, by using Equation (6).
- Step 5.3: Evaluate the new solutions and update the light intensity I_i at new position x_i.
- Step 5.4: Update *countGen_2* = *countGen_2* + 1.

  If *countGen_2* < 8, go to Step 5.2. Otherwise, go to Step 6.

Step 6: KHM Approach

For each observation i,

- Step 6.1: Take the result from FA as the initial cluster centers for KHM algorithm.
- Step 6.2: Calculate the objective function, membership and weight function, by using Equations (1) - (3).
- Step 6.3: Update the cluster center c_i according to Equation (4).

Step 6.4: Update *countGen_3* = *countGen_3* + 1.

If *countGen_3* < 4, go to Step 6.2. Otherwise, go to Step 7.

Step 7: Update *countGen_1* = *countGen_1* + 1.

If *countGen_1* < *Itercount*, go to Step 4. Otherwise, assign the observation to the cluster i with the largest $m(c_i|x_j)$

It is of interest to note that for each loop of FAKHM, four iterations of KHM are applied to the fireflies obtained after every eight cycles, in order to improve its fitness value. These values are chosen based on the previous studies [4, 5]. The value of p in Equations (1)-(3) is chosen empirically as 2.

V. EXPERIMENTAL SIMULATIONS

To validate the feasibility and validity of the proposed hybrid FAKHM algorithm as compared to the FA and KHM algorithms, the experimental studies of these algorithms in two artificial datasets are presented. In addition to the simulated datasets, the publicly available real life datasets (breast cancer, iris and wine datasets) from the UCI machine learning repository (http://archive.ics.uci.edu/ml/datasets.html) are taken into consideration as well.

A. Datasets

The description of the datasets used in this present study is given as follows:

i. Artset1: As illustrated in Figure 1, Artset1 is a two-dimensional artificial dataset with three non-overlapping clusters, where there are 100 observations in each cluster. The observations are independently derived from bivariate normal distribution with means (-2,-2), (2,2), (6,6) and covariance matrix $\sum = \begin{bmatrix} 0.4 & 0.04 \\ 0.04 & 0.4 \end{bmatrix}$.

ii. Artset2: As portrayed in Figure 2, Artset2 is a three-dimensional simulated dataset with a total of 300 observations. The observations are generated from the uniform distribution on the set of (10,25), (25,40), and (40,55) for each cluster, respectively, with 100 observations in each cluster.

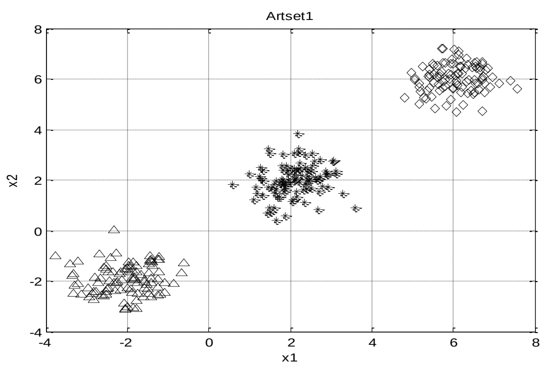

Fig. 1. Data distribution of dataset Artset1, which consists of three clusters and 100 observations in each.

TABLE I. CHARACTERISTICS OF THE DATASETS

Dataset	No. of Classes	No. of Atrributes	Size of dataset (no. of instances in each class is given in parentheses)
Artset1	3	2	300 (100,100,100)
Artset2	3	3	300 (100,100,100)
Breast Cancer	2	9	683 (239, 444)
Iris	3	4	150 (50, 50, 50)
Wine	3	13	178 (59, 71, 48)

iii. Breast Cancer: The Wisconsin breast cancer dataset has 683 instances with 9 attributes corresponding to the clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli and mitoses, and involves a binary classification problem. Partitioning aims to predict whether the given instance is benign (239 instances) or malignant (444 instances), which are linearly separable.

iv. Iris: There are three classes exist in the Iris dataset, where each of them represents a type of Iris plant: Iris Setosa, Iris Versicolour and Iris Virginica. Each class contains 50 instances with 4 attributes refer to the sepal length, sepal width, petal length and petal width.

v. Wine: The chemical analysis results for wines from three different cultivars are summarized in this dataset. The 13 attributes are referring to the amount of different constituents (Alcohol, Malic acid, ash, alkalinity of ash, Magnesium, total phenols, flavanoids, non-flavanoid phenols, Proanthocyanins, colour intensity, hue, OD280/OD315 of diluted wines, Proline) found in the 178 instances (59, 71 and 48 instances for each cultivar, respectively).

The characteristics of the datasets under investigation are summarized in Table 1.

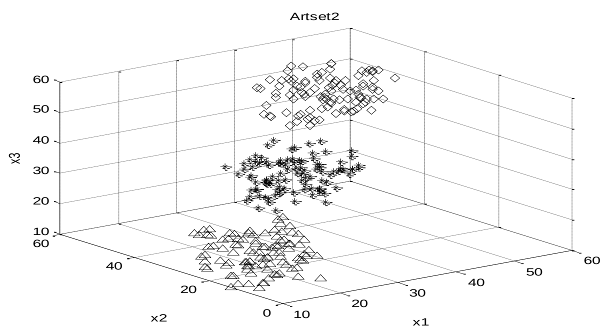

Fig. 2. Data distribution of dataset Artset2, which consists of three clusters and 100 observations in each.

TABLE II. SIMULATION RESULTS OF KHM, FA AND FAKHM IN CLUSTERING THE ARTIFICIAL AND REAL LIFE DATASETS

Dataset	Performance Measure	Clustering Algorithm		
		FA	KHM	FAKHM
Artset1	Sum_{KHM}	**598.492 (0.000)**	600.265 (0.000)	**598.492 (0.000)**
	F-Measure	**1.000 (0.000)**	**1.000 (0.000)**	**1.000 (0.000)**
	Runtime	**0.213 (0.029)**	8.187 (0.155)	1.827 (0.028)
Artset2	Sum_{KHM}	**4.826E+4 (0.000)**	3.611E+5 (1.756E+3)	**4.826E+4 (0.000)**
	F-Measure	**1.000 (0.000)**	0.684 (0.005)	**1.000 (0.000)**
	Runtime	**0.455 (0.035)**	0.638 (0.018)	4.004 (0.310)
Breast Cancer	Sum_{KHM}	2.983E+4 (0.000)	1.188E+5 (8.483E+3)	**4.654E+3 (656.422)**
	F-Measure	0.951 (0.000)	0.902 (0.015)	**0.962 (0.001)**
	Runtime	**0.675 (0.051)**	6.225 (0.304)	19.645 (0.878)
Iris	Sum_{KHM}	**181.728 (0.000)**	1375.842 (1.798)	185.667 (0.000)
	F-Measure	0.892 (0.000)	0.864 (0.031)	**0.933 (0.000)**
	Runtime	**0.262 (0.052)**	0.346 (0.010)	4.607 (0.030)
Wine	Sum_{KHM}	5.388E+6 (0.000)	1.658E+7 (6.831E+6)	**1.085E+5 (1.061E+5)**
	F-Measure	0.686 (0.000)	0.613 (0.045)	**0.718 (0.004)**
	Runtime	**0.870 (0.107)**	1.010 (0.197)	5.610 (1.880)

a. The results shown in this table are the means and standard deviation (in parenthesis) for 10 independent runs. Bold values are the best performances achieved by the algorithms.

B. Performance Evaluation

For quantitative evaluation purpose, two performance criteria are utilized in order to quantify the superiority of a clustering solution, which are:

i. Sum_{KHM}: It represents the summation from all observations for the harmonic average of the Euclidean distance between an observation and all the cluster centers. The value for Sum_{KHM} can be calculated by using Equation (1) wherein, a better clustering solution will be indicated by a lower Sum_{KHM} score.

ii. F-measure: The F-measure is defined as [4, 5, 7]:

$$F = \sum_i \frac{n_i}{n} \max_j \{F(i, j)\}, \tag{7}$$

where

$$F(i, j) = \frac{(b^2 + 1) p(i, j) r(i, j)}{b^2 p(i, j) + r(i, j)}. \tag{8}$$

$p(i,j)$ and $r(i,j)$ are precision and recall for each class i (from reference solution) and cluster j (from clustering solution) which are given by:

$$p(i, j) = \frac{n_{ij}}{n_j}, \tag{9}$$

and

$$r(i, j) = \frac{n_{ij}}{n_i}, \tag{10}$$

where n_j, n_i and n_{ij} is the number of observations assigned to cluster j, number of observations in class i, and number of observations of class i within cluster j. The value of b is chosen as 1 so that an equal weighting for both precision and recall are obtained. A better clustering solution will be indicated by a higher F-measure, where the score 1 means that the perfect clustering is attained.

For each dataset, the experimental simulation of KHM, FA and FAKHM algorithms is computed for 10 times. The obtained results are averaged, where the means and standard deviation for each algorithm in addition to its total runtime are summarized in Table 2.

C. Results and Discussion

As evident in Table 2, the KHM, FA and FAKHM algorithms provide appropriate partitioning results for dataset Artset1, which is clearly indicated by the attained F-measure = 1 for all cases. This is not surprising since Artset1 is a low-dimensional simple dataset which possesses non-overlapping data distribution, as illustrated in Figure 1. In fact, these three algorithms are able to converge to the optimal cluster centers successfully, where the solutions of {(-2.0463, -1.9499), (2.1003, 2.0136), (6.0481, 6.0313)}, {(-2.1119, -1.9886), (2.0716, 2.0096), (6.0509, 6.0189)} and {(-2.0443, -1.9482), (2.1057, 2.0204), (6.0470, 6.0304)} are generated by KHM, FA and FAKHM, respectively, and hence, their obtained average Sum_{KHM} values are almost identical. It thus suggested that the three algorithms may perform equally when the dataset under investigation is well-separated; however, the KHM allows the fastest computation. For the Artset2, both of the KHM and FAKHM algorithms are comparable in terms of Sum_{KHM} and F-measure, but the former gives faster convergence to the global optima.

For the real life datasets, Table 2 shows that the utilization of FA with KHM could lead to a better partitioning result, as indicated by the highest F-measure obtained by the FAKHM for the concerned datasets. At the same time, the average values of Sum_{KHM} for FAKHM are practically perfectly or even superior to those achieved by KHM, while the performance of algorithm is presumable preferable for clustering a high-dimensional complex dataset, especially when the underlying interaction between the attributes is indistinct.

Performance assessment of the proposed FAKHM algorithm and other methods reported in literature applied to the same datasets are made in terms of F-measure, which is presented in Table 3. It can be deduced from this table that the FAKHM offers more promising solutions, as it surpassed the other metaheuristic algorithms, specifically, the ACA, ACAKHM, GSAKHM, PSO, and PSOKHM, by achieving the highest F-measure for all the datasets, from which the superiority and feasibility of the FAKHM can be validated.

TABLE III. PERFORMANCE COMPARISON OF FAKHM ALGORITHM WITH OTHER RESULTS IN LITERATURE

Dataset	Method	F-Measure	Reference
Artset1	ACA	0.35(0.2)	[7]
	ACAKHM	**1.000(0.000)**	[7]
	GSAKHM	**1.000(0.000)**	[4]
	PSO	**1.000(0.000)**	[5]
	PSOKHM	**1.000(0.000)**	[5]
	FAKHM	**1.000(0.000)**	
Artset2	ACA	0.39(0.37)	[7]
	ACAKHM	**1.000(0.000)**	[7]
	GSAKHM	**1.000(0.000)**	[4]
	PSO	0.681(0.093)	[5]
	PSOKHM	**1.000(0.000)**	[5]
	FAKHM	**1.000(0.000)**	
Breast Cancer	GSAKHM	**0.862(0.000)**	[4]
	PSO	0.820(0.046)	[5]
	PSOKHM	0.835(0.003)	[5]
	FAKHM	**0.962(0.001)**	
Iris	ACA	0.31(0.16)	[7]
	ACAKHM	**0.80(0.07)**	[7]
	GSAKHM	0.766(0.000)	[4]
	PSO	0.740(0.025)	[5]
	PSOKHM	0.765(0.004)	[5]
	FAKHM	**0.933(0.000)**	
Wine	ACA	0.21(0.2)	[7]
	ACAKHM	0.53(0.02)	[7]
	GSAKHM	**0.553(0.000)**	[4]
	PSO	0.530(0.039)	[5]
	PSOKHM	**0.553(0.000)**	[5]
	FAKHM	**0.718(0.004)**	

a. The results shown in this table are the means and standard deviation (in parenthesis). Bold values are the best performances achieved by other methods in literature and the FAKHM in each dataset.

VI. CONCLUSION

In this paper, a hybrid clustering algorithm based on KHM and a recent developed swarm intelligence technique, specifically, the firefly algorithm is proposed. The validity of the resulted FAKHM algorithm is assessed empirically, through simulation on artificial and real life datasets. The obtained results demonstrate the beneficial potential of FAKHM, as it provides adequate clustering results for the concerned datasets which possess different structures. Its superiority is especially noteworthy when dealing with the high-dimensional complex datasets wherein, it outperforms the other nature inspired approaches in most cases. Although the proposed FAKHM alleviates the local optima problem in KHM, the computation of FAKHM is onerous and thus, increasing the search efficiency of FAKHM will be an interesting issue to pursue in the future work.

REFERENCES

[1] Z. Zainuddin, and P. Ong, "Design of wavelet neural networks based on symmetry fuzzy C-means for function approximation", Neural Comput & Applic, vol. 2013, pp. 1-13.

[2] L. Zhang, L. Zhang, W. Pedrycz, W. Lu, and X. Liu, "An interval weighed fuzzy c-means clustering by genetically guided alternating optimization", Expert Systems with Applications, vol. 41, 2014, pp. 5960-5971.

[3] Z. Gungor, and A. Unler, "K-harmonic means data clustering with simulated annealing heuristic", Appl Math Comput, vol. 184, 2007, pp. 199-209.

[4] M.H. Yin, Y.M. Hu, F.Q. Yang, X.T. Li, and W.X. Gu, "A novel hybrid K-harmonic means and gravitational search algorithm approach for clustering", Expert Systems with Applications, vol. 38, 2011, pp. 9319-9324.

[5] F.Q. Yang, T.E.L. Sun, and C.H. Zhang, "An efficient hybrid data clustering method based on K-harmonic means and Particle Swarm Optimization", Expert Systems with Applications, vol. 36, 2009, pp. 9847-9852.

[6] B. Zhang, M. Hsu, and U. Dayal, "K-Harmonic means - A data clustering algorithm", HP Laboratories Technical Report, vol. 1999.

[7] H. Jiang, S.H. Yi, J. Li, F.Q. Yang, and X. Hu, "Ant clustering algorithm with K-harmonic means clustering", Expert Systems with Applications, vol. 37, 2010, pp. 8679-8684.

[8] Z. Gungor, and A. Unler, "K-harmonic means data clustering with Tabu-search method", Applied Mathematical Modelling, vol. 32, 2008, pp. 1115-1125.

[9] A. Alguwaizani, P. Hansen, N. Mladenovic, and E. Ngai, "Variable neighborhood search for harmonic means clustering", Applied Mathematical Modelling, vol. 35, 2011, pp. 2688-2694.

[10] Y.C. Liu, and Y. Shen, "Data clustering with cat swarm optimization", Journal of Convergence Information Technology, vol. 5, 2010, pp. 1-8.

[11] A. Song, X. Bai, X. Ding, and W. Zhang, "A novel multiobjective K-harmonic means clustering algorithm using Levy Flight Cuckoo Search", Journal of Computational Information Systems, vol. 9, 2013, pp. 9953-9964.

[12] X.S. Yang, "Firefly algorithms for multimodal optimization", Stochastic Algorithms: Foundations and Applications, vol. 5792, 2009, pp. 168-178.

[13] X.S. Yang, "Firefly algorithm, Levy flights and global optimization", Research and Development in Intelligent Systems Xxvi, vol. 2010, pp. 209-218.

[14] J. Senthilnath, S.N. Omkar, and V. Mani, "Clustering using firefly algorithm: Performance study", Swarm and Evolutionary Computation, vol. 1, 2011, pp. 164-171.

[15] A. Baykasoğlu, and F.B. Ozsoydan, "An improved firefly algorithm for solving dynamic multidimensional knapsack problems", Expert Systems with Applications, vol. 41, 2014, pp. 3712-3725.

[16] A. Kavousi-Fard, H. Samet, and F. Marzbani, "A new hybrid Modified Firefly Algorithm and Support Vector Regression model for accurate Short Term Load Forecasting", Expert Systems with Applications, vol. 41, 2014, pp. 6047-6056.

[17] C. Sudheer, S.K. Sohani, D. Kumar, A. Malik, B.R. Chahar, A.K. Nema, B.K. Panigrahi, and R.C. Dhiman, "A Support Vector Machine-Firefly Algorithm based forecasting model to determine malaria transmission", Neurocomputing, vol. 129, 2014, pp. 279-288.

[18] X.S. Yang: Engineering optimization: An introduction with metaheuristic applications, John Wiley & Sons 2010, pp. 221-229.