Hybridization of harmonic search algorithm in training radial basis function with dynamic decay adjustment for condition monitoring

Hue Yee Chong1 · Shing Chiang Tan2 · Hwa Jen Yap1

Accepted: 9 June 2021 / Published online: 19 June 2021 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
In recent decades, hybridization of superior attributes of few algorithms was proposed to aid in covering more areas of complex application as well as improves performance. Condition monitoring is a major component in predictive maintenance which monitors the condition and identifies significant changes in the machinery parameter to perform early detection and prevent equipment defects that could cause unplanned downtime or incur unnecessary expenditures. An effective condition monitoring model is helpful to reduce the frequency of unexpected breakdown incidents and thus, facilitates in maintenance. ANN has shown effective in various condition monitoring and fault detection applications. ANN is popular due to its capability of identifying the complex nonlinear relationships among features in a large dataset and hence, it can perform with an accurate prediction. However, a drawback is that the performance of ANN is sensitive to the parameters (i.e., number of hidden neurons and the initial values of connection weights) in its architecture where the settings of these parameters are subject to tuning on a trial-and-error basis. Hence, a wide range of studies have been focused on determining the optimal weight values of ANN models and the number of hidden neurons. In this research work, the motivation is to develop an autonomous learning model based on the hybridization of an adaptive ANN and a metaheuristic algorithm for optimizing ANN parameters so that the network could perform learning and adaptation in a more flexible way and handle condition classification tasks more accurately in industries, such as in power systems. This paper presents an intelligent system integrating a Radial Basis Function Network with Dynamic Decay Adjustment (RBFN-DDA) with a Harmony Search (HS) to perform condition monitoring in industrial processes. RBFN-DDA performs incremental learning wherein its structure expands by adding new hidden units to include new information. As such, its training can reach stability in a shorter time compared to the gradient-descent based methods. To achieve optimal RBFN-DDA performance, HS is proposed to optimize the center and the width of each hidden unit in a trained RBFN. By integrating with the HS algorithm, the proposed metaheuristic neural network (RBFN-DDA-HS) can optimize the RBFN-DDA parameters and improve classification performances from the original RBFN-DDA by 2.2% up to 22.5% in two benchmarks datasets, which are numerical records from a bearing and steel plate system and a condition-monitoring system in a power plant (i.e., the circulating water (CW) system). The results also show that the proposed RBFN-DDA-HS is compatible, if not better than, the classification performances of other state-of-the-art machine learning methods.

Keywords Metaheuristic · Radial Basis Function Network · Condition monitoring · Harmony search · Optimization

1 Introduction
Maintenance involves carrying out all technical and associated administrative activities to keep all components in an operational system to perform their function properly (Stephens 2010). If any equipment breakdown occurs in a well-maintained operational system, minor problems can be detected and corrected by conducting the short daily inspection, cleaning, and lubricating activities. An
effectively maintenance action requires company-wide participation and support by every personnel in order to make plant and equipment more reliable (Waeyenbergh and Pintelon 2002).

Basically, maintenance can be divided into two types, preventive maintenance and corrective maintenance (Blanchard 2004). Preventive maintenance includes all planned maintenance actions e.g., periodic inspection, condition monitoring, etc. that are implemented to avoid the equipment from breaking down unexpectedly. Corrective maintenance includes all planned and unplanned maintenance actions to rectify failures before restoring equipment back to its operational condition. Condition monitoring is a major component in preventive maintenance which monitors the condition and the significant changes in the machinery parameter to perform early detection and eliminate equipment defects that could cause unplanned downtime or incur unnecessary expenditures (Mohanty 2014). It is essential to develop efficient condition monitoring techniques for (1) quickly identifying the faulty components before breakdown. (2) Reducing costly repairs caused by unexpected failure. (3) Optimizing the scheduling of preventive and predictive maintenance operations without routine inspections, which may require periodic shutdowns in a plant. Condition monitoring consists of two sequential processes: feature extraction and condition classification (Marwala and Vilakazi 2007). Feature extraction requires the use of signal processing and/or post-processing techniques; it is a mapping process from the signal space to the feature space. The condition classification is a process of classifying the obtained features into different categories. The traditional approach of condition monitoring relies on human expertise to relate the extracted features to the faults, which is usually time-consuming and unreliable, particularly when multiple features are referred for fault diagnostics and when the data are affected by noises (Gertler 1998). Fault diagnosis coupled with artificial intelligence (AI) can potentially overcome the shortcomings of traditional signal processing techniques by incorporating the human-like thought abilities such as learning, reasoning, and self-correction (Kok, 2009). The widely used AI tools in condition monitoring include artificial neural network (ANN), fuzzy logic system, support vector machine, extreme learning machine, etc. (Ali 2018). Literature shows that ANNs with learning capabilities are useful models for tackling condition-based maintenance problems (Tallam et al. 2003). They learn from data samples without a requirement for building an exact mathematical model. However, the performance of ANNs is highly dependent on their parameter settings. For this concern, numerous global optimization methods are introduced and utilized to train ANNs to achieve a better network performance. Many of these global optimization methods are metaheuristic algorithms, which initiate a search process to explore in the search space to obtain near optimal solutions. The metaheuristic algorithms are mechanisms that imitate certain strategies inspired from nature, social behavior, physical laws, etc. Some well-known metaheuristic algorithms include Genetic Algorithm (GA) (Holland 1975), Particle Swarm Optimization (PSO) (Sheng et al. 2011), Artificial Bee Colony Algorithm (ABC) (Karaboga 2005) and harmony search (HS) (Geem et al. 2001). Other new metaheuristic algorithms are also introduced in the literature, for instances, Aquila Optimizer (AO) (Abualigah 2021a) and Arithmetic Optimization Algorithm (AOA) (Abualigah 2021b). Metaheuristic algorithms have been applied in a wide range of application domains such as medical (Ahmad et al. 2017), engineering (Bhoskar 2015), finance (Daliri 2020), document clustering (Abualigah et al. 2018a, b) and due to their high efficiency and flexibility characteristic. Literature showed that many metaheuristic algorithms have been successfully used to optimize the ANN’s design and the parameters (Yao 1993). Many reviews on ANN-based condition monitoring have been presented (Singh 2003; Lei 2014). Notably, the application of metaheuristic ANNs to condition monitoring is still relatively few. An effective metaheuristic ANN-based condition monitoring model is helpful to reduce the frequency of unexpected breakdown incidents and thus, facilitate in maintenance. Thus, this motivates a research on developing an autonomous learning model based on the hybridization of an adaptive ANN and a metaheuristic algorithm for handling condition classification tasks in industries, such as in power systems. In this paper, a hybrid model of an incremental radial basis ANN and metaheuristic algorithm is developed. A metaheuristic algorithm (i.e., harmony search (HS) algorithm) is proposed to improve the learning parameters of the ANN. The research work is aimed at developing a metaheuristic ANN for monitoring the operating states of a system more accurately in a power plant.

The paper’s organization is as follows: Sect. 2 presents a review of condition monitoring and fault diagnosis techniques. In addition, the state-of-the-art of ANN and metaheuristic ANN models used for performing condition monitoring and fault detection will be described. Section 3 describes two machine learning components [namely, Radial Basis Function Network with Dynamic Decay Adjustment (RBFN-DDA), Harmony Search (HS)] before describing the proposed RBFN-DDA-HS in details. The effectiveness of RBFN-DDA-HS in condition monitoring and fault detection is evaluated using two benchmark datasets (namely, bearing fault and steel plate fault datasets) and a real dataset that is collected from a circulating water system in a power plant. Section 4 presents and compares the results of RBFN-DDA-HS and other machine
learning methods in the mentioned condition monitoring and fault detection tasks. Finally, Sect. 5 draws concluding remarks and suggestions for further work.

2 Related work

In this section, we first present an overview of the condition monitoring and fault detection approaches, which are the model-based and the data-driven approaches. The strengths and limitations of each approach are briefly explained. The proposed method in this paper is developed from a data-driven, artificial intelligence (AI) approach. In particular, the proposed method is an example of artificial neural network (ANN) that is integrated with a metaheuristic algorithm. Therefore, the focus of the second section of literature review is on the state-of-the-art of ANN and metaheuristic ANN in condition monitoring and fault detection.

2.1 Condition monitoring and fault detection techniques

In general, condition monitoring and fault detection techniques can be broadly classified into model-based approach (Isermann 1997) and data-driven approach (Patton et al. 2013), which are shown in Fig. 1. Many review articles are available to survey the application of techniques from data-driven and model-based approaches in process monitoring and fault detection (Tidriri 2016; Miljković 2011; Ding 2011). The model-based approach involves a modeling process, which is based on fundamental understanding about the physics of the process. The model can be either of a qualitative or a quantitative type. Qualitative modeling (Venkatasubramanian et al. 2003a) relies on the use of qualitative functions centered around different units in a process to express the relationship between the input and output of the system while quantitative modeling (Venkatasubramanian 2003b) is built by the use of mathematical function. The performance and reliability of the diagnostic system are assessed by the accuracy of the model. Thus, these approaches are usually applied only when an accurate analytical model is available. However, an accurate mathematical model is usually difficult to derive to represent complex mechanical systems, especially when the machines operate in uncertain and noisy environments (Tan et al. 2007). Besides, model-based approaches also rely on human expertise to relate the extracted features to the faults, which is usually time-consuming and unreliable, particularly when multiple features are utilized for fault diagnostics and when the data are affected by noises (Gertler 1998). Data-driven approach overcomes the above-mentioned limitation of the model-based approaches and is more viable to monitor industrial processes that are inherently automated. Data-driven approach requires for real data obtained from a data acquisition system to monitor or forecast the process behavior of the system under monitoring. The commonly used data-driven methods can be divided into statistical (Yin 2014) and AI approaches (Ali 2018). Methods under the statistical approach perform condition monitoring/fault detection by identifying the relationships among the data and these relationships are referred for making classification and prediction in the future. The commonly used statistical methods such as control chart, principle component analysis and partial least squares are easy to build and they can provide fast detection on abnormal situations (Venkatasubramanian 2003c). To further improve the diagnosis performance, fault diagnosis coupled with AI with the human-like thought abilities such as learning, reasoning and self-correction (Kok et al. 2009) are extensively developed. Comprehensive reviews on the use of AI techniques for fault detection and condition monitoring on mechanical components such as induction machine (Singh 2003) and planetary gearboxes (Lei 2014) have been presented. The reviews show AI is useful to enhance the precision and accuracy of the condition monitoring/fault detection system. AI condition monitoring/fault detection systems have been developed by using, for examples, ANN, fuzzy logic system, support vector machine, extreme learning machine and etc. (Ali 2018). Among them, ANN is popular due to its capability of identifying the complex nonlinear relationships among features in a large dataset and hence, it can perform with an accurate prediction.

2.2 ANN and metaheuristic ANN methods for condition monitoring and fault detection

ANNs have been widely applied as condition classification tool in combination with signal processing techniques for feature extraction. For example, bearing has been identified as one of the main fault components in rotating machinery.
From the literature, extensive research work on bearing fault detection and diagnosis have been reported (Cerrada 2018). A single hidden layer feed forward ANN was proposed to perform health diagnostics of ball bearings in direct-drive motors (Cocconcelli et al. 2011). The feed forward ANN overcomes the problem of continuous change in rotational speed of the motor and identifies the health of ball bearing in different speed applications. A singular spectrum analysis (SSA) is integrated with an ANN based classifier to perform fault detection and diagnosis (Muruganatham 2013). Feature extraction by using existing time series methods could be affected by noise and sample sizes, but the proposed method could overcome the drawback of the existing time series methods and achieve a high classification rate. A feature extraction method based on empirical mode decomposition (EMD) was appended to an ANN based classifier to categorize bearing defects (Ben Ali 2015). The authors showed that the ANN can perform degradation assessment on the bearing condition automatically without human intervention. Kankar et al. (2012) compared the effectiveness of ANN and support vector machine (SVM) in detecting faults in a rotor bearing system. They used statistical methods to extract features that were input to both ANN and SVM classifiers. Results showed that ANN classifier was able to achieve higher accuracy than SVM. An adaptive algorithm based on wavelet transform was applied to perform feature extraction and extracted features were input to ANN and KNN, where ANN showed more effective in classifying bearing faults (Gunerkar et al. 2019). Effective condition monitoring technique is useful to enhance a machining process. Siddhpura and Paurobally (2013) presented a review of different classifiers (such as ANN, fuzzy logic, neurofuzzy, hidden markov model and SVM) used for predicting tool wear in metal cutting such. Among them, ANN classifier was the most frequently used classifier due to its noise tolerant and high fault adaptability characteristics. A hybrid machine-learning classifier between Gaussian classifier and an adaptive resonance theory (ART) neural network was proposed to implement online tool wear monitoring (Wang et al. 2013). This hybrid machine-learning classifier could memorize new knowledge in online training and achieves high classification accuracy. The performance of an ANN classifier was demonstrated to outperform SVMand KNN (Hesser and Markert 2019) in monitoring the condition of tool wear of a retrofitted CNC milling machine.

Although ANN has shown effective in various condition monitoring and fault detection applications, a drawback is that the performance of ANN is highly relied on the parameters in its architecture such as the number of hidden nodes and connection weights. A survey on different combination of ANNs and evolutionary algorithms (EAs) was first made by Xin (1999). Yao explained that connection weights, learning rules and architecture of ANNs could be evolved using EAs. After a decade, Azzini and Tettamanzi (2011) extended the Yao’s survey by updating the most recent related literatures. Ding et al. (2013) reviewed and presented some of the problems associated to the integration of ANNs with EAs. In machine learning community, many metaheuristic algorithms have been introduced to integrate with ANNs to enhance learning for achieving a better generalization performance (Devikanniga et al. 2019). On the other hand, a number of comprehensive reviews on the application of ANNs in condition monitoring and fault detection are available in the literature; however, the work on integrating ANNs with metaheuristic algorithms for performing condition monitoring and fault detection is relatively few. Genetic algorithm (GA) is one of the popular evolutionary algorithms applied to perform fault detection. A Dual GA loop has been applied to optimize the structure and the connection weights of a three-layered feed-forward ANN for fault detection in power systems (Bi et al. 2000). Besides, GA has been used together with Levenberg–Marquardt (LM) algorithm to train an ANN in an electrical machine fault detection application. The proposed hybrid method between GA and LM overcame the weak local search ability of GA (Zaiping et al. 2008). Besides, the performance of various optimization methods including ANN, fuzzy, GA and Ant Colony Optimization (ACO) and their hybrid models is compared when predicting accidents caused by repair and maintenance in oil refineries (Zaranezhad et al. 2019). Among those computing models, ANN-GA outperformed the rest. Three methods including multi-layer perceptron (MLP), neural network, radial basis function (RBF) neural network, and KNN were hybridized with GA for detecting gear faults (Lei 2010). Result showed that the performance of hybrid method was better than the single machine learning methods. Apart from evolutionary algorithm, swarm based metaheuristic algorithms are also always applied to train an ANN in fault detection applications. For example, particle swarm optimization (PSO) was used to hybridize with ANN in predicting drilling fluid density (Ahmadi 2018). Result showed that PSO-ANN could achieve a better fault detection performance if compared to fuzzy inference system (FIS) and GA-FIS. PSO was also proposed to optimize the weights and threshold parameters of a back propagation neural network (BPNN) that was applied for drilling fluid density prediction (Zhou et al. 2016). In a fault detection of multilevel inverter, GA and PSO were applied separately to train an ANN in order to optimize its connection weights. PSO tended to predict more accurately and faster than GA (Sivakumar and Parvathi 2014). On the other hand, four metaheuristic algorithms including PSO,
GA, Tabu Search (TS) and Cuckoo Search (CS) have been applied to optimize the weight and the bias values of an ANN in a multilevel inverter fault diagnosis application (Manjunath and Kusagur 2018). A modified evolutionary PSO with time varying acceleration coefficient (MEPSO-TVAC) was proposed to optimize regression coefficient of the ANN before performing transformer fault detection (Illias et al. 2016). Fault diagnosis performance could be enhanced on the use of a feature extraction method in addition to a decision-making tool. Kernel linear discriminant analysis (KLDA) was proposed to extract the optimal features from a fault dataset of analog circuits, and PSO was applied to tune the ANN parameters and structures (Xiao and Feng 2012). ACO was proposed to optimize the weights of an ANN in a rolling bearing fault detection application (Shi 2010). An improved Gravitational Search Algorithm (GSA) was to optimize the weights and the bias settings of an ANN trained by back propagation algorithm in machine vibration (Do 2017). The performance of metaheuristic algorithm could be affected by the exploitation and exploration characteristics of the algorithm. Ideally, a metaheuristic algorithm should achieve a good balance of search between exploration and exploitation. One of the ways to tackle this issue is by hybridizing two or more metaheuristic algorithms in order to achieve such balance between two modes of search. A novel fault diagnosis model for sensor nodes was proposed by optimizing the weights of a feed-forward ANN using a hybrid metaheuristic algorithm of GSA and PSO (Khilar and Dash 2020).

Table 1 shows a summary of aforementioned research work about the state-of-the-art of metaheuristic ANNs (Bi et al. 2000; Zaiping et al. 2008; Zaranezhad et al. 2019; Lei 2010; Ahmadi 2018; Zhou et al. 2016; Sivakumar and Parvathi 2014; Manjunath and Kusagur 2018; Illias et al. 2016; Xiao and Feng 2012; Shi 2010; Do 2017; Khilar and Dash 2020) in condition monitoring and fault detection. All of these metaheuristic ANNs optimize the ANN’s parameters (i.e., centers/widths, connection weights the number of hidden nodes or structure) from randomly-initialized settings. On the contrary, in this paper, we introduce a hybrid model of a radial basis ANN and HS algorithm (RBFN-DDA-HS) that combines the advantage of incremental learning with the RBFN-DDA, and search and adaptation with the HS. RBFN-DDA performs incremental learning wherein its structure expands by adding new hidden units to include new information. As such, its training can reach stability in a shorter time compared to the gradient-descent based methods to learn information (or knowledgebase) directly from a dataset. In this case, the knowledgebase represents solutions to a problem at hand and is not initialized randomly (that is implemented in other metaheuristic ANNs). To achieve optimal RBFN-DDA performance, HS is proposed to further optimize the knowledgebase that is represented in terms of a set center and width from all hidden units in a trained RBFN.

3 Methods

3.1 RBF-D-DDA

Learning in RBFs is governed by many parameter settings such as the number of nodes in the hidden layers, the type of activation functions, the center and width of a neuron. Normally, in a fixed architecture, the number of nodes in the hidden layer must be determined or fixed before training process begins. It is important to determine the optimized number of neuron since redundant hidden neuron can result in poor generalization and overlearning situation. On the hand, insufficient number of hidden neuron can result in inadequate of learning information from data (Liu 2004). To enhance the performance of the network by determining optimized number of hidden neuron, the Dynamic Decay Adjustment (DDA) algorithm is applied. Two unique characteristics of DDA algorithm are the constructive nature of probabilistic extension of restricted coulomb energy algorithm (P-RCE) (Hudak 1992) and the independent adaptation of each prototype by referring to a decay factor. An area of conflict is defined by using two thresholds including positive thresholds ($\theta^+$) and negative thresholds ($\theta^-$) during training as shown in Fig. 2 (Berthold and Diamond 1998). $\theta^+$ determines the lower bound of activation value for the training patterns of correct class so that no new neuron is committed while $\theta^-$ represents the upper bound of activation value for the neuron tolerating with neurons of conflicting class. Thus, the area of conflict refers to those sections where neither matching nor conflicting training patterns are allowed to reside. This algorithm constructs an RBFN structure dynamically during training process by an idea that new hidden neurons will only be inserted to the hidden layer when necessary. With this growing structure during training process, to the RBFN can reach stability in a shorter time compared to ANNs trained with gradient-descent based methods. Besides, the DDA algorithm computes the decay factor based on neighbors’ information.

An RBFN trained with the DDA Algorithm (RBFN-DDA) consists of three layers, namely, the input, hidden and output layers (Fig. 3). The input layer corresponding to network input features which represents the dimensionality of the input space and this layer is fully connected to the hidden layer. Next, the hidden layer of RBFN-DDA network consists of radial Gaussian units as an activation function and these units will only be added during training.
when necessary. Each hidden RBF unit is connected to only one output unit. The number of the output unit is determined by the number of the possible class. The output unit with the highest activation determines the class value. During the DDA-Algorithm training, all RBFs weights are first set to zero to avoid any duplication of the information about training samples. Then, all training samples are presented to the network. If a new training sample is classified correctly, then the weight of the RBF unit of correct class is increased by one. On the other hand, if the training sample is misclassified, a new RBF unit with an initial weight of one is introduced. The hidden RBF unit is assigned with a reference vector that is the same as the new training instance. The last step of the algorithm reduces the radii of all conflicting RBF units in a width shrinking process. The DDA algorithm defines an area of conflict by referring on two thresholds \( h^+ \) and negative thresholds \( h^- \). In this study, \( h^+ \) and \( h^- \) are

![Fig. 2 The DDA algorithm with two thresholds](image)

| Model                  | Method                                                                 | References                  |
|------------------------|------------------------------------------------------------------------|-----------------------------|
| ANN—dual GA loop       | Dual GA loops are applied to optimize both ANN’s structure and the connection weights. The number of hidden neuron in ANN is set manually and the connection weights are randomly initialized before GA optimization | Bi et al. (2000)            |
| ANN-GA-LM              | GA is proposed to optimize the number of hidden neuron with the initial value is set with experience. Then, LM is used to further optimize the ANN’s connection weights from random initialization | Zaiping et al. (2008)       |
| ANN-GA/ANN-ACO         | GA/ACO is used to optimize the weights of neurons on the hidden layer from random initialization | Zaranezhad et al. (2019)    |
| Multiple classifier—GA | GA is applied to optimize the weight of multiple classifiers from random initialization | Lei (2010)                  |
| PSO—ANN/GA-FIS         | GA/PSO is applied to optimize the relevant parameters of a regarded FIS/ANN from random initialization | Ahmadi (2018)               |
| PSO-BPNN               | PSO is used to optimize network weights, threshold parameters with the number of hidden neuron is set manually and the connection weights are randomly initialized before PSO optimization | Zhou et al. (2016)          |
| ANN-PSO/ANN-GA         | PSO/GA is applied to optimize the connection weights of ANN from random initialization | Sivakumar and Parvathi (2014) |
| ANN—PSO/ANN—GA/ANN—TS/ANN—CS | PSO/GA/TS/CS is used to optimize the weight and the bias values of an ANN from random initialization | Manjunath and Kusagur (2018) |
| MEPSO-TVAC-ANN         | MEPSO-TVAC is proposed to optimize the learning rate (LR) and momentum constant (MC) of ANN. The number of hidden neuron in ANN is set manually | Illias et al. (2016)        |
| ANN—PSO               | PSO is used to optimize the number of hidden neurons, weights and biases of network. The number of hidden neuron in ANN is set manually and the weights and biases of network are randomly initialized before PSO optimization | Xiao and Feng (2012)        |
| ANN-AOC               | ACO is proposed to optimize the weights of ANN from random initialization | Shi (2010)                  |
| ANN-Improved GSA       | Improved GSA is proposed to optimize the weights and the bias settings of ANN from random initialization | Do (2017)                  |
| ANN-GSA-PSO           | Hybrid metaheuristic algorithm of GSA and PSO is proposed to optimize the weight of ANN from random initialization | Khilar and Dash (2020)      |

![Fig. 3 The structure of RBFN-DDA (Robnik-Šikonja 2016)](image)
set to 0.4 and 0.2. To commit a new RBF unit, the activation of existing RBFs of the correct class must not be above positive thresholds \( \theta^+ \) and during shrinking the width of an RBF unit of a conflicting class must not be above negative thresholds \( \theta^- \). Figure 4 shows the RBFN-DDA training in a single epoch. Usually, the training process of RBFN-DDA involves several epochs before completion. However, in this study, the center and radius of neuron shall be optimized using EA, thus the ANN training of RBFN-DDA is run for one epoch only.

### 3.2 Harmony search (HS)

An interesting musical inspired algorithm called HS algorithm was first introduced to solve optimization problems (Geem et al. 2001). Musicians always search for an ideal state of harmony in their performances. Music improvisation is executed iteratively to obtain optimal harmony by considering three rules (Geem et al. 2001): (1) memory consideration—a new music is improvised from the existing harmony memory (HM); (2) pitch adjustment—a new music is improvised by slightly adjusting the pitch; (3) randomization—a new music is improvised on a random basis. HS mimics a musician’s behavior of searching for an ideal state of harmony for finding the best solution to an optimization problem. This algorithm is widely applied for solving optimization problems and training of ANN. In Geem et al. (2002), HS was employed to solve several water resources problems and it gave better near-optimal solutions with faster convergence in most problems. HS was also employed to optimize the parameters of a multi-layer feed-forward ANN such as the number of neurons in the hidden layer, the learning rate, and the momentum rate in predicting the AC power from a photovoltaic system (Kassim et al. 2014). HS was applied to determine the optimal initial weights of the ANN in the prediction of the stability number of breakwater armor stones and it was more efficient than the conventional ANN model (Lee et al. 2016). The HS algorithm was proposed in predicting the best structure for ANN in the financial fraud detection application and it is able to achieve a high accuracy (Daliri 2020). A study showed that HS could achieve a better overall recognition performance than BP or GA in a feed-forward ANN (Kattan and Abdullah 2013). The ability to explore the search space of HS is highly depended on the pitch adjustment and randomization. Pitch adjustment warrants that new solution is improved from existing good solutions, whereas randomization provides a search for new solution within the search space. Harmony memory size (HMS), harmony memory considering rate (HMCR), pitch adjusting rate (PAR), and the termination criterion shall be specified. The ability of HS in finding optimal solution from a search space is highly relied on a harmony memory acceptance rate. The following HS parameter ranges are recommended to produce an optimal solution (Lee 2005) 0.70–0.95 for HMCR, 0.20–0.50 for PAR, and 10–50 for HMS. The procedure of HS algorithm is shown in Fig. 5.

### 3.3 The proposed hybrid RBFN-DDA-HS

In literature, many training algorithms were proposed to train RBFN, including gradient descent algorithm (GD)
Karayiannis (1999) and Kalman filtering (KF) (Simon 2002). However, these methods exhibit poor convergence and time consuming before obtaining optimal solution (Kurban and Beşdok 2009). Evolutionary algorithm such as GA has been applied to optimize RBFN (Barreto et al. 2002). GA could perform a robust search algorithm to avoid itself from being stuck in local minima. However, its search algorithm is computational expensive to find optimal solution (Hamadneh, 2012). On the other hand, various global optimization algorithms include particle swarm optimization (PSO) algorithm (Liu 2004), artificial immune system (AIS) algorithm (Castro and Zuben 2001), differential evolution (DE) algorithm (Yu and He 2006), firefly algorithm (FA) (Horng et al. 2012). The purpose of applying these algorithms is to find the best settings of the center and the width of the hidden units in RBFN to achieve optimal network performance (Simon 2002). In this paper, HS is proposed to optimize the center and the width of each hidden unit in a trained RBFN for optimizing its recognition performance. The reason why HS is adopted in this work is that conventional EA such as GA utilizes two parental vectors to generate new solution vectors, whereas HS involves all existing vectors. Therefore, HS has a higher ability in obtaining better solutions when compared to those conventional EAs (Mahdavi et al. 2007).

The training procedure of RBFN-DDA-HS model can be summarized as follows:

a. **RBFN-DDA Training** The proposed model begins with the training of RBFN with the DDA Algorithm as described in Sect. 3.1 by using a training data set. After completing the RBFN-DDA training process, the trained solution vector is formed as a set of center and radius of hidden units \( \tilde{c}_{\text{train}}, \tilde{r}_{\text{train}} \).

b. Defining objective function and setting the HS parameters. The fitness of the solution vector is evaluated in term of accuracy and hence the objective function is defined as below:

Fig. 5 The HS algorithm procedure

| Table 2 | The HS parameter settings |
|---------|---------------------------|
| Harmony search parameter | Value |
| HMS | 10 |
| HMCR | 0.9 |
| PAR | 0.5 |
| Maximum improvisation | 100 |
Maximum improvisation is referred to as the largest number of iteration of the evolution in an HS. A common setting for maximum improvisation of HS is occasionally greater than 10,000 for achieving high accuracy rates. The reason is, the initial solution of HS is set randomly because prior knowledge about the problem is not known. A high setting in maximum improvisation is a way adopted by HS to search for solutions to achieve good accuracy rates. However, in this study, the initial harmony memory is not randomly set, it explores from a set of knowledge of the problem learned by the RBFN-DDA. RBFN-DDA that performs incremental learning can absorb information about the problem from the data and provides high quality initial solution. HS is employed to perform exploration on this solution with an aim to achieve high performance in terms of accuracy rates. Notably, the number of iteration of the HS component in RBFN-DDA-HS is set with a small number to avoid overtraining. Another advantage of RBFN-DDA-HS is that it can shorten computation time as compared to performing search using HS alone.

c. Forming an initial harmony memory from a trained RBFN. An initial population called harmony memory (HM) is generated from a set of parameters from a trained RBFN instead of generating this population randomly. By referring Sect. 3.2, the HMS is set as 10. Thus, 10 sets of solution vector stored in the HM matrix \( \mathbf{z} \) as shown below:

\[
\begin{bmatrix}
\mathbf{z}_1^1 & \mathbf{z}_2^1 & \mathbf{z}_{m-1}^1 & \mathbf{z}_m^1 \\
\mathbf{z}_1^2 & \mathbf{z}_2^2 & \mathbf{z}_{m-1}^2 & \mathbf{z}_m^2 \\
\vdots & \vdots & \vdots & \vdots \\
\mathbf{z}_1^g & \mathbf{z}_2^g & \mathbf{z}_{m-1}^g & \mathbf{z}_m^g \\
\mathbf{z}_1^h & \mathbf{z}_2^h & \mathbf{z}_{m-1}^h & \mathbf{z}_m^h
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
\mathbf{r}_1^1 \\
\mathbf{r}_2^1 \\
\vdots \\
\mathbf{r}_1^g \\
\mathbf{r}_2^g \\
\vdots \\
\mathbf{r}_1^h \\
\mathbf{r}_2^h
\end{bmatrix}
\]

where \( \mathbf{z}_i^j \) is, respectively, the center and radius of RBFN; \( m \) is the number of hidden units in RBFN. The other solution vectors \( \mathbf{z}_i^r \) are developed according to:

\[
\mathbf{z}_i^r = \mathbf{z}_i^z + \text{rand}(0,1) \times \text{RMF}
\]

(2)

A Relative Multiple Factor (RMF) value \( \in [0,1] \) is applied on the trained solution vectors to control the variation of the developed center and width. All solution vectors stored in the harmony memory are then evaluated by using the objective function defined in step (b).

d. Evolving solution vectors. A new solution vector \( \mathbf{z}^{\text{new}}_i \) is generated either by slightly adjusting the solution candidate from any existing solution candidate in HM or by creating a new solution candidate on a random basis as described in Session 3.2. The fitness of new solution vector is computed as in step (b). If the new solution is better than the worst one, then it will be included in the HM while the worst one will be removed.

e. The procedure is either continued as in step (d) or is terminated if the maximum number of improvisations has been reached. A flowchart of the RBFN-DDA-HS model is listed in Fig. 6.

### 4 Experiments and results

Fault detection is essential to ensure the effectiveness and reliability of machinery. In rotating machinery, bearing has been identified as one of the main fault components. Hence, fault detection of bearing has attracted a great attention from researchers. Besides, extensive research have been also focused on a data classification approach to fault detection in the manufacturing industry such as tool wear monitoring and machining parameter prediction that are described in Session 2.2. Condition monitoring and fault detection techniques are applied in power generation industry with an intention to avoid sudden breakdowns which may result in costly repair and machine unavailability. The effectiveness of the proposed model is evaluated in two benchmark fault detection problems, which are bearing fault classification and steel plate fault detection problems. This study also involves the application of the proposed RBFN model to perform condition monitoring in a real case study, which is a circulating water (CW) system in a power generation plant. The dataset in each problem was randomly divided into both training and testing data set. All the attributes in the dataset were normalized between 0 and 1. To compare classification performances of RBFN-DDA and RBFN-DDA-HS statistically in monitoring the operating condition of the CW system, a Wilcoxon signed rank test was employed at a level of significance \( \alpha = 0.05 \). The classification performance of the proposed RBFN-DDA-HS was also compared with other machine learning methods in all three problems for which the results of these machine learning methods are taken from Tan et al. (2007), Kavathkar et al. (2016), Tan and Lim (2015) and Wong (2015), respectively.

#### 4.1 Fault detection applications

##### 4.1.1 Benchmark dataset 1: Bearing data set

The RBFD-DDA-HS is applied to classify the bearing faults. The benchmark dataset is taken from Case Western
Reserve University (CWRU) (Loparo 2003). The data were generated from a setup consisting of a 2HP motor, a dynamometer and bearing support (Fig. 7). Data were collected from an accelerometer at 48,000 Hz which means 48,000 samples/second was collected in the bearing experiments. The load was varied from 0 to 3 HP with fault dimensions varying from 0.1761 to 0.7044 mm. Ten features were extracted from drive end signal, which were mean, standard deviation, variance, root mean square value, skewness, kurtosis, minimum value, peak value, crest factor and form factor. The bearing dataset contains 55 records that indicate four types of bearing fault condition including health bearing, inner race (IR) defect, ball defect and outer race (OR) defect. Details of this dataset are explained by Vakharia et al. (2016).

4.1.2 Benchmark dataset 2: Steel plate data set

The data set for steel plate fault detection problem can be downloaded from the UCI web repository (Lichman et al. 2013). A total of 27 numerical attributes of a steel-plate image are used to classify each sample into one of the seven types of steel plate’s faults, namely pastry, Z-scratch, K-scratch, stains, dirtiness, bumps, and others. The experiment is repeated for 30 times, with different training data to obtain average results. The data set consists of 1941 records which are divided into a training set of 1457 samples and a test set of 484 samples in the study.
4.1.3 Case study: the CW system

The proposed RBFN-DDA-HS is applied to perform condition monitoring by learning and classifying a set of real data collected from a CW system in a power generation plant in Penang, Malaysia. The CW system operates to provide cooling water continuously to the main turbine condenser to condense steam from the turbine exhaust (Berhad and System description and operating procedures. 1999). The overall water steam cycle efficiency in power plant is highly relied on the operating condition of the CW system. Intelligent system such as RBFN-DDA-HS could be helpful to monitor the operating conditions of the CW system and reduce the frequency of unexpected breakdown. Figure 8 shows an overview of the CW system. The CW system includes all piping and equipment (such as condensers and drum strainer) between seawater intake and the outfall where water is returned to the sea. The description of the method of seawater processing and how it is transferred into the CW system can be referred in Tan and Lim (2004). A targeted 80 MW power generation is the fundamental environment to establish the database. For every 5 min, an input sample of 12 temperature and pressure measurements was collected at the inlet and outlet points of the condenser. The operating conditions of the CW system were identified in four classes, which are listed in Table 3. The database had a total number of 2500 input samples.

| Class | Operation condition |
|-------|---------------------|
| 1     | Heat transfer in the condenser is efficient and there is no significant blockage in the piping system |
| 2     | Heat transfer in the condenser is not efficient and there is no significant blockage in the piping system |
| 3     | Heat transfer in the condenser is efficient and there is significant blockage in the piping system |
| 4     | Heat transfer in the condenser is not efficient, and there is significant blockage in the piping system |

Fig. 8 The circulating water (CW) system

4.2 Results and discussion

4.2.1 Performance comparison between RBFN-DDA and RBFN-DDA-HS

The experiment was conducted by repeatedly training the classifiers using different training and testing data before computing average classification results. We followed the experimental setup as mentioned in Kavathekar et al. (2016), Tan and Lim (2015) and Tan et al. (2007). In this case, for steel plate (Tan and Lim 2015) problem and CW system case study (Tan et al. 2007), the experiment was repeated for thirty and ten times, respectively. However, in experiment using the bearing dataset (Kavathekar et al. 2016), the number of repetition of machine learning in training and testing was not mentioned. In our work, we repeat the experiment using RBFN-DDA and RBFN-DDA-HS for thirty times. All average results of training and testing accuracy of RBFN-DDA and RBFN-DDA-HS are
shown in Table 4, RBFN-DDA-HS performs with higher accuracy rates than RBFN-DDA in all three dataset. These results signify the performance of RBFN-DDA has been improved after its learning is integrated with a search and adaptation process by the HS algorithm.

A Wilcoxon signed rank test (Woolson 2007) is applied to compare statistically the classification performance between RBFN-DDA and RBFN-DDA-HS at a level of significance $\alpha = 0.05$ in all three problems. In this test, the null hypothesis is that the testing accuracy of RBFN-DDA-HS is the same as that of RBFN-DDA. The alternative hypothesis is that the testing accuracy of RBFN-DDA-HS is different from RBFN-DDA. By referring to the results in Table 5, all $p$ values are smaller than 0.05. This means that the classification performances of RBFN-DDA and RBFN-DDA-HS in terms of testing accuracy are statistically different. In this regard, RBFN-DDA-HS achieves better accuracy rates than RBFN-DDA in handling three tasks related to condition monitoring and fault detection.

### 4.2.2 Performance comparison with other machine learning methods

The classification performance of RBFN-DDA-HS is compared with other machine learning methods in all three problems Kavathekar et al. (2016), Tan and Lim (2015) and Tan et al. (2007) from which their accuracy rates are referred. Table 6 shows the results. In bearing dataset, the result of RBFN-DDA-HS is compared with machine learning classifiers equipped with a feature selection algorithm (i.e., Random forest and Rotation Forest), and without a feature selection algorithm (i.e., ANN, SVM, Decision Tree, RBFN-DDA) (Kavathekar et al. 2016). The classification accuracy of RBFN-DDA (i.e., 64.35%) is higher than the classifiers without feature selection such as ANN, SVM and Decision Tree, but is less accurate than classifiers with feature selection (i.e., Random forest and Rotation Forest). When RBFN-DDA classifier is integrated with HS, the proposed RBFN-DDA-HS model outperforms all other classifiers by achieving the highest classification accuracy, i.e., 78.81%.

Next, in steel plate dataset the classification performance of RBFN-DDA-HS is compared with a metaheuristic classifier (FAM-GSA) and others single machine learning classifiers (FAM, MLP and RBF) (Tan and Lim

| Method | (a) Bearing | (b) Steel plate | (c) CW system case study |
|--------|-------------|-----------------|-------------------------|
|        | Training accuracy (%) | Testing accuracy (%) | Training accuracy (%) | Testing accuracy (%) | Training accuracy (%) | Testing accuracy (%) |
| RBFN-DDA | 65.66 | 64.35 | 90.58 | 68.41 | 93.16 | 91.94 |
| RBFN-DDA-HS | 80.06 | 78.81 | 92.59 | 69.94 | 96.56 | 96.20 |

#### Table 5 Results from the Wilcoxon rank tests (at the level of significance $\alpha = 0.05$)

| Dataset       | $p$ value          |
|---------------|--------------------|
| CW system     | 0.002              |
| Bearing       | $1.66 \times 10^{-6}$ |
| Steel plate   | $3.86 \times 10^{-6}$ |

#### Table 6 Performance comparison in all three problems

| Model         | (a) Bearing | (b) Steel plate | (c) CW system case study |
|---------------|-------------|-----------------|-------------------------|
|               | Testing accuracy (%) | Testing accuracy (%) | Testing accuracy (%) |
| ANN           | 57.14       | FAM             | FAM-RecBFN (Tan et al. 2007) |
| SVM           | 46.42       | FAM-GSA         | ELM (RBF) (Wong 2015) |
| Decision tree | 58.92       | MLP             | C-ELM (RBF) (Wong 2015) |
| Random forest | 66.07       | RBF             | RBFN-DDA |
| Rotation Forest | 75.00   | RBFN-DDA       | RBFN-DDA-HS |
| RBFN-DDA      | 64.35       | RBFN-DDA-HS    | 96.20       |
| RBFN-DDA-HS   | 78.81       |                 |             |
Based on the results in Table 6, the testing accuracy of RBFN-DDA is higher than other single machine learning classifiers. Note that FAM-GSA contributes only a small improvement in testing accuracy, i.e., 0.9% from FAM. Both FAM-GSA and RBFN-DDA-HS are metaheuristic classifiers. In comparison, the proposed RBFN-DDA-HS is more effective than FAM-GSA in detecting steel plate faults. The HS can help improve RBFN-DDA-HS by a testing accuracy of 2.2% from RBFN-DDA.

The performance of the RBFN-DDA-HS in CW system is compared with fuzzy ARTMAP (FAM) rectangular basis function network (RecBFN), RBF-based Extreme Learning Machine and RBF-based Constrained Optimization Extreme Learning Machine (C-ELM). By observing the results in Table 6, the testing accuracy of RBFN-DDA is the lowest if compared to other machine learning methods. However, when RBFN-DDA is integrated with the HS algorithm to form the RBFN-DDA-HS, its testing accuracy improves greatly and outperforms all other machine learning methods. The HS algorithm is effective to search for the optimal parameter settings.

5 Summary

In this study, a hybrid classification model of RBFN-DDA and the HS algorithm is proposed to improve the RBFN-DDA learning parameters. HS is adopted in the study due to its simplicity, high flexibility and search efficiency. The effectiveness of the proposed algorithm is demonstrated in the condition monitoring of circulating water (CW) system and two fault detection benchmark datasets, which are bearing and steel plate. The proposed RBFN-DDA-HS achieves the highest classification accuracy if compared to other machine learning methods. Result shows RBFN-DDA-HS can improve classification performance in terms of testing accuracy that has a range between 2.2 and 22.5% from RBFN-DDA. Besides, the results from the Wilcoxon signed rank test show that the classification performances of RBFN-DDA and RBFN-DDA-HS are statistically different for all three condition-monitoring case studies where the latter outperforms the former. HS algorithm is a musical inspired metaheuristic algorithm. In the future, the classification performances of RBFN-DDA integrated with other types of metaheuristic algorithm (e.g., from biological, physical or chemical type) would be investigated. In addition, research may also focus on combining RBFN-DDA with two metaheuristic algorithms. The integration of two metaheuristic algorithms may explore the search space for optimal solutions more efficiently and effectively.

Acknowledgements The authors would like to acknowledge the Ministry of Higher Education of Malaysia for the financial support under the Fundamental Research Grant Scheme (FRGS), Grant No: FP061-2015A and also the Ministry of Education Malaysia for the financial support under the program myBrain15.

Data availability 1. The bearing dataset analyzed during the current study is available in the bearing data center of Case Western Reserve University (CWRU) https://csegroups.case.edu/bearingdatacenter/home. 2. The steel plate dataset analyzed during the current study is available in the UCI repository https://archive.ics.uci.edu/ml/datasets/steel+plates+faults. 3. The circulating water (CW) system dataset analyzed during the current study is not publicly available due the copyright is owned by the power generation plant but is available from the corresponding author on reasonable request.

Code availability NA.

Declaration

Conflict of interest The authors declare that there is no conflict of interests regarding the publication of this paper.

References

Abualigah LM, Khader AT, Hanandeh ES (2018a) A combination of objective functions and hybrid Krill herd algorithm for text document clustering analysis. Eng Appl Artif Intell 73:111–125
Abualigah LM, Khader AT, Hanandeh ES (2018b) A new feature selection method to improve the document clustering using particle swarm optimization algorithm. J Comput SciScn 25:456–466
Abualigah L et al (2021a) Aquila optimizer: a novel meta-heuristic optimization algorithm. Comput Ind Eng 157:107250
Abualigah L et al (2021b) The arithmetic optimization algorithm. Comput Methods Appl Mech Eng 376:113609
Ahmad R, Akhtar N, Choubey N (2017) Applications of artificial bee colony algorithms and its variants in health care. Bio-Chem Indian J 11(1):110
Ahmadi MA et al (2018) An accurate model to predict drilling fluid density at wellbore conditions. Egypt J Pet 27(1):1–10
Ali YH (2018) Artificial intelligence application in machine condition monitoring and fault diagnosis. Artif Intell Emerg Trends Appl 275
Azzini A, Tettamanzi AG (2011) Evolutionary ANNs: a state of the art survey. Intell Artif 5(1):19–35
Barreto AM, Barbosa HJ, Ebecken NF (2002) Growing compact RBF networks using a genetic algorithm. In: VII Brazilian symposium on neural networks, 2002. SBRN 2002. Proceedings. IEEE
Ben Ali J et al (2015) Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. Appl Acoust 89:16–27
Berthold MR, Diamond J (1998) Constructive training of probabilistic neural networks. Neurocomputing 19(1):167–183
Bhoskar MT et al (2015) Genetic algorithm and its applications to mechanical engineering: a review. Mater Today Proc 2(4–5):2624–2630
Bi TS et al (2000) A novel ANN fault diagnosis system for power systems using dual GA loops in ANN training. In: 2000 power engineering society summer meeting (Cat. No.00CH37134)
Blanchard BS (2004) Logistics engineering and management. Pearson Prentice Hall, Hoboken
Cerrada M et al (2018) A review on data-driven fault severity assessment in rolling bearings. Mech Syst Signal Process 99:169–196
Singh GK (2003) Induction machine drive condition monitoring and diagnostic research—a survey. Electric Power Syst Res 64(2):145–158
Sivakumar M, Parvathi R (2014) Application of neural network trained with meta-heuristic algorithms on fault diagnosis of multi-level inverter. Res J Appl Sci 9(6):369–375
Stephens MP (2010) Productivity and reliability-based maintenance management. Purdue University Press, West Lafayette
Tallam RM, Habetler TG, Harley RG (2003) Stator winding turn-fault detection for closed-loop induction motor drives. IEEE Trans Ind Appl 39(3):720–724
Tan SC, Lim CP (2004) Application of an adaptive neural network with symbolic rule extraction to fault detection and diagnosis in a power generation plant. IEEE Trans Energy Convers 19(2):369–377
Tan SC, Lim CP (2015) Evolving an adaptive artificial neural network with a gravitational search algorithm. In: Intelligent decision technologies. Springer, pp 599–609
Tan SC, Lim CP, Rao MVC (2007) A hybrid neural network model for rule generation and its application to process fault detection and diagnosis. Eng Appl Artif Intell 20(2):203–213
Tenaga Nasional Berhad (1999) System description and operating procedures. Prai Power Station Stage 3
Tidriri K et al (2016) Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: a review of researches and future challenges. Annu Rev Control 42:63–81
Vakharia V, Gupta VK, Kankar PK (2016) A comparison of feature ranking techniques for fault diagnosis of ball bearing. Soft Comput 20(4):1601–1619
Venkatasubramanian V, Rengaswamy R, Kavuri SN (2003a) A review of process fault detection and diagnosis. Comput Chem Eng 27(3):313–326
Venkatasubramanian V et al (2003b) A review of process fault detection and diagnosis. Comput Chem Eng 27(3):293–311
Venkatasubramanian V et al (2003c) A review of process fault detection and diagnosis. Comput Chem Eng 27(3):327–346
Waeyenbergh G, Pintelon L (2002) A framework for maintenance concept development. Int J Prod Econ 77(3):299–313
Wang G, Yang Y, Guo Z (2013) Hybrid learning based Gaussian ARTMAP network for tool condition monitoring using selected force harmonic features. Sens Actuators A 203:394–404
Wong SY (2015) Novel extreme learning machines for pattern classification, data regression, and rules generation, in College Of Engineering, Universiti Tenaga Nasional, Kajang
Woolson R (2007) Wilcoxon signed-rank test. Wiley, Hoboken, pp 1–3
Xiao Y, Feng L (2012) A novel neural-network approach of analog fault diagnosis based on kernel discriminant analysis and particle swarm optimization. Appl Soft Comput 12(2):904–920
Xin Y (1999) Evolving artificial neural networks. Proc IEEE 87(9):1423–1447
Yao X (1993) A review of evolutionary artificial neural networks. Int J Intell Syst 8(4):539–567
Yin S et al (2014) A review on basic data-driven approaches for industrial process monitoring. IEEE Trans Ind Electron 61(11):6418–6428
Yu B, He X (2006) Training radial basis function networks with differential evolution. In: Proceedings of IEEE international conference on granular computing. 2006. Citeseer
Zaiping C et al (2008) Neural network electrical machine faults diagnosis based on multi-population GA. In: 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)
Zaranezhad A, Mahabadi HA, Dehghani MR (2019) Development of prediction models for repair and maintenance-related accidents at oil refineries using artificial neural network, fuzzy system, genetic algorithm, and ant colony optimization algorithm. Process Saf Environ Prot 131:331–348
Zhou H et al (2016) Effective calculation model of drilling fluids density and ESD for HTHP well while drilling. In: IADC/SPE Asia Pacific drilling technology conference, 2016. Society of Petroleum Engineers

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.