Information Sciences 330 (2016) 175–185

Contents lists available at ScienceDirect

Information Sciences

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

Granular-computing based hybrid collaborative fuzzy clustering for long-term prediction of multiple gas holders levels

Zhongyang Han, Jun Zhao*, Quanli Liu, Wei Wang

School of Control Sciences and Engineering, Dalian University of Technology, Dalian, Liaoning Province, China

ARTICLE INFO

Article history:
Received 14 December 2014
Revised 26 June 2015
Accepted 4 October 2015
Available online 19 October 2015

Keywords:
LDG system
Multi-output modeling
Long-term prediction
Granular computing

ABSTRACT

Linz–Donawitz converter Gas (LDG), regarded as an essential secondary energy resource, plays a significant role for the entire production process of steel industry. In a LDG system, the gas holders are crucial equipment for temporary energy storage and buffers connecting with the gas generation units and the gas users. The accurate long-term prediction for the holders levels of such a system would be very necessary for energy scheduling and its optimal decision making. Given the practical characteristics of the LDG system in a steel plant, a granular-computing (GrC)-based hybrid collaborative fuzzy clustering (HCFC) algorithm is proposed in this study for the long-term prediction of the multiple holders levels. The hybrid structure considers the features regarding to a gas holder, of which the horizontal part elaborates the mutual influences among different time spaces of a holder level, while the vertical one describes them among the influence factors (denoting the gas generation units or the users). Then, the modeling algorithm is also explicitly derived in this study. To verify the performance of the proposed approach, two groups of simulation are carried out by employing the real-world industrial data coming from this plant, in which the single-output method and the iterative computing-based one are comparatively analyzed. The results indicate that the proposed approach provides a remarkable accuracy for such an industrial application.

1. Introduction

LDG is a class of byproduct energy resource generated during steel making process which is viewed as an important secondary fuel that can be recycled and consumed by the manufacturing procedures in steel industry. The gas holders levels in LDG system are very significant index monitored for reasonable energy utilization and scheduling.

Due to the complex and long distance distributed gas pipeline networks in steel plant and the nonlinear features of the gas consumption units, it is practically impossible to establish a physics-based model [25,26] for the LDG holder level prediction. Given the accumulative production data coming from the energy management information system, a number of data-driven methods were reported in literature for solving such a problem. For instance, an echo state network (ESN) was modeled in [31] to perform a two-stage prediction for a blast furnace gas (BFG) system of steel energy system. Similarly, a multi-kernel learning algorithm for estimating the gas holders levels in a BFG system was constructed in [32], where the generalization of the model was dynamically improved along with the training time reduction. And, [8] proposed an improved least square support vector machine (LSSVM) to predict the gas holders levels for a LDG system and the corresponding multi-output model was explicitly formulated. However, the existing studies mentioned above can only provide accurate results with relative short term period

* Corresponding author. Tel.: +86 411 84707582.
E-mail address: zhaoj@dlut.edu.cn, jun.z@163.com (J. Zhao).

http://dx.doi.org/10.1016/j.ins.2015.10.020
0020-0255/

(approximately 60–90 min in the near future) because of the iterative prediction mechanism. With a practical and detailed study on-site, it is discovered that a long-term prediction for the LDG holders levels also exhibits great demands for the energy scheduling operations.

In terms of data-based long-term prediction, an auto-regressive and moving average (ARMA) model based approach was reported in [20], where a soft computing technique was designed for a time series forecasting. And, a hybrid model combining a maximal overlap discrete wavelet transform and ARMA was presented in [10] to deal with the non-stationary and long range dependence time series. Nevertheless, the ARMA model exhibited poor generalization ability when applying to the time series with intensive data fluctuation. Moreover, it also required a time-consuming calculation for determination of the multiple parameters. Besides, a wind speed long-term prediction by using a radial basis function (RBF) neural network was designed in [29], in which the linear and nonlinear patterns were individually analyzed along with the consideration of seasonal influence. However, it is very tough for such a method to be applied to this study owing to the difficulties on the pattern and feature recognition of the energy system in steel industry. Recently, a hybrid model combining with intuitive fuzzy set, genetic algorithm and LSSVM was constructed in [9] for a long-term business cycle forecasting. Nevertheless, as for the nonlinear and complex pipeline network characteristics of this study, it is rather hardly to establish a corresponding intuitive fuzzy set for such a LDG system by using the similar methods mentioned above. Although [5] proposed a GrC based method for long-term forecasting on time series which set the prediction horizon as data segments instead of single point, that model primarily analyzed single data set. However, the problem of this study involves lots of objects and data sets. Besides, for industrial data that usually presents the features with irregular fluctuations, that algorithm was still limitative. In the perspective of modeling structure, given that what we concentrate are the multiple gas holders levels rather than a single one, it is also necessary to design a model capable of handling a class of multi-output problem. The existing methodologies, such as Neural Networks (NN) [6,11,21], Takagi-Sugeno (TS) fuzzy model [18,19,27], data processing methods [2,23], did not directly concern the impact of the same variables on different time feature space or between various data objects. Therefore, the above mentioned approaches are not suitable for this study.

Based on the literature review above, a GrC-based hybrid collaborative clustering model for the long-term prediction on multiple gas holders levels of the LDG system is proposed in this study, where the prediction scale is effectively extended to a few hours, even a day. A thorough derivation on the model construction of the proposed hybrid collaborative clustering is reported, which comprehensively considers the mutual impacts not only on the levels of the same holder in different time period, but also between the influence factors, i.e., the objects that influence the current holder level. Two groups of experiments are carried out for explicitly verifying the performance of the proposed approach. In the first group of experiments, compared to the single-output one, the advantage of the multi-output and the mutual influence in modeling is clearly demonstrated. The other group elaborates the effectiveness of the long-term prediction by making contrast with the iterative computing-based approach. The entire results exhibit the satisfactory accuracies helpful to energy scheduling and optimization.

The rest of this paper is organized as follows. In Section 2, the underlying practical problem for the LDG system is described. A multi-output long-term prediction model based on GrC and HCFC are proposed in detail in Section 3. In Section 4, the two categories of experiments are conducted to validate the proposed approach by using the practical data coming from the energy data center of this steel plant. Finally, Section 5 draws the conclusions for this study.

2. Problem descriptions

The structure of a LDG system in steel industry is illustrated in Fig. 1, where two subsystems respectively consist of the gas generation units, the gas consumers, and some transportation units. The six converters stand for the gas generation units, and the gas consumers involve blast furnaces, cold/hot rolling, lime kiln, etc. The pipeline networks, along with four gas holders (two for each subsystem), the gas mixture and the pressure stations serve for the gas transportation. Generated by the converters and then temporarily stored in the gas holders, the gas is delivered to the consumers. Connecting the gas generation units and the consumers, the holders can be deemed as the buffers which balance the gas flow in the two sides, and regarded as the suppliers for the gas consumption units.

Considering such a structure of LDG system in this study, the gas holder plays a significant role in the whole LDG system. In practice, the steel plant needs expert knowledge for acquiring or monitoring the trend of the gas holder level. Due to the low accuracy and the difficulty on estimation, a number of low level scheduling works usually bring about energy waste and manufacturing cost increase. On the other hand, prediction in short-term provides insufficient information for the scheduling works. In a word, if one could accurately predict the holders levels for large time horizon, then it would be very helpful for the energy scheduling operators to arrange and plan for energy optimization.

With a long time period investigation and analysis on-site, it is discovered that the influence factors that impact the holders levels fundamentally including two categories, the holders levels at previous time point, the current gas generation and the consumption flow.

3. Granular-computing based hybrid collaborative fuzzy clustering for multi-holder levels prediction

As for the industrial prediction mentioned in this study, there are two problems that should be resolved.

(1) On the respect of practical requirement, the long-term prediction for the multi-output model (two gas tank levels) is on highly demand for the practical energy application. It is largely helpful for the further energy optimization and scheduling so as to reduce the production cost for enterprises.