Applied Soft Computing 12 (2012) 2012–2022

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

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

Integration of growing self-organizing map and continuous genetic algorithm for grading lithium-ion battery cells

R.J. Kuo[a,*], C.F. Wang[b], Z.Y. Chen[c]

[a] Department of Industrial Management, National Taiwan University of Science and Technology, No. 43, Section 4, Kee-Lung Road, Taipei, Taiwan
[b] Department of Computer Integrated Manufacture, Lextar Electronics Corporation, No. 3, Gongye E. 3rd Road, Hsinchu Science Park, Hsinchu, Taiwan
[c] Department of Business Administration, DE LIN Institute of Technology, No. 1, Ln. 380, Qingyun Road, Tucheng Dist, New Taipei City, Taiwan

ARTICLE INFO

Article history:
Received 20 November 2010
Received in revised form 9 January 2012
Accepted 25 January 2012
Available online 20 March 2012

Keywords:
Cluster analysis
Self-organizing map neural network
Continuous genetic algorithm
Growing self-organizing map

ABSTRACT

This study attempts to employ growing self-organizing map (GSOM) algorithm and continuous genetic algorithm (CGA)-based SOM (CGASOM) to improve the performance of SOM neural network (SOMnn). The proposed GSOM + CGASOM approach for SOMnn is consisted of two stages. The first stage determines the SOMnn topology using GSOM algorithm while the weights are fine-tuned by using CGASOM algorithm in the second stage. The proposed CGASOM algorithm is compared with other two clustering algorithms using four benchmark data sets, Iris, Wine, Vowel, and Glass. The simulation results indicate that CGASOM algorithm is able to find the better solution. Additionally, the proposed approach has been also employed to grade Lithium-ion cells and characterize the quality inspection rules. The results can assist the battery manufacturers to improve the quality and decrease the costs of battery design and manufacturing.

1. Introduction

In recent years, cluster analysis has been widely applied in many applications. Cluster analysis is the process of identifying natural groupings or clusters within multidimensional data, based on some similar measures, like Euclidean distance [1]. Its main purpose is to group samples with the same statistical characteristics together into the same cluster in order to achieve higher similarities within same clusters. Also, there are more significant differences between different clusters [2]. Cluster analysis is implemented by using statistical or neural network (NN) methods. Due to its high analytical value, cluster analysis is widely applied in variety of areas including business, education, social sciences, genetics and biology.

Among the major clustering algorithms, unsupervised neural network is one of the most representative methods while self-organizing map neural network (SOMnn) [3] is the most frequently applied one since it can be used for applications like image processing and mode identification. SOMnn mainly searches for different clusters within the data through sample training. Some researches have been made to propose different versions of SOMnns. Thus, this study aims to develop an enhanced cluster analysis approach which is able to group data which share similar characteristics to discover potential data traits and usable information based on SOMnn.

The proposed approach is a two-stage cluster analysis algorithm. In the first stage, the growing SOM (GSOM) algorithm [4] uses input data to determine the suitable SOMnn topology. In the second stage, continuous genetic algorithm (CGA) [5] is integrated with conventional SOM (i.e., CGA-based SOM (CGASOM)) algorithm in searching for the optimal weight vectors of SOMnn. To verify the proposed CGASOM approach for SOMnn, GSOM algorithm and four benchmark data sets (i.e., Iris, Glass, Vowel, and Wine) are employed. Then, the proposed CGASOM algorithm is further applied for grading batteries while incorporating real battery measurement results to obtain a comprehensive assessment. The result can be used as the foundation for automated battery grading evaluation system.

The rest of this paper is organized as follows. Section 2 presents the general background related to this study, while the proposed CGASOM algorithm is explained in Section 3. Sections 4 and 5 show the simulation results and the model evaluation results, respectively. The concluding remarks are finally made in Section 6.

2. Literature review

This section will briefly presents general backgrounds regarding self-organizing map (SOM) neural network, genetic algorithm and hybrid network.

2.1. Self-organizing map neural network (SOMnn)

The attractiveness of neural networks come from the remarkable information processing characteristics of a biological system

* Corresponding author. Tel.: +886 2 27376328; fax: +886 2 27376344.
E-mail addresses: rjkuo@mail.ntust.edu.tw (R.J. Kuo), winyboys@gmail.com (C.F. Wang), keyzyc@gmail.com (Z.Y. Chen).

1568-4946/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2012.01.018

such as non-linearity, high parallelism, robustness, fault and failure tolerance, learning, ability to handle imprecise and fuzzy information, and their capability to generalize [6]. Self-organizing maps (SOMs) are well-known clustering methods proposed by Kohonen [3], and have proven to be extremely useful for input data with high dimensionality and complexity [7]. Some recent investigations have adopted SOM as a key data clustering technology [8].

The SOM architecture was originally motivated by the topological maps and the self-organization of sensory pathways in the brain [9]. The standard SOM algorithm is a computationally powerful abstraction of the actual biological mechanism. This algorithm realizes a self-organizing process following two simple rules [10]: (a) start a competition among nodes finding the best matching unit (the winner) and (b) adapt the synaptic weights of the winner and topological neighborhood of it.

The main focus of the SOM is to summarize information while preserving topological relationships [11]. The objective of SOM is to represent high-dimensional input patterns with weight vectors that can be visualized in a usually two-dimensional (2D) lattice structure [51]. SOM input patterns are transformed into a one or usually 2D lattice structure, consisting of nodes associated with different clusters [12]. Each unit in the lattice is called a neuron, and adjacent neurons are connected to each other, which gives the clear topology of how the network fits itself to the input space. Input patterns are fully connected to all neurons via adaptable weights, and during the training process, neighboring input patterns are projected into the lattice, corresponding to adjacent neurons. A neuron is iteratively updated during training based on the learning vectors so a well-trained SOM represents a distribution of the input data over a 2D surface preserving topology. In this context, a cluster can be defined as a group of neurons with short distances between them and long distances to the other neurons [13].

To overcome the static architecture of Kohonen's model, several improved SOM algorithms have been proposed. They are described as follows:

2.1.1. Incremental SOM

The incremental grid growing (IGG) [14], growing cell structures (GCSs) [15], and growing neural gas (GNG) [16] are some of the best known incremental SOM. Additionally, several improved SOM and SOM-related algorithms have been proposed in recent years [17].

2.1.2. Growing SOM

A dynamic feature map model called the growing SOM (GSOM) is proposed in [4]. The need for a measure for controlling the growth of the GSOM (or any such algorithm) is highlighted, and an indicator called the spread factor (*SF*) is presented as a method of achieving such a control. The GSOM uses the basic concepts of self-organization as the SOM but has a dynamic structure that is generated during the training process itself [4].

2.1.3. Hierarchical SOM

The hierarchical SOM (HSOM) is another variation of SOM [18]. The basic idea of HSOM is to use multiple SOMs from the low-resolution level to the high resolution level, but the number of neurons in each layer is predefined [19].

2.1.4. Growing HSOM

Growing HSOM (GHSOM) [20] is another HSOM and it can grow neurons horizontally at each level or vertically for the whole structure under some condition. Their approach combines individually growing SOMs with a hierarchical architecture and has successfully been applied to the organization of document collections and music repositories [19].

2.2. Genetic algorithms (GAs) in clustering

As an approach to global optimization, GAs have been found to be applicable to optimization problems that are intractable for exact solutions by conventional methods [21]. GAs were pioneered in 1975 by Holland, since GAs are good at searching method, it can cluster the data according to their similarities. Its concept is to mimic the natural evolution of a population by allowing solutions to reproduce and to create new solutions, which then compete for survival in the next iteration [22].

GAs are a set-based search algorithm where in each iteration it simultaneously generates a number of solutions. In each step, a subset of the current set of solutions is selected based on their performance and these solutions are combined into new solutions. For GAs and other evolutionary methods the defining element is the innovative manner in which the crossover and mutation operators define a neighborhood of the current solution. This allows the search to quickly and intelligently traverse large parts of the solution space [23].

Some studies [21,24–26] show that the clustering results of GA are better than the those of the K-means algorithm. Several GA-based clustering algorithms have been developed [27,28]. Also, Wu et al. [29], Shi et al. [30], Guo and Ning [31], and Sha and Che [32] pointed out that GAs are applicable to solving complicated problems and could obtain globally optimal solutions effectively. Next, Das et al. [33] used a novel representation scheme (i.e., automatic clustering differential evolution (ACDE) algorithm) for the search variables to determine the optimal number of clusters. De et al. [34] also presented an automatically clustering method using GA. In recent years, GAs have been widely used in cluster analysis [35]. Owing to its promising results, this study employs GAs for cluster analysis.

2.3. Hybrid algorithm of SOMnn and GAs

One important characteristic of the SOM is that the feature map preserves neighborhood relations of the input pattern [36]. Due to its fast learning, self-organized and graphical inherence, literature indicates that the SOM network can be effectively employed to narrow down initial design options [37]. In addition, the main objective of the SOM is to perform a smooth mapping from the high dimensional input space onto the low dimensional output space that preserves the topology of the input data, so that close points in the input space are mapped on close points in the output space according to the defined neighborhood relationship [38]. The SOM algorithm determines the method of creating the input space mapping. This is an iterative process of training that adapts the synaptic vectors values of the network to perform a particular task [39].

On the other hand, adaptations of GAs to the continuous optimization problems have been proposed [40]. Most existing local optimization algorithms may be chosen to cooperate in that way with GAs [5]. Furthermore, the parallel searching mechanism is the main advantage of GAs since they cannot easily get trapped in local minima [41]. The power of GAs has recently been noticed due to its powerful search for identification of optimum parameters [42].

Since a binary version of SOM (BSOM)-based [43] reproduction can generate many kinds of chromosomes with high fitness values, Kubota et al. [44] proposed the new updating method of weight vectors to achieve a more effective search of the GA employing BSOM-based reproduction comparing with that employing the traditional BSOM [44]. Moreover, Kuo et al. [21] proposes a novel 2-stage method, which first uses SOMs to determine the number of clusters and then employs a GA-based clustering method to find the final solution. Next, In addition, to enable the generation of explicit knowledge, Chang and Liao [11] constructs fuzzy rule bases with the aid of a SOM and GA. The GA used for this study is a simple GA