Optimal Classification of Epileptic EEG Signals Using Neural Networks and Harmony Search Methods

[1, 2] Xiao-Zhi Gao, [3, 5]Jing Wang, [4] Jarno M. A. Tanskanen, [3]Rongfang Bie, [2] Xiaolei Wang, [3]Ping Guo, [2] Kai Zenger

[1]College of Information Engineering, Shanghai Maritime University, China
[2]Department of Automation and Systems Technology, Aalto University School of Electrical Engineering, Finland
[3]Laboratory of Image Processing and Pattern Recognition, Beijing Normal University, China
[4]Department of Biomedical Engineering, Tampere University of Technology, Finland
[5]School of Foundational Education, Peking University Health Science Center, China
xiao-zhi.gao@aalto.fi, wang_jing@bjmu.edu.cn, tanskanen@ieee.org,
rfbie@bnu.edu.cn, xiaolei.wang@aalto.fi, pguo@ieee.org, kai.zenger@aalto.fi

***Abstract*—In this paper, the Harmony Search (HS)-aided BP neural networks are used for the classification of the epileptic electroencephalogram (EEG) signals. It is well known that the gradient descent-based learning method can result in local optima in the training of BP neural networks, which may significantly affect their approximation performances. Three HS methods, the original version and two new variations recently proposed by the authors of the present paper, are applied here to optimize the weights in the BP neural networks for the classification of the epileptic EEG signals. Simulations have demonstrated that the classification accuracy of the BP neural networks can be remarkably improved by the HS method-based training.**

***Index Terms*—Harmony Search (HS) method, ElectroEncephaloGram (EEG), BP neural networks, optimization, Opposition-Based Learning (OBL), memetic computing, bee foraging algorithm, signal classification.**

I. INTRODUCTION

Epilepsy is a chronic neurological disorder that affects approximately 1% of the world's population, which is characterized by recurrent unprovoked seizures caused by abnormal electrical discharges in the brain. The ElectroEncephaloGram (EEG) is an electrical signal recorded from the scalp or intracranial, and reflects the mass activity of neurons and their interactions. The EEG is widely used by physicians to assist diagnosing many neurological disorders, especially the epilepsy. The detection of epileptic seizures in the EEG signals is very important in the diagnosis of epilepsy. In the past decade, interpretation of the EEG has been limited to only visual inspection by neurophysiologists, individuals trained to qualitatively make a distinction between normal and abnormal EEG. Unfortunately, detection of epilepsy that needs visual inspection of long recordings of the EEG is usually a time-consuming and high-cost process. Therefore, several diagnostic aid approaches for automatically detecting epileptic seizures from the EEG signals have been proposed and studied during the recent years.

Various techniques have been developed in the literature for the detection of epileptic seizures in the EEG [1-14]. All of the seizure detection schemes generally consist of two principal stages. In the first phase, features are extracted from the raw EEG data in the time domain, frequency domain, or time-frequency domain. In the second phase, the features extracted from the EEG are used for training classifiers that differentiate between the normal and epileptic EEG. Actually, numerous classifiers have been proposed and employed, including the Bayesian classifiers [1], Support Vector Machine (SVM), and different kinds of artificial Neural Networks (NNs) [2-6], artificial neuro-fuzzy inference system and dynamic fuzzy NN [7, 8]. In addition to the features for classification, the performance of the epilepsy detection is heavily dependent on the classifiers deployed.

The Harmony Search (HS) method is inspired by the underlying principles of the harmony improvisation [15]. Similar to the Genetic Algorithms (GA) [16], Particle Swarm Optimization (PSO) [17], Differential Evolution (DE) [18] and other computational swarm intelligence systems [19], the HS method is a stochastic search technique. It does not require any prior domain information beforehand, such as the gradient of the objective functions. However, different from many population-based evolutionary approaches, it only utilizes a single search memory to evolve. Thus, the HS method has the interesting characteristics of algorithm simplicity. In the HS, the harmony memory is used to store potential solution candidates, which can considerably reduce the possibility of being trapped into local optima.

The BP neural networks have been extensively employed in such important areas as control, optimization, signal processing, prediction, data classification, etc. Unfortunately, the gradient descent-based learning algorithm used usually results in the local optima of the weights in

doi:10.4304/jsw.9.1.230-239

the BP neural networks. Hence, it is always advantageous to apply some global optimization methods in order to acquire the optimal weights so that the training performances can be enhanced. Motivated by this idea, we apply the HS-based training method of the BP neural networks to classify the epilepsy and normal signals in the paper.

The structure of this paper is as follows: the principles of the HS method together with two modified versions are given in Sections II, III, and IV, respectively. In Section V, the epileptic EEG signal classification using the HS-based BP neural networks is proposed and discussed in details. Section VI demonstrates the numerical simulation results of applying our signal classification scheme. Finally, a few remarks and conclusions are given in Section VII.

II. PRINCIPLES OF HS METHOD

As we know that when musicians compose the harmony, they usually try various possible combinations of the music pitches stored in their memory. This kind of efficient search for a perfect harmony is indeed analogous to the procedure of finding the optimal solutions to many engineering problems. Hence, the HS method is inspired by the underlying principles of the harmony improvisation [15]. Figure 1 shows the flowchart of the basic HS method, in which there are four principal steps involved.

Step 1. Initialize the HS Memory (HM). The initial HM consists of a given number of randomly generated solutions to the optimization problems under consideration. For an n-dimension problem, an HM with the size of HMS can be represented as follows:

$$\mathrm{HM} = \begin{bmatrix} x_1^1, x_2^1, \cdots, x_n^1 \\ x_1^2, x_2^2, \cdots, x_n^2 \\ \vdots \\ x_1^{HMS}, x_2^{HMS}, \cdots, x_n^{HMS} \end{bmatrix}, \tag{1}$$

where n is the dimension of the problem, $\left[x_1^i, x_2^i, \cdots, x_n^i\right]$ ($i = 1,2,\cdots,HMS$) is a solution candidate, and HMS is typically set to be between 50 and 100.

Step 2. Improvise a new solution $\left[x_1', x_2', \cdots, x_n'\right]$ from the HM. Each component of this solution, x_j' , is obtained based on the Harmony Memory Considering Rate (HMCR). The HMCR is defined as the probability of selecting a component from the present HM members, and 1-HMCR is, therefore, the probability of generating it randomly. If x_j' comes from the HM, it is chosen from the j^{th} dimension of a random HM member, and it can be further mutated according to the Pitching Adjust Rate (PAR). The PAR determines the probability of a candidate from the HM to be mutated. Obviously, the improvisation of $\left[x_1', x_2', \cdots, x_n'\right]$ is quite similar to the production of the offspring in the GA [16] with the mutation and crossover operations. However, the GA usually create fresh chromosomes using only one (mutation operator) or two (simple crossover operator) existing ones, while the generation of the new solutions in the HS method makes full use of all the HM members on a probability basis.

Step 3. Update the HM. The new solution from Step 2 is evaluated. If it yields a better fitness than that of the worst member in the HM, it will replace that one. Otherwise, it is eliminated.

Step 4. Repeat Step 2 to Step 3 until a preset termination criterion, e.g., the maximal number of iterations, is met.

The HS method is a random search technique. It does not require any prior domain information beforehand, such as the gradient of the objective functions. However, different from those population-based evolutionary approaches, it only utilizes a single search memory to evolve. Thus, the HS method has the interesting characteristics of computation simplicity.

A few modified HS methods have been developed and reported in the literature. For example, the authors of the present paper study a fusion of the HS and Cultural Algorithm (CA), HS-CA, in which the search knowledge stored in the CA is utilized to guide the mutation direction and size of the HS. This HS-CA is further used to effectively cope with an optimal wind generator design problem [20].

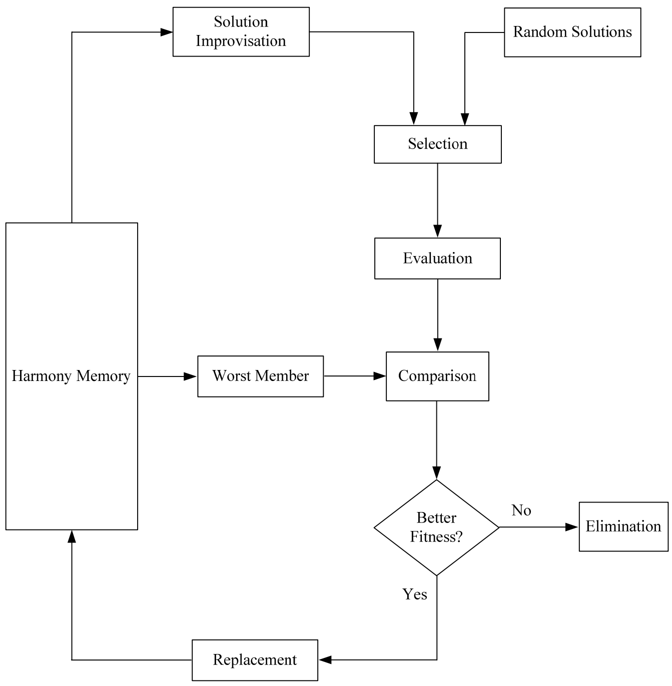


Fig. 1. HS method.

III. A MODIFIED HS METHOD BASED ON OPPOSITION-BASED LEARNING (OBL)

A. Opposition-Based Learning (Obl)

Proposed by Tizhoosh, the OBL is a new approach to machine intelligence, which has been extensively employed in optimization, neural networks training, and reinforcement learning [21]. For dealing with optimization problems, the OBL is based on the utilization of the opposition numbers of the current search directions. More precisely, suppose $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ is a single search point in the n-dimension solution space, and $x_i \in \left[a_i, b_i\right]$,

$i = 1,2,\cdots,n$. Only the continuous variables $\mathbf{x}$ are considered here. The opposition number $\mathbf{x}^* = (x_1^*, x_2^*, \cdots, x_n^*)$ of $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ is defined as:

$$x_i^* = a_i + b_i - x_i \,,\; i = 1,2,\cdots,n \,. \tag{2}$$

The principle of the OBL for optimization is that the search for the optimal solutions should be on the basis of both $\mathbf{x}$ and $\mathbf{x}^*$ as follows:

In every iteration, $\mathbf{x}^*$ is calculated from $\mathbf{x}$, and let $\mathrm{f}(\mathbf{x})$ and $\mathrm{f}(\mathbf{x}^*)$ represent the fitness of $\mathbf{x}$ and $\mathbf{x}^*$, respectively. The iterations proceed with $\mathbf{x}$, if $\mathrm{f}(\mathbf{x}) \geq \mathrm{f}(\mathbf{x}^*)$, otherwise, with $\mathbf{x}^*$. Note that " $\geq$ " here means "better than or equal to with regard to the objective function $\mathrm{f}(\mathbf{x})$ ". An illustrate example of the OBL in the simple one-dimension ($n = 1$) optimization case is given in Fig. 2, where k is the iteration step.

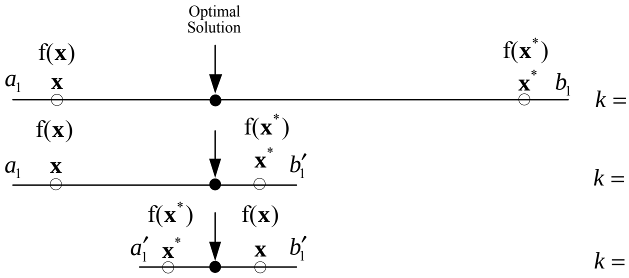


Fig. 2. Opposition-based Learning (OBL) in one-dimension ($n = 1$) optimization.

As Fig. 2 shows, with the growth of OBL iterations, the search interval can be *recursively* shrunk by half by choosing the solution candidate as the better one between $\mathbf{x}$ and $\mathbf{x}^*$. This procedure will ultimately converge, when $\mathbf{X}$ has approached to be close enough with $\mathbf{x}^*$. From these descriptions, it is concluded that the counterpart of $\mathbf{x}$ is utilized in the OBL so that the efficiency of search can be improved. Particularly, the employment of the OBL in the GA, reinforcement learning, DE, and HS method has been investigated in [22]-[24], respectively.

As a matter of fact, hybridization of different optimization schemes has shown to converge drastically faster than the standalone algorithms under certain application circumstances [25]. Inspired by this idea, we propose a hybrid HS method, so called DUAL-HS, by merging the HS and OBL together. A secondary memory is introduced in the DUAL-HS, and the OBL is incorporated for the evolution of this auxiliary memory so that the overall convergence speed can be accelerated [26].

B. A Hybrid HS Method with Dual Memory: DUAL-HS

As previously discussed, the quality of the HM members plays an important role in the convergence of the original HS method. Therefore, we introduce an OBL-based secondary harmony memory for providing alternative solution candidates in our hybrid HS method: DUAL-HS. Figure 3 illustrates the structure and evolution procedure of our DUAL-HS, and its principles can be explained as follows. At the beginning of the DUAL-HS search, a counterpart of the initial harmony memory HM_0, HM_0^*, is obtained using the OBL according to (2). Note that $[a, b]$ used here are the originally given variable ranges of the optimization problems. HM_0^* is first evaluated in the same way as HM_0, and it is then combined with HM_0. Only the top HMS members of $HM_0 + HM_0^*$ are selected as the initial HM to be started with. Actually, the use of HM_0^* results in an improved starting point for our DUAL-HS. In Fig. 3, N is the number of the iterations in the evolution interval of the DUAL-HS. Suppose there are N iterations in the evolution from the harmony memories HM_k to HM_{k+N}. For HM_{k+N}, we first compare the members in HM_{k+N} and HM_k, and merely choose those new members to compose a temporary memory $\overline{HM}_{k+N}$. That is to say, the high-quality fresh HM members created by the HS search during N iterations are employed. Note that the size of $\overline{HM}_{k+N}$ is usually much smaller than that of HM_{k+N}, especially when the DUAL-HS approaches to convergence. Next, the secondary memory in the DUAL-HS is built up by applying the OBL to $\overline{HM}_{k+N}$. It should be emphasized that $[a, b]$ used this time to calculate $\overline{HM}_{k+N}^*$ is based on the present ranges of the members in $\overline{HM}_{k+N}$. In other words, with the shrinkage of the solution candidates in $\overline{HM}_{k+N}$, the convergence of the DUAL-HS can be guaranteed. Similarly, $\overline{HM}_{k+N}^*$ is evaluated and combined together with HM_{k+N}. Similarly to $HM_0 + HM_0^*$, only the best HMS members from the combination of $\overline{HM}_{k+N}^*$ and HM_{k+N} are retained to replace HM_{k+N} so as to continue the search of the DUAL-HS. The above iteration procedure is repeated until a preset termination criterion is satisfied.

It can be observed from these descriptions that the secondary memory in the DUAL-HS, $\overline{HM}_k^*$, acts as an auxiliary storage to the primary harmony memory HM_k. The members generated by the OBL in $\overline{HM}_k^*$ can provide alternative solution candidates for the DUAL-HS to utilize, which may result in a superior convergence property over the regular HS method.

IV. A MEMETIC-INSPIRED HARMONY SEARCH METHOD: m-HS

The memetic computing has recently gained growing interest from different communities. The past decade has witnessed the great successes of applying the memetic algorithms in coping with large-scale, combinatorial, constrained, and multi-objective optimization problems [27]. As a matter of fact, the memetic algorithms represent a wide class of evolutionary computation methods with an inherent local search capability. More precisely, in the memetic algorithms, some local search techniques are incorporated into the meta-heuristics framework, and they operate only at certain cycles of the main-stream computation. The interesting characteristics of the memetic algorithms are that the local search used can efficiently improve the overall quality of the solution candidates, thus accelerating the convergence procedure. Indeed, the memetic computing can provide a useful guideline for researchers to modify the existing evolutionary computation schemes so as to design alternative optimization methods. However, the following important issues have to be carefully addressed when developing a memetic algorithm:

(1) Types of the local search strategies utilized.
(2) Components in the memetic algorithms chosen for the local search.
(3) Frequency of applying the local search.
(4) Depth size (search range) of the local search.

Among all these issues, selecting an appropriate local search method can significantly affect the optimization performance of the memetic algorithms.

As we know that the swarm of bees can simultaneous explore various directions from their nests and find a lot of food sources. However, they are well capable of successfully locating the nearest flower patches with the largest amount of nectar or pollen. Actually, the bee foraging starts with randomly sending out a colony of scout bees for searching for potential flower patches. Based on the information (directions, distances, and qualities of the flower patches) collected by the returned scout bees, the bee colony evaluates the merits of different patches, and then send more scout bees to those more promising areas. During this positive feedback procedure, the bee swarm can gradually find the best food sources to harvest.

The bee foraging algorithm is a kind of popular population-based search method, which is inspired by the aforementioned food foraging behavior of honey bees [28]. In the bee foraging algorithm, the scout bees that return back with the best fitness are first chosen. The sites explored by these bees are considered as the most potential areas, where the optimal solutions may exist. Therefore, the local search is next performed on such sites in order to obtain more promising solution candidates. In the local search, a given number of bees are assigned to the selected sites according to their fitness for neighborhood search. That is to say, the sites with better fitness are going to receive more scouted bees to explore. Among all the scouted bees, only the best bee is selected from each patch as the representative local search result. Nevertheless, at the same time, some bees are also randomly scouted in the whole solution space. With this unique local search capability, the bee foraging algorithm has been proved to be an effective optimization method [29]. In the next section, we propose a memetic HS method, so-called m-HS, by incorporating the bee foraging-based local search strategy into the regular HS method.

A few interesting approaches to merging the HS method and bee foraging algorithm so as to develop novel memetic algorithms have been proposed and studied in the literature. For example, in [30], the authors propose a hybrid HS method, namely HHSABC, by incorporating the Artificial Bee Colony (ABC) algorithm and its variants. The harmony memory of the HHSABC is optimized by the ABC so that both the overall optimization accuracy and convergence rate can be significantly enhanced. The uniform design experiment is employed to verify and demonstrate the superiority of this hybrid HS. Based on the fusion of the HS method, hill climbing, and PSO, another hybrid version, HHSA, is developed [31]. In the HHSA, the local optimizer of the hill climbing and global-best approach of the PSO are complementary to each other, which can strike an appropriate balance between the exploration and exploitation in the search space. Compared with a total of 27 published methods, it is well capable of achieving the best results for most of the data sets for dealing with the popular benchmark problem of the university course timetabling. The hybrid HS method introduced by the authors in [32] combines the HS and bee algorithm together. It actually involves two serial optimization phases of the neighborhood search and global search implemented by the bee algorithm and HS, respectively. The effectiveness of their hybrid technique is examined using 14 real-world data instances of the university course timetabling problem. It can indeed outperform a few famous meta-heuristics methods, such as the variable neighborhood search, Tabu search as well as the original bee algorithm.

We propose and study a new memetic HS algorithm, m-HS, by incorporating the bee foraging like search strategy into the HS method [33]. In our m-HS, the bee foraging-inspired local search technique is applied periodically to improve the quality of the harmony memory. The structure of this m-HS is shown in Fig. 4. More precisely, after *N* steps of the HS evolution, the neighborhood search is performed on only some selected HM members, as illustrated in Fig. 5. Suppose $\mathbf{x}'$ is one of the top *e* members $[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_e]$ chosen from HM_{K+N}. The neighborhood search can produce *ne* offspring from $\mathbf{x}'$ as follows:

$$\begin{aligned} \mathbf{x}'_1 &= \mathbf{x}' \pm rand \times ngh \\ \mathbf{x}'_2 &= \mathbf{x}' \pm rand \times ngh \\ &\vdots \\ \mathbf{x}'_{ne} &= \mathbf{x}' \pm rand \times ngh \end{aligned}, \tag{3}$$

where $rand$ is a random number within $[-1,1]$, and ngh is the local search range applied. $[\mathbf{x}'_1, \mathbf{x}'_2, \cdots, \mathbf{x}'_{ne}]$ are then evaluated, and only the one with the best fitness is retained. Note, to simplify our presentation here, the search performed by the randomly scouted bees is not used, and *ne* is chosen to be fixed. In other words, it is not proportional to the fitness of those selected HM members. This neighborhood search and selection applies to every member of $[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_e]$. Therefore, there are a total of *e* solution candidates with better fitness resulted from the local search, which may replace $[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_e]$ and enter the harmony memory HM_{K+N}. The above evolution procedure is terminated when a preset criterion is met.

It is concluded that the bee foraging like local search approach employed in the m-HS can indeed lead to enhanced solution candidates. That is to say, the quality of the harmony memory is improved by exploring the neighborhood of the top members. Additionally, the choices of parameters of *e*, *ne*, and *ngh* provide desired flexibilities to the m-HS so that its performance can be further fine-tuned. However, the local search used might increase the computational complexity of the HS method.

V. Epileptic EEG Signal Classification with HS-based BP Neural Networks

A. BP Neural Networks

The BP neural networks, also named multi-layer perceptron networks, are an important class of neural networks, due to their simple topology and powerful approximation capability [34]. A simplified BP neural network with only three layers, i.e., input, hidden, and output layer, is illustrated in Fig. 6. There are adjustable weights connecting each two adjacent layers. The back-propagation of approximation error is utilized to train these weights. In general, one iteration of the back-propagation learning algorithm can be written as:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \alpha_k g_k , \tag{4}$$

where $\mathbf{w}_k$ is a vector of the weights at iteration k, α_k is the learning rate, and g_k is the calculated error gradient. It has been proved that a BP neural network with sufficient hidden nodes can approximate any nonlinear function to arbitrary degree of accuracy [35]. Therefore, the BP neural networks are usually regarded as universal function approximators as well as good candidates for classification, modeling, and prediction.

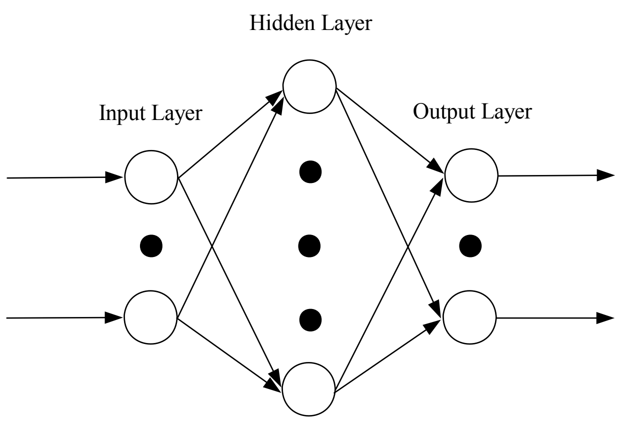


Fig. 6. A BP neural network with three layers.

B. EEG Signal Classification Using BP Neural Networks with HS Method-based Training

In this study, the epileptic seizure detection in the EEG signals can be considered as a classification problem. It includes the data acquisition, feature extraction, and classification steps [36]. With the consideration of the fact that EEG is sparse in Gabor dictionary, feature extraction method described in paper [1] is applied here. The HS-based BP neural networks are used as data classifiers to differentiate the normal EEG from the epilepsy signal [37]. The procedure of our method can be summarized as follows:

Step 1: Divide every EEG signal into segments.
Step 2: Extract feature vector based on the sparse representation [1].
Step 3: Formulate the training and testing signal sets for classification.
Step 4: Obtain the optimal weights of the BP neural networks using the HS method.
Step 5: Examine the classification performance of the BP neural networks using the testing set.

The simulation results of the proposed epileptic EEG signal classification scheme are demonstrated in the following section.

VI. Simulations

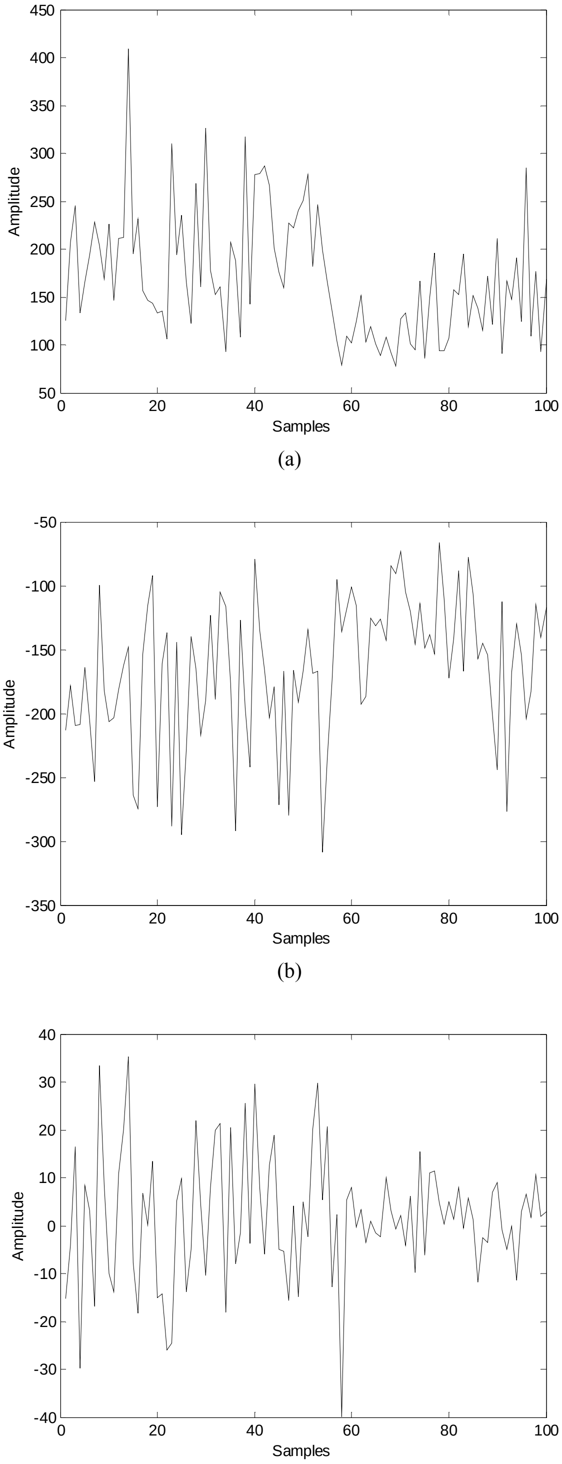
A. Data Sets

The public available data described in [38] is deployed. The complete data set consists of five sets (denoted as Z, O, N, F and S), and each contains 100 single-channel EEG segments. The dimension of the raw data is 4,096. Sets Z and O consist of segments, which are taken from the surface EEG recordings that are carried out on five healthy volunteers using a standardized electrode placement scheme. Volunteers are relaxed in an awake state with eyes open (Z) and eyes closed (O), respectively. Sets N, F and S originate from the EEG archive of presurgical diagnosis. Segments in set F are recorded from the epileptogenic zone, and those in set N from the hippocampal formation of the opposite hemisphere of the brain. While set N and F contain only the activity measured during seizure free intervals, set S only contains the seizure activity. Here, the segments are selected from all the recording sites exhibiting ictal activity.

The above data are digitized at 173.61 samples per second using 12 bit resolution. The band-pass filter settings are 0.53-40 Hz (12dB/oct). The dataset Z includes the signals from normal people and S contains signals with epileptic patient's seizure activity. In this paper, two data sets (Z and S) of the complete data set are used.

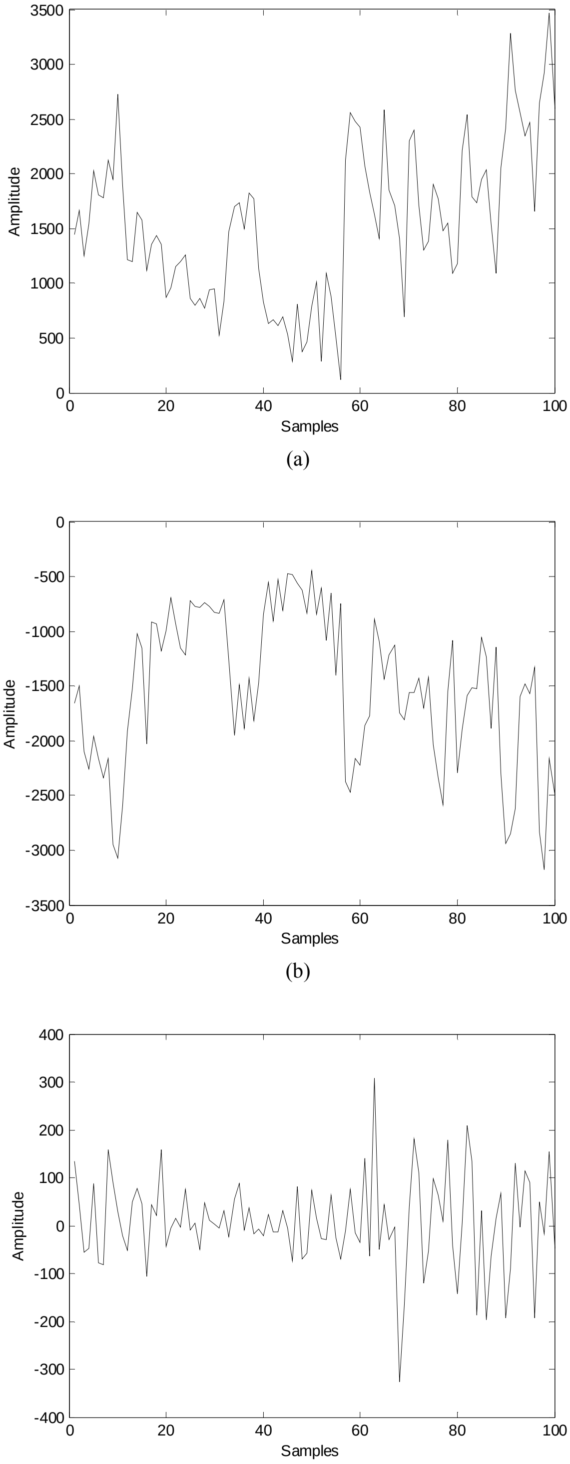
B. Simulation Results

In our simulations, because the dimension of the raw data set is very high, every sample is divided into 17 sub-samples. Thus, the dimension of 4,096 is reduced to 241. The parameters needed in feature extraction based on the sparse representation are the same as in [38]. The normal EEG signals from healthy volunteers and epileptic EEG signals from patients are shown in Figs. 7 and 8, respectively. Note that to simplify our presentation, only the first three sub-samples are given here (in (a), (b), and (c)). The desired classification outputs of the normal and abnormal EEG signals are denoted as -1 and 1, respectively, as illustrated in Fig. 9.



(a) (b) (c)

Fig. 7. Normal EEG signals from healthy people.



(a) (b) (c)

Fig. 8. Epileptic EEG signals from patients.

The structure of the BP neural network used is 3-5-1. That is, there are three, five, and one nodes in the input layer, hidden layer, and output layer, respectively. Therefore, a total of 26 weights/biases need to be optimized. We first compare the optimization performances of the original HS method, DUAL-HS, and m-HS, all of which have 100 HM members, i.e., $HMS = 100$. Their common parameters are given as follows: $\mathrm{HMCR} = 0.8$, and $\mathrm{PAR} = 0.6$. However, in the DUAL-HS, the OBL coefficient, P^{OBL}, is chosen to be $P^{OBL} = 0.35$, and $N = 10$. In the m-HS, $N = 100$, $e = 10$, $ne = 10$. It is also pointed out that all the optimization results presented are based on the average of 100 independent trials.

The classification error of the epileptic EEG signals is used as the fitness for the HS method to optimize. A targeted optimization goal is chosen for the HS, DUAL-HS, and m-HS, which are all terminated after the goals, as given in Table 1, are reached. The iteration steps used by these three algorithms are compared with each other, and the comparison results are presented in Table 1. Obviously, compared with the original HS method, both the DUAL-HS and m-HS use less iterations to achieve the same targeted optimization goals. In other words, the enhanced convergence of these two modified HS methods results in an improved optimization capability. It is also worth pointing out that the m-HS can converge slightly faster than the DUAL-HS. The classification results of the BP neural networks with the HS-based training are further shown in Fig. 10. The testing EEG signals instead of training signals are used this time to examine their generalization capability. It is clearly visible that they are well capable of separating the epileptic EEG signals from the normal EEG signals.

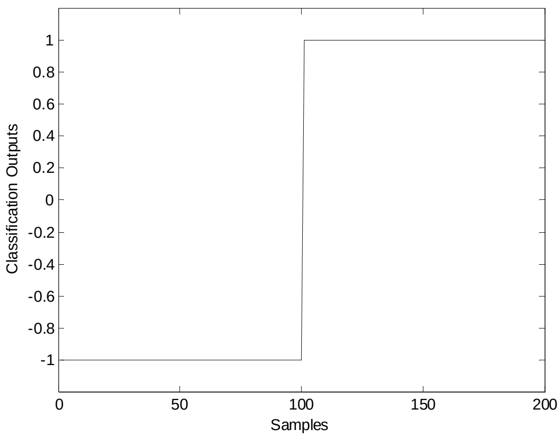


Fig. 9. Desired classification outputs of EEG signals.

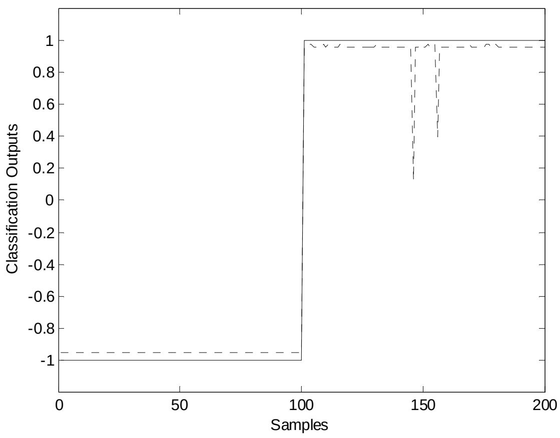


Fig. 10. Classification results of EEG signals using HS-optimized BP neural networks (solid line: desired classification outputs, dotted line: BP neural networks classification outputs).

VII. CONCLUSIONS

In this paper, the HS method together with two variants are used to acquire the optimal weights in the BP neural networks, which are employed as efficient data classifiers for the classification of the EEG signals. The performances of the regular HS method, DUAL-HS, and m-HS are compared in this case-study. Simulations show that both the two modified HS methods can yield an improved optimization accuracy in the training of the BP neural networks. With the proposed HS-based training strategies, the BP neural networks are capable of classifying the epilepsy EEG signals in a satisfactory way. Our future work includes how to apply the HS method in optimizing other kinds of data classifiers so that their performances can be enhanced.

ACKNOWLEDGMENTS

The research work in this paper was supported by the grants from a joint call by the National Natural Science Foundation of China (Project No. 60911130513) and Academy of Finland (Grants 135225, 127299, and 137837). X. Z. Gao's research is also supported by the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning.

REFERENCES

[1] J. Wang and P. Guo, "Epileptic electroencephalogram signal classification based on sparse representation," in *Proceedings of the International Conference on Neural Computation Theory and Applications*, Paris, France, pp. 24-26, October 2011.

[2] P. Guo, J. Wang, X. Z. Gao, and J. Tanskanen, "Epileptic EEG signal classification with marching pursuit based on harmony search method," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Seoul, Korea, October, 2012, pp. 283-288.

[3] V. Nigam and D. Graupe, "A neural-network-based detection of epilepsy," *Neurological Research*, vol. 26, pp. 55-60, 2004.

[4] V. Srinivasan, C. Eswaran, and N. Sriraam, "Artificial neural network based epileptic detection using time-domain and frequency-domain features," *Journal of Medical Systems*, vol. 29, (6), pp. 647-660, 2005.

[5] N. Guler, E. D. Ubeyli, and I. Guler, "Recurrent neural networks employing Lyapunov exponents for EEG signals classification," *Expert Systems with Applications*, vol. 29(3), pp. 506-514, 2005.

[6] A. Subasi, "Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients," *Expert Systems with Applications*, vol. 28(4), pp. 701-711, 2005.

[7] A. Subasi, "Epileptic seizure detection using dynamic wavelet network," *Expert Systems with Applications*, vol. 29, pp. 343-355, 2005.

[8] A. Subasi, "Automatic detection of epileptic seizure using dynamic fuzzy neural networks," *Expert Systems with Applications*, vol. 31(2), pp. 320-326, 2006.

[9] A. Subasi, "Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction," *Computers in Biology and Medicine*, vol. 37(2), pp. 227-244, 2007.

[10] A. Subasi, “EEG signal classification using wavelet feature extraction and a mixture of expert model,” *Expert Systems with Applications*, vol. 32 (4), pp. 1084-1093, 2007.

[11] E. Ubeyli, “Combined neural network model employing wavelet coefficients for EEG signals classification,” *Digital Signal Processing*, vol. 19(2), pp. 297-308, 2009.

[12] H. Ocak, “Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy,” *Expert Systems with Applications*, vol. 36(2), pp. 2027-2036, 2009.

[13] E. D. Ubeyli, “Lyapunov exponents/probabilistic neural networks for analysis of EEG signals,” *Expert Systems with Applications*, vol. 37(2), pp. 985-992, 2010.

[14] L. Guo, D. Rivero, J. Dorado, C. R. Munteanu, and A. Pazos, “Automatic feature extraction using genetic programming: An application to epileptic EEG classification,” *Expert Systems with Applications*, vol. 38(8), pp. 10425-10436, 2011.

[15] Z. W. Geem, J. H. Kim, and G. V. Loganathan, “A new heuristic optimization algorithm: harmony search,” *Simulation*, vol. 76 (2), pp. 60-68, 2001.

[16] R. Poli and W. B. Langdon, *Foundations of Genetic Programming*, Berlin, Germany: Springer-Verlag, 2002.

[17] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of the IEEE International Conference on Neural Networks*, Perth, Australia, December 1995, pp. 1942-1945.

[18] R. Storn and K. Price, “Differential evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces,” *Journal of Global Optimization*, vol. 11, pp. 341-359, 1997.

[19] A. P. Engelbrecht, *Fundamentals of Computational Swarm Intelligence*. West Sussex: John Wiley & Sons Ltd, 2005.

[20] X. Z. Gao, X. Wang, T. Jokinen, S. J. Ovaska, A. Arkkio, and K. Zenger, “A hybrid optimization method for wind generator design,” *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 6, pp. 4347-4373, 2012.

[21] H. R. Tizhoosh, “Opposition-based learning: a new scheme for machine intelligence,” in *Proceedings of the International Conference on Computational Intelligence for Modelling Control and Automation*, Vienna, Austria, November, 2005, pp. 695-701.

[22] S. Rahnamayn, H. R. Tizhoosh, and M. Salama, “A novel population initialization method for accelerating evolutionary algorithms,” *Computers and Mathematics with Applications*, vol. 53, no. 10, pp. 1605-1614, 2007.

[23] H. R. Tizhoosh, “Opposition-based reinforcement learning,” *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 10, no. 5, pp. 578-585, 2006.

[24] S. Rahnamayan, H. R. Tizhoosh, M. M. A. Salama, “Opposition-based differential evolution,” *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 1, pp. 64-79, 2008.

[25] X. Z. Gao, X. Wang, S. J. Ovaska, and K. Zenger, “A hybrid optimization method of harmony search and opposition-based learning,” *Engineering Optimization*, vol. 44, no. 8, pp. 895-914, 2012.

[26] X. Z. Gao, X. Wang, K. Zenger, and Xiaofeng Wang, “A novel harmony search method with dual memory,” in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Seoul, Korea, October 2012, pp. 177-183.

[27] F. Neria and C. Cottab, “Memetic algorithms and memetic computing optimization: A literature review,” *Swarm and Evolutionary Computation*, vol. 2, pp. 1-14, 2012.

[28] D. T. Pham and M. Castellani, “The bees algorithm: modelling foraging behaviour to solve continuous optimization problems,” *Proceedings of Institution of Mechanical Engineers*, Part C, vol. 223, pp. 2919-2938, 2009.

[29] D. Karaboga and B. Basturk, “On the performance of artificial bee colony algorithm,” *Applied Soft Computing*, vol. 8, no. 1, pp. 687-697, 2008.

[30] B. Wu, C. Qian, W. Ni, and S. Fan, “Hybrid harmony search and artificial bee colony algorithm for global optimization problems,” *Computers and Mathematics with Applications*, vol. 64, no. 8, pp. 2621-2634, October 2012.

[31] M. A. Al-Betar, A. T. Khader, and M. Zaman, “University course timetabling using a hybrid harmony search metaheuristic algorithm,” *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, vol. 42, no. 5, pp. 664-681, September 2012.

[32] K. Nguyen, P. Nguyen, and N. Tran, “A hybrid algorithm of harmony search and bees algorithm for a university course timetabling problem,” *International Journal of Computer Science Issues*, vol. 9, no. 1, pp. 12-17, January 2012.

[33] X. Z. Gao, X. Wang, K. Zenger, and Xiaofeng Wang, “A bee foraging-based memetic harmony search method,” in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Seoul, Korea, October 2012, pp. 184-189.

[34] S. Haykin, *Neural Networks, A Comprehensive Foundation*. Second Edition, Upper Saddle River, NJ: Prentice-Hall, 1999.

[35] K. Hornik, M. Stinchcommbe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, pp. 359-366, 1989.

[36] X. Z. Gao, J. Wang, J. Tanskanen, R. Bie, and P. Guo, “BP neural networks with harmony search method-based training for epileptic EEG signal classification,” in *Proceedings of the International Conference on Computational Intelligence and Security*, Guangzhou, China, November 2012, pp. 252-257.

[37] K. C. Zikidis and A. V. Vasilakos, “ASAFES2: A novel, neuro-fuzzy architecture for fuzzy computing, based on functional reasoning,” *Fuzzy Sets and Systems*, vol. 83, no. 1, pp. 63-84, 1996.

[38] R. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. Elger, “Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state,” *Physical Review E*, vol. 64, 2001.

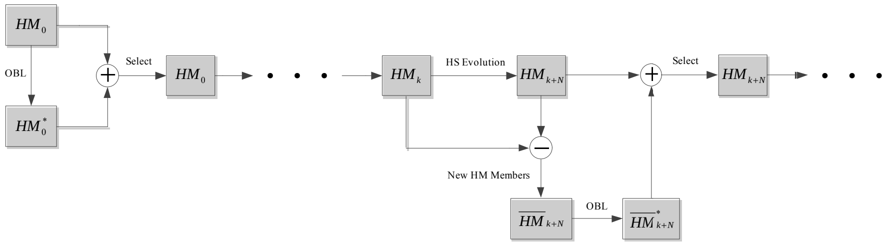


Fig. 3. A new HS method with dual memory: DUAL-HS.

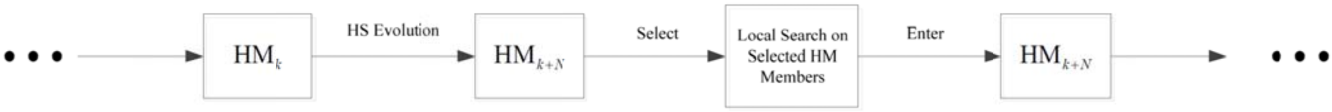


Fig. 4. A memetic HS method: m-HS.

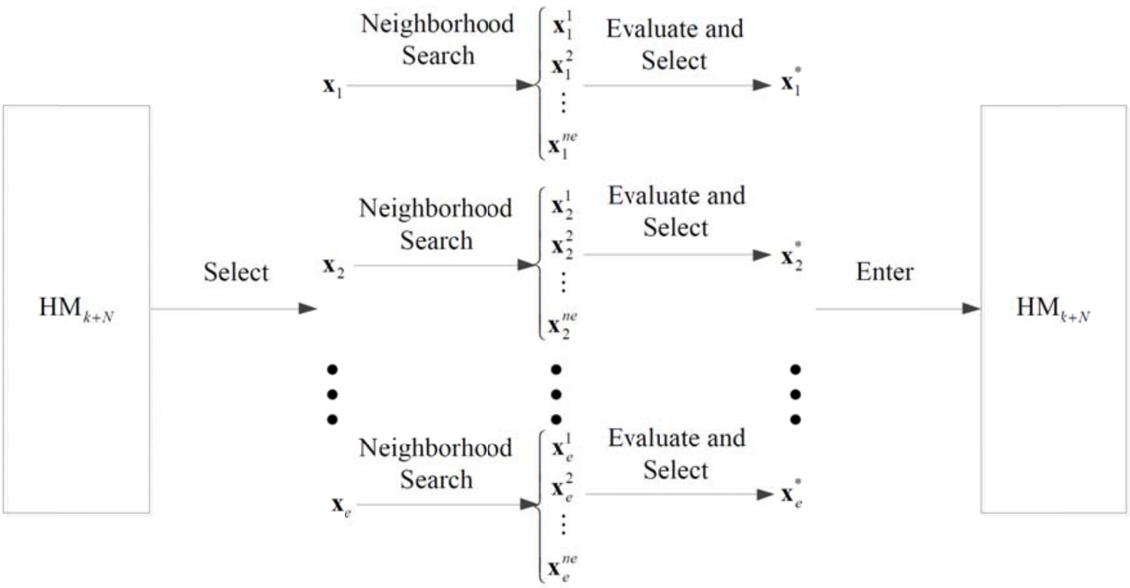


Fig. 5. Local search in m-HS.

TABLE 1.
ITERATIONS USED BY HS, DUAL-HS, AND M-HS IN BP NEURAL NETWORKS TRAINING FOR ACHIEVING OPTIMIZATION GOALS.

Optimization Goals	HS	DUAL-HS	m-HS
0.25	3.5316×10^4	3.0738×10^4	3.0228×10^4
0.5	2.6097×10^4	2.4230×10^4	2.1780×10^4
0.75	2.1423×10^4	1.9011×10^4	1.7347×10^4
1	1.7869×10^4	1.5792×10^4	1.5202×10^4
1.25	1.7359×10^4	1.5669×10^4	1.4014×10^4
1.5	1.3840×10^4	1.2188×10^4	1.1275×10^4

Xiao-Zhi Gao received his D.Sc. (Tech.) degree from the Helsinki University of Technology (HUT), Finland in 1999. Since 2004, he has been appointed as a Docent of Soft Computing Methods and Applications at the Aalto University (formerly HUT). His current research interests are nature-inspired computing methods with applications in optimization, prediction, data mining, signal processing, and control. He has published more than 260 technical papers in refereed journals and international conferences.

Jing Wang received the M.Sc. degree in computer software theory and applications from the Beijing Normal University, China in 2004. She is currently a Ph.D. student at the same university. Her research interests include computational intelligence methods with applications in medical signal processing.

Jarno M. A. Tanskanen received his D.Sc. (Tech.) degree from the Helsinki University of Technology (HUT), Finland in 2000. He is now working as a Senior Research Fellow at the Department of Electronics and Communications Engineering, Tampere University of Technology (TUT), Finland. His main research interests are biomedical and electrophysiological signal processing and analysis.

Rongfang Bie is the leader of IReG Group on the Internet of Things (IoT) at the College of Information Science and Technology of the Beijing Normal University (BNU), China. She received her M.Sc. degree in 1993 and Ph.D. degree in 1996 in Mathematics from the Beijing Normal University. She was with the Computer Laboratory at the University of Cambridge as a visiting faculty from March 2003 for one year. She is the author or co-author of more than 40 academic papers. Her current research aims to establish theories and models on database management model, knowledge representation, knowledge automatic acquisition, and knowledge sharing for the IoT, to develop the corresponding platforms, and to develop some typical applications on IoT education, health monitoring methods for civilians, and mechanical structures using the proposed theoretical models and IoT related technologies, such as sensing systems.

Xiaolei Wang received her D.Sc. (Tech.) degree from the Helsinki University of Technology (HUT), Finland in 2009. She is now working as a Postdoc Research Fellow at the Aalto University School of Electrical Engineering (formerly HUT). Her research interests are hybrid computational intelligence methods for optimization.

Ping Guo is an IEEE Senior Member and Professor at the College of Information Science and Technology of the Beijing Normal University (BNU), China. From 1980 to 1983, he studied at the Physics Department of the Peking University for M.Sc. degree, majoring in Optics. From 1997 to 2001, he studied at the Computer Science and Engineering Department of the Chinese University of Hong Kong for his Ph.D. degree. From 1993 to 1994, he was a visiting faculty at Computer Science and Engineering Department of Wright State University in the United States. He is a PC member of several international conferences, such as the ISNN and ICIC. As an author or co-author, he has published over 200 academic papers.

Kai Zenger received his M.Sc., L.Sc. and D.Sc. degrees in electrical engineering, computer technology, and automation and systems technology in 1986, 1992, and 2003, respectively. From 1983 to 1989, he worked as an automation engineer in MKT Finland and HT Automation. Since 1989, he had several positions related to teaching and research in the Control Engineering Laboratory at the Helsinki University of Technology, Finland. Currently, he is working as a Professor of Automation Technology in the Aalto University School of Electrical Engineering, Finland. Prof. Zenger's main research areas are Control Engineering and System Theory with applications in chemical process engineering, power electronics and mechanical engineering. He has specialized in the research of time-varying linear systems, periodic systems, and adaptive and robust control methods.