Information Sciences 258 (2014) 108–121

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Hybridising harmony search with a Markov blanket for gene selection problems

Salam Salameh Shreem, Salwani Abdullah *, Mohd Zakree Ahmad Nazri

Data Mining and Optimisation Research Group (DMO), Centre for Artificial Intelligent Technology (CAIT), Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 16 June 2012
Received in revised form 23 July 2013
Accepted 14 October 2013
Available online 22 October 2013

Keywords:
Gene selection
Filter approach
Harmony search algorithm
Markov blanket
Wrapper approach

ABSTRACT

Gene selection, which is a well-known NP-hard problem, is a challenging task that has been the subject of a large amount of research, especially in relation to classification tasks. This problem addresses the identification of the smallest possible set of genes that could achieve good predictive performance. Many gene selection algorithms have been proposed; however, because the search space increases exponentially with the number of genes, finding the best possible approach for a solution that would limit the search space is crucial. Metaheuristic approaches have the ability to discover a promising area without exploring the whole solution space. Hence, we propose a new method that hybridises the Harmony Search Algorithm (HSA) and the Markov Blanket (MB), called HSA-MB, for gene selection in classification problems. In this proposed approach, the HSA (as a wrapper approach) improvises a new harmony that is passed to the MB (treated as a filter approach) for further improvement. The addition and deletion of operators based on gene ranking information is used in the MB algorithm to further improve the harmony and to fine-tune the search space. The HSA-MB algorithm method works especially well on selected genes with higher correlation coefficients based on symmetrical uncertainty. Ten microarray datasets were experimented on, and the results demonstrate that the HSA-MB has a performance that is comparable to state-of-the-art approaches. HSA-MB yields very small sets of genes while preserving the classification accuracy. The results suggest that HSA-MB has a high potential for being an alternative method of gene selection when applied to microarray data and can be of benefit in clinical practice.

1. Introduction

Gene expression microarray dataset technology offers an appropriate and effective way to assess the pathological diagnosis and classification of cancer diseases [1]. However, it is not a practical tool for use in clinical diagnosis because the number of genes exceeds thousands or tens of thousands, which makes it impossible for doctors or biologists to examine the expression of all genes. Therefore, over the past few decades, many studies have applied machine learning approaches to analyse the gene selection for microarray datasets to enable pathological diagnosis to be used in clinical diagnosis [14]. However, most of the proposed learning algorithms suffer from an inevitable problem, that of over-fitting, which is caused by the high dimensionality of the microarray datasets and the small sample sizes [11]. To address this problem, feature selection (or gene selection in the context of microarray data) is recommended to select a small number of appropriate gene subsets and to improve the classification accuracy.

* Corresponding author. Tel.: +60 389216183.
E-mail addresses: salam@ftsm.ukm.my (S.S. Shreem), salwani@ftsm.ukm.my (S. Abdullah), mzan@ftsm.ukm.my (M.Z.A. Nazri).

0020-0255/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2013.10.012

In this context, gene selection is considered to be a necessary preprocessing step to reduce the dimensionality of these data. This step helps to improve the prediction performance of the induction algorithm, provide a better understanding and analysis of the data, and reduce the computational cost. Thus far, two main approaches have been proposed for gene selection: the filter and the wrapper [31]. In the case of filter approaches, the gene selection process is not dependent on the induction algorithm. Instead, these approaches are based on a criterion that depends only on the general characteristics of the data and that evaluates the relevance or importance of each gene for class label discrimination [16]. In the case of wrapper approaches, the classifier is used to evaluate the generated subset of genes each time [31]. However, wrapper approaches are more computationally intensive than filter approaches because the classifier is employed on each candidate subset.

Gene selection is an active research topic in areas such as data mining, pattern recognition, machine learning and health [56]. Gene selection is an NP-hard problem because the search space expands exponentially with the increase in the number of genes 2^N, where N represents the number of genes [17]. Moreover, the search for the optimal solution among a very large number of possible solutions is very difficult if not impossible because exploring the whole solution space is a computational task that involves a prohibitive amount of computational time and cost [13]. Thus, it cannot be guaranteed that the optimal subset of genes will be acquired. Additionally, a solution space typically has a large number of local optimal solutions that usually cannot be tackled by using the classical methods [34]. Therefore, the appropriate way to find the best solution from the enormous search space is to use metaheuristic techniques to obtain a good solution without exploring the whole space of solutions. Various metaheuristic techniques have been applied in many optimisation problems, and they have been proved to have superior performance and are efficient in addressing optimisation problems. However, the problem with the expression data is not only related to the dimensionality issue; it is also related to redundancy and noise. Therefore, an effective gene selection technique is needed to search for the most important subset of genes by eliminating spurious or non-predictive genes from the original dataset without sacrificing or decreasing the accuracy of the classification [29].

The Harmony Search Algorithm (HSA) is a relatively new population-based method that was proposed by Geem in 2001 [25]. HSA has been successfully applied to different optimisation problems such as structural design [24,33], energy system dispatch [46], music composition [23], nurse rostering [26], Sudoku puzzle [22], manufacturing optimisation problems [52], document clustering [21], webpage clustering [35], soil stability analysis [12], ground water modelling [7,45], heat exchanger design [20], medical physics [39], medical image analysis [5], timetabling [2,3], and image segmentation [4]. Other applications of HSA can be found in [18,38,41,47,49] due to its simplicity, generality and flexibility and its lower parameter sensitivity [15,50]. Unfortunately, an empirical study has shown that the HSA usually suffers from slow convergence [10].

The contribution of this work is that it proposes a new hybrid approach that embeds a filter, namely, the Markov Blanket (MB), into HSA (as a wrapper) to address gene selection problems. This hybrid method employs the advantages of both the HSA wrapper and the MB filter, where the HSA addresses the exploration process while the MB enhances the exploitation process in obtaining the smallest subset of genes and simultaneously gaining high classification performance. Note that the MB is a filter technique that can identify the irrelevant and redundant genes by using cross-entropy [32]. The exploitation of candidate subsets by the MB is guided by the information provided by the HSA. The improvised solution generated by the HSA is sent to the MB algorithm for further improvement; the MB fine-tunes the search by adding relevant genes and deleting irrelevant genes based on information regarding gene ranking. The MB can also accelerate the convergence speed in this hybridised approach. Experimental results show that the performance of the hybrid HSA-MB is superior to that of the HSA wrapper alone, and it is also comparable to previous approaches presented in the literature in terms of classification accuracy and the number of selected genes.

The remainder of this paper is organised as follows: Section 2 presents the related studies on gene selection problems. Section 3 describes the proposed method, and Section 4 discusses the experimental results. Finally, Section 5 presents the conclusions of this study.

2. Related work

In recent years, various techniques for gene selection problems have been proposed. Among these, metaheuristic techniques have thus far been the most extensively used, and their performance has been proved to be one of the better performing techniques that have been used for solving gene selection problems [8,55]. Notable techniques include Agrawal and Bala [1], who proposed a filter–wrapper approach that hybridised the gene-ranking method as a filter technique (to rank the genes) and a Genetic Algorithm (GA) as a wrapper technique. The accuracy of the selected genes was measured using a multi-class support vector machine. Zhu et al. [57] presented a memetic algorithm that embeds an MB filter with a GA wrapper (coded as MBEGA). In each GA generation, the MB filter is used to fine-tune the search by adding the relevant genes or removing the redundant and/or irrelevant genes from the elite chromosome, i.e., the chromosome with the highest or the best fitness function value from the population. Talbi et al. [44] designed two hybrid population-based metaheuristics, i.e., Geometric Particle Swarm Optimisation (GPSO) and the Genetic Algorithm using the Support Vector Machine classifier (GASVM) for gene selection. The results of the experiments conducted showed that the GASVM approach obtains better solutions in terms of selected genes. However, GPSO shows better performance than the GASVM in terms of the classification accuracy. Both of these hybrid methods can obtain results that are comparable to those of state-of-the-art approaches. El Akadi et al. [19] proposed a two-stage algorithm for microarray data; in the first stage,

Minimum Redundancy Maximum Relevance (MRMR) is used to filter the genes, while in the second stage, a GA is used. Later, Huang [28] presented the ant colony optimisation method, which uses the SVM to improve the classification accuracy with a small subset of genes. Chuang et al. [13] proposed a hybrid method that combines Binary Particle Swarm Optimisation (BPSO) and a Combat GA (CGA) (coded as BPSO-CGA) for gene selection, and a K-Nearest Neighbour (K-NN) with LOOCV serving as an evaluator for gene expression data selection and classification problems. In another proposal for a two-stage algorithm for microarray data by Chuang et al. [14], the first stage is a Correlation-based Gene Selection (CFS), which is used to filter genes and is then merged with a wrapper Taguchi-GA (TGA) gene selection method to create a new hybrid method, while the K-NN with the Leave One Out Cross-Validation (LOOCV) method serves as a classifier for 11 classification profiles. Kannan [30] proposed a novel correlation based on a memetic framework, which is a combination of a GA wrapper and a correlation-based filter ranking approach as a local search heuristic (coded as MA-C). The local search filter approach is used to fine-tune the population of the GA solution by adding or deleting genes based on the Symmetrical Uncertainty (SU) measurement. Then, the GA operators are applied to generate the next population while the Naïve Bayes (NB) is used as a classifier. Bonilla-Huerta et al. [9] proposed a hybrid filter–wrapper method for microarray datasets, where initially the between-group to within-group sum of square ratio filter (BSS/WSS) is used to filter and select top-ranking genes, and then, Fisher's Linear Discriminate Analysis (LDA) with a multiparent recombination operator GA is applied to the most predictive gene, which reduces the search space. Ruiz et al. [43] present a new heuristic for selecting relevant gene subsets. Their method is based on incremental ranking that occurs in two stages, i.e., filter and wrapper. In the filter stage, genes are ranked independently based on their statistical significance with respect to a class. In the wrapper stage, the first gene in the rank is added to the partial of the solution, and then, the wrapper algorithm iteratively attempts to add the next gene in the rank into the partial solution by evaluating the goodness of the new subset.

Most of the methods applied to gene selection problems use a combined filter and wrapper approach [37]. However, in the filter approach, the goodness of the genes is estimated based on a statistical gene selection according to its correlation with the class label without depending on the induction algorithm. The advantage of the filter approach is that it needs less computational time. However, the drawback of this approach is that it ignores the effect of the subset of genes in the induction algorithm, while on the other hand, in the wrapper approach, the gene subset is obtained by using the induction algorithm. The wrapper approach evaluates the quality of the selected subset of genes, and there is no doubt that the results obtained from the wrapper approach are better than those obtained from the filter approach alone [36]. However, the wrapper approach requires more computational time because a classifier is used to evaluate the generated subset of genes during each iteration. Hence, when combined, the filter and the wrapper can complement each other to obtain the most relevant subset of genes and show a higher classification accuracy. These findings motivate us to further investigate other hybridisation approaches to identify solutions for gene selection problems.

3. Harmony search algorithm with a Markov blanket

Our proposed method is a hybrid wrapper–filter, where the HSA is used as a wrapper and the MB is treated as a filter. This hybridised algorithm is coded as HSA-MB. The components and process are discussed in the following subsections.

3.1. Wrapper approach: Harmony search algorithm

The HSA has the following five steps:

- Step 1: Initialise the parameters of HSA.
- Step 2: Initialise harmony memory (HM).
- Step 3: Improvise new harmony (G').
- Step 4: Update harmony memory.
- Step 5: Check the stopping criterion.

Step 1: Initialise the parameters of the Harmony Search Algorithm

The parameters involved comprise the Harmony Memory Size (HMS), the Harmony Memory Consideration Rate (HMCR), the Pitch Adjusting Rate (PAR), the maximum number of selected genes, and the number of improvisations (NI). The values for HMCR and PAR are in the range of [0,1], and NI is treated as a stopping criterion.

Step 2: Initialise the harmony memory

Harmony memory is a two-dimensional matrix with the size HMS; Harmony memory is composed of randomly generated solutions. Each row in the HM represents one chromosome (solution), as shown in Fig. 1. Solutions in HM are arranged in reverse order based on the fitness values (refer to $f(G)$ in Fig. 1). In this work, the fitness value is referred to as the classification accuracy (which is obtained using the NB classifier based on the selected genes).

$$HM = \begin{bmatrix} g_1^1 & g_2^1 & \cdots & g_N^1 & \rightarrow & f(G^1) \\ g_1^2 & g_2^2 & \cdots & g_N^2 & \rightarrow & f(G^2) \\ \vdots & \vdots & \cdots & \vdots & \rightarrow & \vdots \\ g_1^{HMS-1} & g_2^{HMS-1} & \cdots & g_N^{HMS-1} & \rightarrow & f(G^{HMS-1}) \\ g_1^{HMS} & g_2^{HMS} & \cdots & g_N^{HMS} & \rightarrow & f(G^{HMS}) \end{bmatrix}$$

Fig. 1. Harmony memory.

Step 3: Improvise a new harmony

A new harmony $G' = (g'_1, g'_2, \ldots, g'_N)$ is improvised based on three rules, i.e., memory consideration, pitch adjustment and random consideration, as follows:

(i) Memory consideration

In memory consideration, a new solution is generated based on the HMCR. A random number, R, is generated between [0,1]. If R < HMCR, then the first gene is selected from $(g'_1, \ldots, g_1^{HMS})$. The next gene, g'_2, is chosen from $(g'_2, \ldots, g_2^{HMS})$, and the process goes on; otherwise, the gene is determined based on the random consideration process (see below).

(ii) Pitch adjustment

Every gene obtained by the HMCR is examined to determine whether it should be tuned ('pitch-adjusted') with the probability of PAR or left unchanged with the probability (1-PAR). The adjustment here will mutate the gene from 0 to 1 or vice versa. This operation uses the PAR as in:

$$g'_i \leftarrow \begin{cases} mutate\, g'_1 & \text{w.p} \quad \text{PAR} \\ g'_1 & \text{w.p} \quad \text{1-PAR} \end{cases} \tag{1}$$

(iii) Random consideration

In random consideration, genes that are not selected from the HM with the probability of (1-HMCR) are selected randomly according to their possible range of values, as shown in Eq. (2) (the possible range value in our problem is either 0 or 1). This procedure helps to increase the diversity of the solutions, and thus various solutions can be explored to obtain the global optimum.

$$g'_i \leftarrow \begin{cases} g'_i \in \{g_i^1, g_i^2, \ldots, g_i^{HMS}\} & \text{w.p} \quad \text{HMCR} \\ g'_i \in X^i & \text{w.p} \quad \text{(1-HMCR)} \end{cases} \tag{2}$$

Step 4: Update Harmony Memory

Harmony Memory is updated by replacing the worst solution in the HM with the improvised harmony $G' = (g'_1, g'_2, \ldots, g'_N)$ if the quality of the new harmony is better. Otherwise, the new harmony is ignored.

Step 5: Check the stopping criterion

The maximum NI (or classification accuracy = 100%) is treated as the stopping criterion.

3.2. Filter approach: Markov blanket

The MB was proposed by Koller and Sahami [32]. The MB is a cross-entropy technique that considers the relevance between different genes; it is used to identify and eliminate the irrelevant and redundant genes in a general classification problem. The main idea of the MB is to safely remove gene g_i from the existing gene subset without degrading or improving their performance (classification accuracy). The MB of a gene, g_i, is defined as follows:

Definition 1 (*Markov Blanket*). Let F be a full set of genes, let C be the class, and let M be a subset of genes that does not contain the gene g_i, $M \subseteq F$ and $g_i \in M$. M is an MB of g_i if g_i is conditionally independent of $(F \bigcup C) - M - \{g_i\}$ given M, i.e., $P(F - M - \{g_i\}, C \setminus g_i, M) = P(F - M - \{g_i\}, C \setminus M)$. It is clear that if M is an MB of g_i, then C is also conditionally independent of F_i given M.

Two genes, g_i and g_j, are conditionally independent given gene subsets with a specific size K, if $P(g_i \setminus K, g_j) = P(g_i \setminus K)$; in other words, g_j gives no information about g_i beyond what is already in K. If a gene, g_i, has a MB of M within the currently selected gene subset, then g_i gives no more information beyond M about C and other selected genes; therefore, g_i can be removed safely. However, because finding the MB of a gene might be computationally infeasible, Yu and Liu [53,54] considered using only one gene to approximate the MB of g_i.

Definition 2 (*Approximate Markov Blanket*). Given two genes, g_i and g_j ($g_i \neq g_j$), gene g_j is said to be an approximate MB of g_i if $SU_{j,c} \geqslant SU_{i,c}$ and $SU_{i,j} \geqslant SU_{i,c}$, where SU (symmetrical uncertainty) [48] measures the correlation between the genes and the class C (called a C-correlation). A gene is thus considered to be relevant if its C-correlation is higher than a given threshold, which is a user-specific value, i.e., $SU_{i,c} \geqslant \gamma$. In this work, the top 10% of the ranking genes are considered to be the most relevant genes correlated to the class based on the preliminary experiments (10 independent runs) based on 100%, top 50% and top 10% of the ranking genes using the MLL and Colon datasets. The top 10% of the ranking genes gives the best results (in terms of the classification accuracy), as shown in Table 1. The best results are presented in bold.

In this work, the approximate MB is employed to calculate the quality of each gene based on:

$$SU = 2.0 \times \left[\frac{IG(X|Y)}{H(X) + H(Y)}\right] \tag{3}$$

where $IG(X|Y)$ is the information gain [40] between genes X and Y, which is represented by:

$$IG(X|Y) = H(X) - H(X|Y) \tag{4}$$

where $H(X)$ is the entropy of X, which is defined as:

$$H(X) = -\sum_i P(x_i)\log_2(P(x_i)) \tag{5}$$

where $P(x_i)$ is the prior probability for all of the values of X. The entropy of X after observing the values of another variable Y is defined as follows:

$$H(X|Y) = -\sum_j P(x_i)\sum_i (x_i|y_j)\log_2(x_i|y_j)) \tag{6}$$

where $P(x_i)$ is the posterior probability of x_i, given the values of y_i.

3.3. Hybridisation of the Harmony search algorithm and Markov blanket

3.3.1. Solution representation

The solution is encoded as a chromosome that is composed of a binary string that has a length equal to the number of genes + 1 (where the last bit represents the fitness value, which is calculated based on Eq. (7)). Each bit encodes a single gene, where a bit of '1' ('0') implies that the corresponding gene is selected (excluded).

3.3.2. Fitness function

The fitness function is defined by the classification accuracy, as stated in:

$$f(G) = J(S_g) \tag{7}$$

where S_g denotes the selected gene encoded in chromosome G', and $J(S_g)$ is specified as the classification accuracy for S_g. If two chromosomes have similar fitness, then the chromosome with the smaller number of selected genes is chosen.

3.3.3. Neighbourhood structure

Two neighbourhood structures are employed in this work, as follows:

- Nbs_1 (add): selects the gene with the highest rank from the *excluded subset* and adds it to the *selected subset*.
- Nbs_2 (delete): selects the gene with the highest rank g_i from the *selected subset* and deletes the remaining genes from this subset.

The simple pseudo-code for Nbs_1 and Nbs_2 is presented in Figs. 2 and 3, respectively:

Table 1
Preliminary experiment for threshold γ.

	MLL (%)	Colon (%)
100%	94.5	80.7
Top 50%	95.7	87.1
Top 10%	**98.7**	**90.3**

Nbs_1: Add
Begin
Select the highest ranking gene g_i from the *excluded subset*;
Add g_i to the *selected subset*;
End

Fig. 2. Nbs_1 – add neighbourhood structure.

Nbs_2: Delete
Begin
Select the highest ranking gene g_j from the *selected subset*;
Delete other genes (G-g_i) from this subset that are in the approximate MB of g_i, where G is a full set of genes;
End

Fig. 3. Nbs_2 – delete neighbourhood structure.

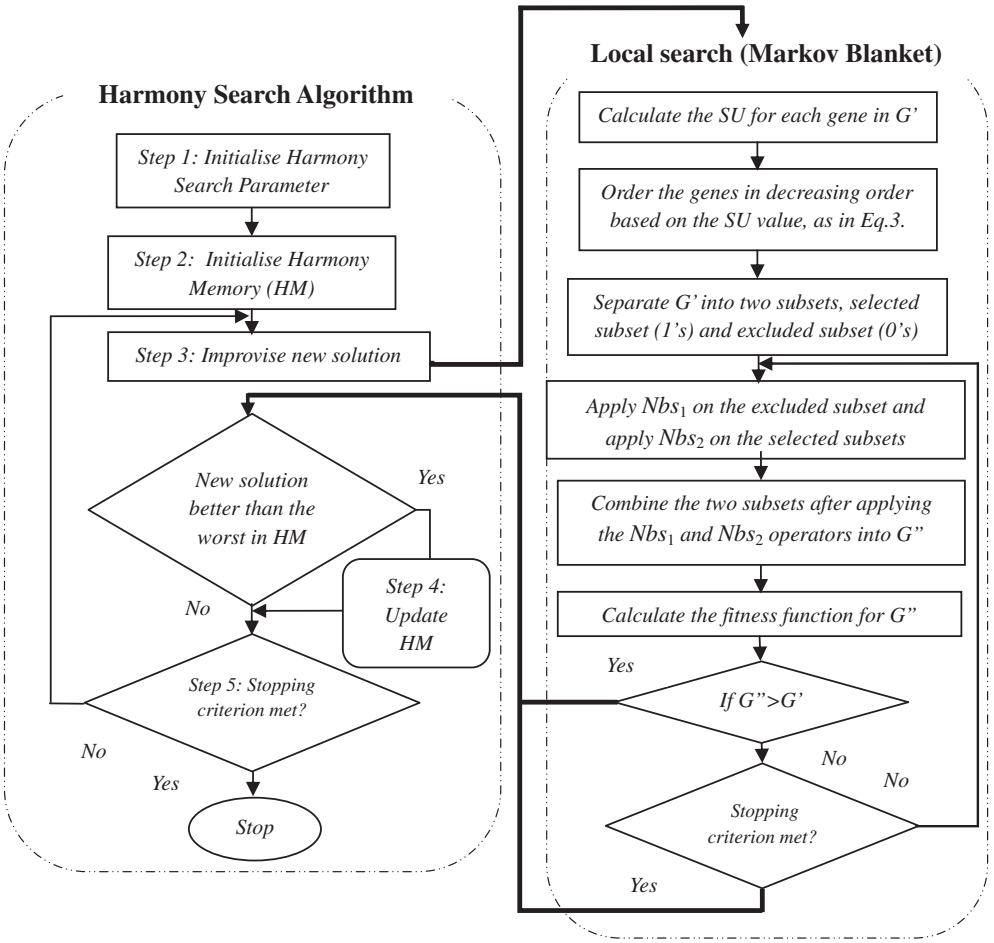

Fig. 4. Harmony search algorithm with Markov blanket.

Table 2
Description of the datasets.

Data set	Genes	Samples	Classes	Description
ALL-AML	7129	72	2	Two acute leukemia, i.e., Acute Myelogenous Leukemia (AML) and Acute Lymphoblastic Leukemia (ALL)
ALL-AML-3C	7129	72	3	AML, ALL B-cell, and ALL T-Cell
ALL-AML-4C	7129	72	4	AML-Bone Marrow, AML-Peripheral Blood, ALL B-cell, and T-Cell
Colon	2000	62	2	40 colon cancer biopsies vs. 22 normal biopsies
CNS	7129	60	2	Outcome of the treatments for 60 central nervous system cancer patients (21 survivors and 39 failures)
Lymphoma	4026	62	3	Three most prevalent adult lymphoid tumors
MLL	12582	72	3	AML, ALL, and mixed-lineage leukemia (MLL)
Breast	24481	97	2	97 samples from breast cancer patients (46 patients developed distance metastases; the other 51 remained healthy after their initial diagnosis for an interval of at least five years)
Ovarian	15154	253	2	Proteomic spectra of 91 normal persons and 162 ovarian cancer patients
SRBCT	2308	83	4	Small, round blue cell tumors (SRBCT) from childhood

3.3.4. HSA-MB algorithm

The hybridisation of HSA and MB is presented in Fig. 4. At the beginning of the search process, all of the parameters used in the harmony search are initialised (Step 1). The parameter settings are presented in Table 2. The HM is initialised by randomly generating a number of initial solutions, which equates to the HMS (Step 2). The quality of each solution in the HM is measured based on the classification accuracy using the Naïve Bayes classifier. A new harmony or solution $G' = (g'_1, g'_2, \ldots, g'_N)$ is improvised based on three rules, which are memory consideration, pitch adjustment and random consideration (Step 3). The improvised harmony (G') is passed onto the MB local search for further improvement. During the MB local search (which is where the filter process takes place), the quality for each gene in G' is calculated based on Eq. (3). These genes are ordered in decreasing order based on the SU value. Next, G' is divided into two subsets, which are referred to as the *selected subset* (1's) and the *excluded subset* (0's). Then, Nbs_1 and Nbs_2 is applied to the (0's) and (1's) subsets, respectively, and these subsets are combined to generate G''. The quality of G'' (in terms of the classification accuracy) is evaluated. If G'' is better than G', and G'' better than the worst solution in HM then the HM is updated (Step 4). The process (in the MB local search) is repeated by re-employing Nbs_1 and Nbs_2 until the termination criterion is met (set as the maximum number of iterations, which is equal to 5). Note that a small value is set here to save some computational time and to provide more opportunities for other solutions to be explored. The flow then goes back to the HSA. This process is repeated until the stopping criterion (which is set to be either a maximum number of iterations or the classification accuracy is equal to 100%) is met (Step 5). The process is illustrated in Fig. 4.

4. Experimental results

The proposed algorithms are programmed using Java in the WEKA environment [27], and simulations are performed on an Intel Core i5-2450M-2.5 GHz CPU with 4 GB of RAM. The WEKA is used as a classifier tool, where an NB classifier with 10-fold cross-validation (as recommended by Ambroise and McLachlan [6]) is applied to validate and assess the generated solutions. Three comparisons are performed, i.e., a comparison of HSA-MB with (i) HSA alone; (ii) with GA (as one example of evolutionary algorithms) with and without MB; and (iii) state-of-the-art methods.

The results reported in this section are obtained by performing a 10-fold cross-validation over each dataset, where the datasets are partitioned into training sets (90%) and independent test sets (10%). In other words, gene subsets are selected from 90% of the training instances, and then, the accuracy is estimated over the unseen 10% of the test instances. This process is performed 10 times. The accuracy and number of selected genes reported in this work are based on 10-fold cross-validation, and they are based on test data.

4.1. Datasets

The performance of the proposed algorithm, HSA-MB, is evaluated by applying it to 10 well-known microarray datasets (refer to Table 2), which can be freely downloaded from http://csse.szu.edu.cn/staff/zhuzx/Datasets.html. There is a large difference between the number of genes and the number of samples across the selected datasets, which means that the experiment reflects the challenge of effectively dealing with such varying dimensionalities.

4.2. Parameter settings

Table 3 shows the parameters for the proposed algorithm; the values for these parameters were determined after some preliminary experiments.

The details of the preliminary experiments are discussed below.

Table 3
Parameter settings.

Parameter	Value
Harmony Memory Size (HMS)	50
Harmony Memory Consideration Rate (HMCR)	0.7
Pitch Adjustment rate (PAR)	0.3
Number of Improvisations (NI)	50
Maximum number of selected genes	50

4.2.1. HMS

We first examine the effect of HMS on three microarray datasets with a different number of genes, i.e., a small dataset with 2000 genes (Colon), a medium dataset with 12582 genes (MILL) and a large dataset with 24481 genes (Breast). In this experiment, we fix the parameter values as HMCR = 0.7, PAR = 0.3 and NI = 100. Table 4 presents the experimental results (average out of 10 runs) for different HMS parameter values. The best results are presented in bold.

Based on the experimental results presented in Table 4, it can be seen that HMS = 50 is the most suitable parameter value compared to other HMS values. Therefore, HMS = 50 is chosen for all of the tested instances.

4.2.2. HMCR and PAR

We examine several HMCR and PAR values, as shown in Table 5. We fix the NI to 50 iterations. In the literature, the recommended value for HMCR is between 0.70 and 0.99, while for PAR, it is between 0.1 and 0.3 [33,42].

From Table 5, it can be seen that the best results (in bold) are obtained when HMCR = 0.7 and PAR = 0.3, i.e., on Colon and MLL. Therefore, these values are chosen in this work.

4.2.3. NI

To determine a suitable parameter value for NI, we perform an experiment with different values of NI and with fixed values of HMS = 50, HMCR = 0.7 and PAR = 0.3. The obtained results are presented in Table 6. It can be seen that there is a significant increase in the quality of the solution during the first 50 iterations. Subsequently (between 50 and 100 iterations), the change in the classification accuracy is small or there is almost no improvement. Thus, NI = 50 is chosen in this work.

4.3. Comparison of HSA-MB and HSA using the Naïve Bayes Classifier

Table 7 shows the results obtained from HSA-MB and from HSA alone. The best results are shown in bold. It can be seen that HSA-MB outperforms HSA alone on all of the tested datasets with a high classification accuracy (>90%) except for the Breast and CNS datasets (80.06% and 84.17%, respectively), and with a 99.99% accuracy for the Lymphoma dataset. In terms of the number of selected genes, again the HSA-MB performs better than HSA on all of the tested datasets. From the results, it

Table 4
Results of using HSA with different HMS parameter values.

HMS	Colon	MLL	Breast
HMS = 2	56.36	82.65	55.17
HMS = 10	64.6	85.64	54.02
HMS = 50	**69.8**	**86.28**	**57.47**
HMS = 100	63.63	84.69	55.17
HMS = 150	65.45	83.26	54.02
HMS = 200	52.45	79.92	52.87

Table 5
Results of using HSA with different HMCR and PAR parameter values.

HMCR	PAR	Colon	MLL	Breast
0.7	0.1	61.03	82.62	56.32
	0.2	63.10	81.52	**57.58**
	0.3	**69.80**	**86.28**	57.47
0.89	0.1	63.63	81.28	56.32
	0.2	65.45	81.25	57.47
	0.3	67.27	81.30	57.47
0.99	0.1	57.23	78.26	51.72
	0.2	57.26	77.62	52.87
	0.3	59.26	80.25	54.02

Table 6
Classification accuracy of solutions in HM with different NI values.

No. of iterations	Colon	MLL	Breast
Initial value	56.36	63.07	51.72
After 10	63.63	67.69	54.02
After 20	65.45	75.38	55.17
After 30	67.27	78.46	56.32
After 40	69.09	81.53	56.32
After 50	69.09	83.07	56.32
After 60	69.09	84.37	56.32
After 70	69.09	84.37	56.32
After 80	69.09	84.95	56.32
After 90	69.09	84.95	56.32
After 100	69.09	84.95	56.32

Table 7
Comparison between HSA-MB and HSA.

Algorithm		Dataset				
		ALL-AML	ALL-AML-3C	ALL-AML-4C	Colon	CNS
HSA-MB	$\|\#G\|$	**5.00**	**5.84**	**6.37**	**4.16**	**7.43**
	ACC	**99.34**	**99.18**	**96.79**	**90.27**	**84.17**
	T	1:42	3:53	2:21	2:22	1:41
HSA	$\|\#G\|$	17.39	26.61	27.88	16.05	10.70
	ACC	93.09	86.25	83.56	73.36	73.08
	T	**1:10**	**2:28**	**1:44**	**1:19**	**1:02**
		Lymphoma	MLL	Breast	Ovarian	SRBCT
HSA-MB	$\|\#G\|$	**3.75**	**6.60**	**5.06**	**5.73**	**8.9**
	ACC	**99.99**	**99.55**	**80.06**	**99.81**	**99.57**
	T	1:32	2:32	12:15	50:00	3:28
HSA	$\|\#G\|$	19.94	28.80	13.01	21.31	22.93
	ACC	97.57	91.59	58.18	94.50	91.40
	T	**1:03**	**1:32**	**5:23**	**11:49**	**2:55**

Note: $|\#G|$: average number of genes; ACC: average classification accuracy (%); T: average time (minutes).

Table 8
p-Values between HSA-MB and HSA.

Datasets	ACC	$\|\#G\|$
ALL-AML	0.000	0.000
ALL-AML-3C	0.000	0.000
ALL-AML-4C	0.000	0.000
Colon	0.000	0.000
CNS	0.000	0.000
Lymphoma	0.000	0.000
MLL	0.000	0.000
Breast	0.000	0.000
Ovarian	0.000	0.000
SRBCT	0.000	0.000

can be seen that HSA-MB can reduce the redundant and irrelevant genes efficiently when compared to HSA. It is believed that this is due to the significant role played by the MB, which acts as a local search that can fine-tune the search space of the new improvised harmony toward the local optima through adding and deleting relevant/irrelevant genes during the exploitation process. It is also believed that HSA-MB obtains genes that are highly correlated and that can render the remaining genes independent of the variable of interest, thus positively influencing the classification accuracy.

The proposed HSA-MB method can find the most suitable genes for the induction algorithm and eliminate the most redundant genes without affecting the classification accuracy because it employs cooperative classification via a wrapper approach and correlation redundancy among genes via a filter ranking approach with an SU measure as a local search heuristics. Thus, HSA-MB can improve the classification accuracy with fewer selected genes compared to HSA because the latter is solely based on the induction algorithm. However, HSA-MB is more computationally demanding because the improvised solution is passed to an MB local search at each iteration.

A Wilcoxon rank test is performed to ascertain whether there is any significant difference between HSA-MB and HSA (in terms of the average classification accuracy, ACC, and the average number of genes, $|\#G|$) with a significance interval of 95% ($\alpha = 0.05$). Table 8 shows the obtained *p*-values.

The *p*-values show that there is a significant difference between the compared algorithms, which leads to the conclusion that the hybridisation of the MB filter with the HSA wrapper (HSA-MB) leads to superior performance when compared to the HSA wrapper approach alone.

4.4. Comparison between HSA and GA with and without MB

In this work, we choose to compare our proposed method with the GA as the evolutionary algorithm. Table 9 shows the parameter settings for the GA [57].

Two comparisons are conducted, i.e., a comparison between HSA and GA in isolation and a comparison between the hybridisation of HSA with MB (HSA-MB) and the hybridisation of GA with MB (GA-MB). To make a fair comparison, we set the stopping criterion to be a maximum number of improvisations or a classification accuracy of 100%. We execute 31 independent runs. The results are presented in Table 10.

The results in Table 10 show that HSA and HSA-MB outperform GA and GA-MB on most of the tested datasets, respectively (as shown in bold). Because of the high dimensionality of the microarray datasets, the basic GA and GA-MB cannot obtain good results as HAS-MB, which we believe is due in part to the drawbacks of the basic GA, which relies on only

Table 9
Parameter settings for the GA.

Parameter	Value
Population size	50
Number of generations	50
Crossover probability	0.6
Mutation rate	0.5

Table 10
Results of the comparisons (i) HSA vs. GA and (ii) HSA-MB vs. GA-MB.

Datasets		HSA vs. GA		HSA-MB vs. GA-MB	
		HSA	GA	HSA-MB	GA-MB
ALL-AML	$\|\#G\|$	**17.39**	26.2	**5.00**	8.5
	ACC	93.09	**93.87**	**99.34**	96.47
	T	**1:10**	1:45	**1:42**	3:52
ALL-AML-3C	$\|\#G\|$	**26.61**	41.7	**5.84**	23.8
	ACC	**86.25**	83.68	**99.18**	93.27
	T	**2:28**	2:48	**3:53**	7:38
ALL-AML-4C	$\|\#G\|$	**27.88**	30.4	**6.37**	27.3
	ACC	83.56	**86.14**	**96.79**	90.65
	T	**1:44**	2:27	**2:21**	3:29
Colon	$\|\#G\|$	16.05	**15.6**	**4.16**	10.9
	ACC	**73.36**	69.72	**90.27**	83.92
	T	**1:19**	1:57	**2:22**	5:07
CNS	$\|\#G\|$	**10.70**	30.32	**7.43**	9.31
	ACC	**73.08**	65.96	**84.17**	79.09
	T	**1:02**	1:18	1:41	**1:25**
Lymphoma	$\|\#G\|$	**19.94**	28.1	**3.75**	14.9
	ACC	**97.57**	91.62	**99.99**	98.08
	T	**1:03**	2:46	1:32	**1:12**
MLL	$\|\#G\|$	**28.80**	44.8	**6.60**	23.1
	ACC	**91.59**	82.14	**99.55**	96.64
	T	**1:32**	1:51	**2:32**	3:05
Breast	$\|\#G\|$	13.01	**8.5**	**5.06**	5.2
	ACC	**58.18**	56.12	**80.06**	68
	T	5:23	**3:56**	**12:15**	16:23
Ovarian	$\|\#G\|$	**21.31**	23.4	**5.73**	19.6
	ACC	**94.50**	90.69	**99.81**	96.30
	T	**11:49**	21:45	**50:00**	87:45
SRBCT	$\|\#G\|$	**22.93**	31.1	**8.9**	29.4
	ACC	**91.40**	85.49	**99.57**	95.47
	T	**2:55**	4:18	**3:28**	5:32

Note: $|\#G|$: average number of genes; ACC: average classification accuracy (%); *T*: average time (minutes).

two parents to produce new offspring [51]. However, GA-MB can obtain better results compared to the basic GA in terms of the classification accuracy and the number of selected genes. This finding is due to the significant role played by the MB, which acts as a local search that can fine-tune the search space of the new solution toward better local optima through adding and deleting relevant/irrelevant genes during the exploitation process.

We further validate our results by conducting a statistical pairwise comparison test to ascertain whether there is any significant difference, using a significance interval of 95% ($\alpha = 0.05$). Table 11 presents the obtained p-values. The bold values show that there is no significant difference.

It is clear from the p-values that there is a significant difference between HSA-MB and GA-MB on 10 and eight datasets, respectively, in terms of the accuracy and the number of selected genes. A significant difference also exists between HSA and GA. This finding is supported by the results previously presented in Table 10; hence, it can be concluded that HSA is better than GA in solving gene selection problems.

Table 11
p-Values between (i) HSA-MB and GA-MB, and (ii) HSA and GA.

Datasets	HSA-MB and GA-MB		HSA and GA	
	ACC	$\|\#G\|$	ACC	$\|\#G\|$
ALL-AML	0.000	**0.296**	0.000	0.000
ALL-AML-3C	0.000	0.000	0.000	0.000
ALL-AML-4C	0.000	0.000	0.000	0.000
Colon	0.000	0.000	0.000	**0.338**
CNS	0.000	0.000	0.000	0.000
Lymphoma	0.000	0.000	0.000	0.000
MLL	0.000	0.000	0.000	0.000
Breast	0.000	**0.155**	0.000	0.000
Ovarian	0.000	0.000	0.000	0.000
SRBCT	0.000	0.000	0.000	0.000

Table 12
Comparison of HSA-MB with state-of-the-art methods.

Datasets		HSA-MB	MBEGA	MRMR-GA	MA-C	BPSO-CGA	GPSO	BIRSW	LDA-GA
ALL_AML	$\|\#G\|$	5.00	12.8	15	387	300	3	**2.5**	3
	ACC	99.34	95.89	**100**	99.56	**100**	97.38	93.04	99.5
	T	**1:42**	1:52	–	–	–	–	–	30–35
ALL-AML-3C	$\|\#G\|$	**5.84**	18.1	–	394	–	–	–	–
	ACC	99.18	96.64	–	**99.53**	–	–	–	–
	T	3:53	**2:56**	–	–	–	–	–	
All-AML-4C	$\|\#G\|$	**6.37**	26.2	–	386	–	–	–	–
	ACC	96.79	91.93	–	**98.61**	–	–	–	–
	T	**2:21**	3:54	–	–	–	–	–	
Colon	$\|\#G\|$	4.16	24.5	15	–	214	**2**	3.5	7
	ACC	90.27	85.66	98.39	–	96.7	**100**	85.48	98.83
	T	2:22	**1:10**	–	–	–	–	–	30–35
CNS	$\|\#G\|$	7.43	20.5	–	374	–	–	–	**4**
	ACC	84.17	72.21	–	**97.78**	–	–	–	**99.3**
	T	1:41	**1:21**	–	–	–	–	–	30–35
Lymphoma	$\|\#G\|$	**3.75**	34.3	15	–	196	–	10.3	–
	ACC	99.99	97.68	98.96	–	**100**	–	82.14	–
	T	**1:32**	2:22	–	–	–	–	–	
MLL	$\|\#G\|$	**6.60**	32.1	–	108	–	–	–	–
	ACC	99.55	94.33	–	**100**	–	–	–	–
	T	**2:32**	3:02	–	–	–	–	–	
Breast	$\|\#G\|$	5.06	14.5	–	183	–	**4**	–	–
	ACC	80.06	80.74	–	**95.26**	–	86.35	–	–
	T	12:15	**4:16**	–	–	–	–	–	
Ovarian	$\|\#G\|$	5.73	9	–	247	–	**4**	–	6
	ACC	99.81	99.71	–	**100**	–	99.4	–	97.4
	T	50:00	44.49	–	–	–	–	–	**1:10**
SRBCT	$\|\#G\|$	**8.9**	60.7	–	526	880	–	–	–
	ACC	99.57	99.23	–	**100**	**100**	–	–	–
	T	**3:28**	4:06	–	–	–	–	–	

Note: $|\#G|$: average number of genes; ACC: average classification accuracy (%); T: average time (minutes); '–': not available.

4.5. Comparison with state-of-the-art methods

Table 12 shows a comparison of the results obtained by the proposed approach and other approaches identified in the literature review. Again, the best results are presented in bold. The approaches that are compared with the proposed method are as follows:

- MBEGA: Markov Blanket-Embedded Genetic Algorithm for gene selection [57].
- MRMR-GA: Minimum Redundancy–Maximum Relevancy with a Genetic Algorithm [19].
- MA-C: Correlation-based memetic framework [30].
- BPSO-CGA: Binary Particle Swarm Optimisation and a Combat Genetic Algorithm [13].
- GPSO: Geometric Particle Swarm Optimisation [44].
- BIRSW: Best Incremental Ranked Subset [43].
- LDA-GA: Fisher's Linear Discriminate Analysis-based GA [9].

The results in Table 12 show that, in terms of the number of selected genes, HSA-MB outperforms MRMR-GA, BPSO-CGA, BIRSW and LDA-GA on three, four, one and two datasets, respectively. However, note that these methods (MRMR-GA, BPSO-CGA, BIRSW and LDA-GA) are not tested on all of the considered instances. With respect to the classification accuracy, HSA-MB is better than MRMR-GA, GPSO, BIRSW and LDA-GA on one, two, three and one instances, respectively. HSA-MB obtains competitive classification accuracy results on three out of four instances (i.e., ALL-AML, Lymphoma, and SRBCT), when compared to BPSO-CGA. Although HSA-MB did not outperform GPSO in terms of the selected genes, it did achieve a higher classification accuracy for two instances (ALL_AML and Ovarian). (Note that the GPSO is tested on six instances, four of which are in common with HSA-MB.)

We are also specifically interested in comparing HSA-MB with MBEGA and MA-C because these methods use a local search as the filter approach. Furthermore, these algorithms have been tested on at least eight datasets (out of 10), in contrast to the other state-of-the-art algorithms, which were tested on only three to four datasets (it is believed that this result occurs because of the complexity of these instances). From Table 12, it is obvious that, in terms of the number of selected genes, HSA-MB can obtain better results on 10 and eight datasets when compared with MBEGA and MA-C, respectively. In terms of the classification accuracy, HSA-MB outperforms MBEGA on all of the tested datasets except for the Breast dataset. When compared to MA-C, it can be seen that HSA-MB obtains competitive classification accuracy on all of the datasets except for Breast and CNS.

From these results, it can be concluded that HSA-MB is unable to outperform MA-C because a larger number of genes (generated by MA-C) is used to measure the classification accuracy. Theoretically, more genes provide a higher classification accuracy (even though this relationship is not applicable in all cases: a higher number of selected genes not only will slow down the learning process but also could result in the selection of some irrelevant genes that might later provide incorrect results). Nevertheless, there is a higher difference (in%) between HSA-MB and MA-C in terms of the selected number of genes (98.04%) and a small difference in terms of the classification accuracy (8.74%). From the results presented in Table 12, it can be observed that the number of selected genes is not necessarily associated with a high classification accuracy; there are cases here in which a small number of genes can obtain competitive or superior classification accuracy (e.g., the performance of HSA-MB on the Lymphoma dataset). Again, this finding is, of course, due to the incorporation of the MB filter, which has the ability to filter redundant and irrelevant genes efficiently and, thus, keeps only relevant and useful genes that can depict the remaining genes independently of the target variable.

5. Conclusions

In this study, a new approach that combines the Harmony Search Algorithm (HSA) and Markov Blanket (MB), called HSA-MB, for gene selection is investigated. This method is a hybrid wrapper–filter approach, where the HSA is used as a wrapper while the MB is employed as the filter. Experimental results show that HSA-MB is better than HSA in isolation. A comparison with GA also reveals that HSA shows better performance on all of the tested datasets. The improved performance is obtained when HSA-MB is compared with GA-MB. Furthermore, a comparison with several state-of-the-art methods shows that our proposed approach can obtain five (out of 10) new best results in terms of the number of selected genes and competitive results in terms of the classification accuracy. This finding occurs because HSA-MB is capable of eliminating irrelevant and redundant genes effectively due to the combination of the HSA wrapper and the MB filter, which leads to the identification of a small set of reliable genes.

References

[1] R. Agrawal, R. Bala, A hybrid approach for selection of relevant features for microarray datasets, International Journal of Computer and Information Science and Engineering 1 (2007) 196–202.
[2] M.A. Al-Betar, A. Khader, I. Liao, A harmony search with multi-pitch adjusting rate for the university course timetabling, Recent Advances in Harmony Search Algorithm (2010) 147–161.
[3] M.A. Al-Betar, A.T. Khader, A harmony search algorithm for university course timetabling, Annals of Operations Research (2008) 1–29.

[4] O.M. Alia, R. Mandava, D. Ramachandram, M.E. Aziz, Dynamic fuzzy clustering using harmony search with application to image segmentation, IEEE International Symposium on Signal Processing and Information Technology (ISSPIT) (2009) 538–543.
[5] O.M. Alia, R. Mandava, D. Ramachandram, M.E. Aziz, Harmony search-based cluster initialization for fuzzy c-means segmentation of MR images, TENCON 2009, in: 2009 IEEE Region 10 Conference, 2009, pp. 1-6.
[6] C. Ambroise, G.J. McLachlan, Selection bias in gene extraction on the basis of microarray gene-expression data, Proceedings of the National Academy of Sciences 99 (2002) 6562–6566.
[7] M.T. Ayvaz, Simultaneous determination of aquifer parameters and zone structures with fuzzy c-means clustering and meta-heuristic harmony search algorithm, Advances in Water Resources 30 (2007) 2326–2338.
[8] P. Bermejo, J.A. Gámez, J.M. Puerta, A GRASP algorithm for fast hybrid (filter–wrapper) feature subset selection in high-dimensional datasets, Pattern Recognition Letters 32 (2011) 701–711.
[9] E. Bonilla-Huerta, B. Duval, J. Hernández, J.K. Hao, R. Morales-Caporal, Hybrid filter–wrapper with a specialized random multi-parent crossover operator for gene selection and classification problems, Bio-Inspired Computing and Applications (2012) 453–461.
[10] P. Chakraborty, G. Roy, B. Panigrahi, R. Bansal, A. Mohapatra, Dynamic economic dispatch using harmony search algorithm with modified differential mutation operator, Electrical Engineering (Archiv fur Elektrotechnik) (2012) 1–9.
[11] S.-C. Chen, S.-W. Lin, S.-Y. Chou, Enhancing the classification accuracy by scatter-search-based ensemble approach, Applied Soft Computing 11 (2011) 1021–1028.
[12] Y. Cheng, L. Li, T. Lansivaara, S. Chi, Y. Sun, An improved harmony search minimization algorithm using different slip surface generation methods for slope stability analysis, Engineering Optimization 40 (2008) 95–115.
[13] L.Y. Chuang, C.H. Yang, J.C. Li, A hybrid BPSO-CGA approach for gene selection and classification of microarray data, Journal of Computational Biology 19 (2011) 1–14.
[14] L.Y. Chuang, C.H. Yang, K.C. Wu, A hybrid feature selection method for DNA microarray data, Computers in Biology and Medicine 41 (2011) 228–237.
[15] C. Cobos, D. Estupiņán, J. Pérez, GHS + LEM: global-best harmony search using learnable evolution models, Applied Mathematics and Computation 218 (2011) 2558–2578.
[16] J. Derrac, C. Cornelis, S. García, F. Herrera, Enhancing evolutionary instance selection algorithms by means of fuzzy rough set based feature selection, Information Sciences 186 (2012) 73–92.
[17] A. Duval, J.-K. Hao, J.C.H. Hernandez, A memetic algorithm for gene selection and molecular classification of cancer, in: Proceedings of the 11th Annual Conference on Genetic and Evolutionary Computation, 2009, pp. 201–208.
[18] M. El-Abd, Performance assessment of foraging algorithms vs. evolutionary algorithms, Information Sciences 182 (2012) 243–263.
[19] A. El Akadi, A. Amine, A. El Ouardighi, D. Aboutajdine, A two-stage gene selection scheme utilizing MRMR filter and GA wrapper, Knowledge and Information Systems 26 (2011) 487–500.
[20] M. Fesanghary, E. Damangir, I. Soleimani, Design optimization of shell and tube heat exchangers using global sensitivity analysis and harmony search algorithm, Applied Thermal Engineering 29 (2009) 1026–1031.
[21] R. Forsati, M. Mahdavi, M. Shamsfard, M. Reza Meybodi, Efficient stochastic algorithms for document clustering, Information Sciences 220 (2013) 269–291.
[22] Z. Geem, Harmony search algorithm for solving sudoku, Knowledge-Based Intelligent Information and Engineering Systems 4692 (2007) 371–378.
[23] Z. Geem, J.Y. Choi, Music composition using harmony search algorithm, Applications of Evolutionary Computing (2007) 593–600.
[24] Z.W. Geem, Harmony Search Algorithms for Structural Design Optimization, Springer-Verlag New York Inc., 2009.
[25] Z.W. Geem, J.H. Kim, G. Loganathan, A new heuristic optimization algorithm: harmony search, Simulation 76 (2001) 60–68.
[26] M. Hadwan, M. Ayob, N.R. Sabar, R. Qu, A harmony search algorithm for nurse rostering problems, Information Sciences (2013) 126–140.
[27] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA data mining software: an update, ACM SIGKDD Explorations Newsletter 11 (2009) 10–18.
[28] C.L. Huang, ACO-based hybrid classification system with feature subset selection and model parameters optimization, Neurocomputing 73 (2009) 438–448.
[29] M.M. Kabir, M. Shahjahan, K. Murase, A new hybrid ant colony optimization algorithm for feature selection, Expert Systems with Applications 39 (2012) 3747–3763.
[30] S.S. Kannan, A novel hybrid feature selection via symmetrical uncertainty ranking based local memetic search algorithm, Knowledge-Based Systems 23 (2010) 580–585.
[31] R. Kohavi, G.H. John, Wrappers for feature subset selection, Artificial Intelligence 97 (1997) 273–324.
[32] D.S. Koller, Mehran, Toward optimal feature selection, in: 13th International Conference on Machine Learning, Stanford InfoLab, Morgan Kaufmann, Bari, Italy, 1996, pp. 284–292.
[33] K.S. Lee, Z.W. Geem, A new structural optimization method based on the harmony search algorithm, Journal of Computers & Structures 82 (2004) 781–798.
[34] S.W. Lin, S.C. Chen, Parameter determination and feature selection for C4. 5 algorithm using scatter search approach, Soft Computing-A Fusion of Foundations, Methodologies and Applications (2012) 1–13.
[35] M. Mahdavi, M.H. Chehreghani, H. Abolhassani, R. Forsati, Novel meta-heuristic algorithms for clustering web documents, Applied Mathematics and Computation 201 (2008) 441–451.
[36] S. Maldonado, R. Weber, A wrapper method for feature selection using support vector machines, Information Sciences 179 (2009) 2208–2217.
[37] S. Mitra, P.P. Kundu, W. Pedrycz, Feature selection using structural similarity, Information Sciences 198 (2012) 48–61.
[38] K. Nekooei, M.M. Farsangi, H. Nezamabadi-Pour, K.Y. Lee, An improved multi-objective harmony search for optimal placement of DGs in distribution systems, IEEE Transactions on Smart Grid 4 (1) (2013) 557–567.
[39] A. Panchal, Harmony search in therapeutic medical physics, Music-Inspired Harmony Search Algorithm (2009) 189–203.
[40] J.R. Quinlan, Induction of decision trees, Machine Learning 1 (1986) 81–106.
[41] R. Rao, V. Savsani, D. Vakharia, Teaching–learning-based optimization: an optimization method for continuous non-linear large scale problems, Information Sciences 183 (2012) 1–15.
[42] V. Ravikumar Pandi, B.K. Panigrahi, Dynamic economic load dispatch using hybrid swarm intelligence based harmony search algorithm, Expert Systems with Applications 38 (2011) 8509–8514.
[43] R. Ruiz, J.C. Riquelme, J.S. Aguilar-Ruiz, Incremental wrapper-based gene selection from microarray data for cancer classification, Pattern Recognition 39 (2006) 2383–2392.
[44] E.G. Talbi, L. Jourdan, J. Garcia-Nieto, E. Alba, Comparison of population based metaheuristics for feature selection: application to microarray data classification, in: International Conference on Computer Systems and Applications, AICCSA 2008, IEEE/ACS, 2008, pp. 45–52.
[45] M. Tamer Ayvaz, Application of harmony search algorithm to the solution of groundwater management models, Advances in Water Resources 32 (2009) 916–924.
[46] A. Vasebi, M. Fesanghary, S. Bathaee, Combined heat and power economic dispatch by harmony search algorithm, International Journal of Electrical Power & Energy Systems 29 (2007) 713–719.
[47] L. Wang, R. Yang, Y. Xu, Q. Niu, P.M. Pardalos, M. Fei, An improved adaptive binary harmony search algorithm, Information Sciences 232 (2013) 58–87.
[48] H. William, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C:, Cambridge university press, 1998.
[49] P. Yadav, R. Kumar, S. Panda, C. Chang, An intelligent tuned harmony search algorithm for optimisation, Information Sciences 196 (2012) 47–72.
[50] P. Yadav, R. Kumar, S.K. Panda, C.S. Chang, An intelligent tuned harmony search algorithm for optimisation, Information Sciences 196 (2012) 47–72.
[51] X.-S. Yang, Harmony search as a metaheuristic algorithm, in: Music-Inspired Harmony Search Algorithm, Springer, 2009, pp. 1–14.

[52] A.R. Yildiz, A comparative study of population-based optimization algorithms for turning operations, Information Sciences 210 (2012) 81–88.
[53] L. Yu, H. Liu, Redundancy based feature selection for microarray data, in: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2004, pp. 737–742.
[54] L. Yu, H. Liu, Efficient feature selection via analysis of relevance and redundancy, The Journal of Machine Learning Research 5 (2004) 1205–1224.
[55] S.C. Yusta, Different metaheuristic strategies to solve the feature selection problem, Pattern Recognition Letters 30 (2009) 525–534.
[56] Z. Zhu, Y.S. Ong, M. Dash, Wrapper–filter feature selection algorithm using a memetic framework, IEEE Transactions on Systems, Man, and Cybernetics, Part B 37 (2007) 70–76.
[57] Z. Zhu, Y.S. Ong, M. Dash, Markov blanket-embedded genetic algorithm for gene selection, Pattern Recognition 40 (2007) 3236–3248.