Causality Diagram-based Scheduling Approach for Blast Furnace Gas System

Feng Jin, Jun Zhao, *Member, IEEE*, Chunyang Sheng, and Wei Wang, *Senior Member, IEEE*

***Abstract*—Rational use of blast furnace gas (BFG) in steel industry can raise economic profit, save fossil energy resources and alleviate the environment pollution. In this paper, a causality diagram is established to describe the causal relationships among the decision objective and the variables of the scheduling process for the industrial system, based on which the total scheduling amount of the BFG system can be computed by using a causal fuzzy C-means (CFCM) clustering algorithm. In this algorithm, not only the distances among the historical samples but also the effects of different solutions on the gas tank level are considered. The scheduling solution can be determined based on the proposed causal probability of the causality diagram calculated by the total amount and the conditions of the adjustable units. The causal probability quantifies the impact of different allocation schemes of the total scheduling amount on the BFG system. An evaluation method is then proposed to evaluate the effectiveness of the scheduling solutions. The experiments by using the practical data coming from a steel plant in China indicate that the proposed approach can effectively improve the scheduling accuracy and reduce the gas diffusion.**

***Index Terms*—Blast furnace gas system, causal fuzzy C-means (CFCM) clustering, causality diagram, scheduling.**

I. INTRODUCTION

BLAST furnace gas (BFG) is one of the most important byproduct gases generated from blast furnace, which is widely used as secondary fuel for some other production processes in iron and steel industry. Rational scheduling of BFG can raise the economic profit, improve the production efficiency, reduce the waste of energy and alleviate the environmental pollution. In practical production processes, scheduling solutions are determined mostly by the experience of the scheduling experts. In such a way, low accuracy and incomplete scheduling might be generated. Thus, studies focused on the scheduling problem were presented in literature. A Bayesian network was used to determine the structure of the adjustable users in [1], and the scheduling amount was obtained by computing the posterior probability. In order to obtain the scheduling amount, a hybrid parameter optimization algorithm was developed to optimize the model for high prediction accuracy in [2]. Also, a Gaussian process-based echo states network (ESN) was proposed to predict the gas tank level, and a certain heuristic method was developed to quantify the users gas adjustment [3]. A dynamic mixed integer linear programming model is established to achieve total scheduling amount in [4] and a benefit and cost (BAC) model-based algorithm was proposed to solve the online byproduct gas scheduling problem in [5]. However, these methods enumerated above had two common limitations. One is that the real-time status of the scheduling units was not being taken into account, and the other is that the total scheduling amount obtained by the gas tank level prediction may cause the accumulation of iteration error.

Fuzzy C-means (FCM) clustering [6] is one of the major clustering algorithms which has been widely used in many fields, such as the long-term prediction for time series [7], the image segmentation [8] and the feature extraction [9]. FCM establishes the uncertainty description of the samples and can objectively reflect the actual situation of industrial data. Based on its characteristics that data can be separated into a number of groups and the objects in each group showing a high similarity, FCM is capable of selecting the historical scheduling solutions most close to the current condition and calculating the total scheduling amount. However, the traditional FCM is not suitable for such an industrial scheduling problem since the reported studies in literature only take the geometric distances among the sample data into account, which failed to clearly describe the causal relationships among the clustering variables, especially for the practical industrial problem.

Causality is a kind of analysis method which is highly consistent with human cognition mode. Pearl established the causality model [10], and employed a directed acyclic graphs (DAGs) to denote the cause and the effect among variables. One can reference a series of typical instances related to the causality-based research. A syntax and semantics of neuron diagrams were formalized to identify the causal effects in [11]. A graphical representation of missing data mechanism was presented in [12]. And then a causal model was reported to solve the problem of estimating the causal relationship from data with missing entries [13]. Similarly, a minimal causal model is presented for the reliable knowledge discovery [14]. In addition, a causal effect criterion was made for model selection [15]. The above mentioned studies indicated the reliability of the causality-based approaches when facing with the data-driven modeling or referencing.

Manuscript received August 23, 2016; accepted May 25, 2017. This work was supported by the National Natural Sciences Foundation of China (61473056, 61533005, 61522304, 61603068, U1560102). Recommended by Associate Editor Naiqi Wu. *(Corresponding author: Jun Zhao.)*

Citation: F. Jin, J. Zhao, C. Y. Sheng, and W. Wang, "Causality diagram-based scheduling approach for blast furnace gas system," *IEEE/CAA J. of Autom. Sinica*, vol. 5, no. 2, pp. 587−594, Mar. 2018.

F. Jin, J. Zhao, and W. Wang are with the Faculty of Electronics Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China (e-mail: jinfeng_1126@126.com; zhaoj@dlut.edu.cn; wangwei@dlut.edu.cn).

C. Y. Sheng is with the College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China (e-mail: chunyang860124@163.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2017.7510715

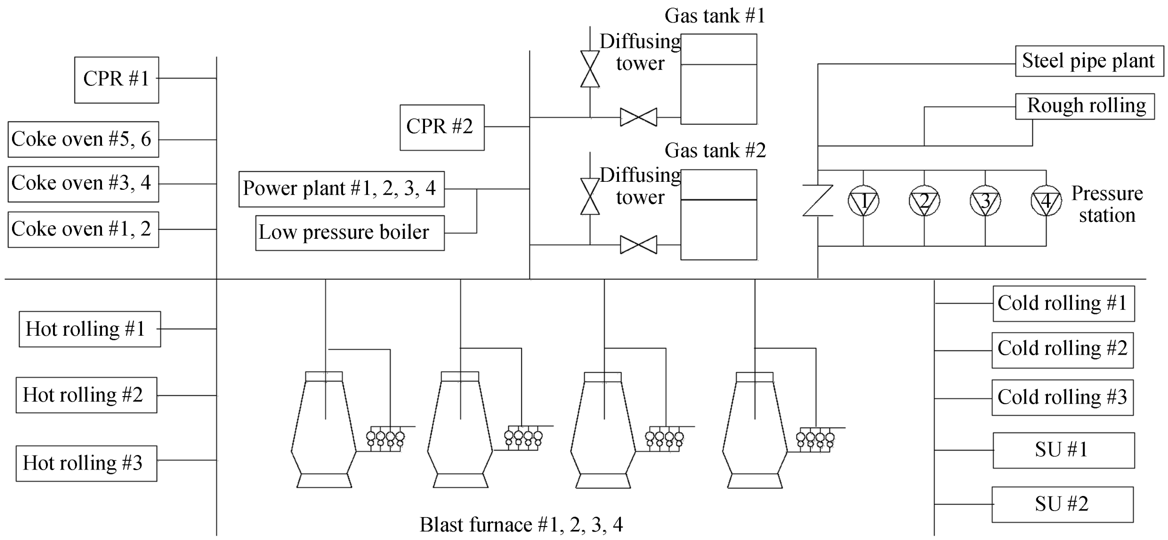


Fig. 1. Structure of BFG system.

In this paper, a causality diagram is established to describe the causal relationships among the decision objective and variables of the scheduling process for the BFG energy system, based on which the total scheduling amount can be computed by using a modified fuzzy C-means clustering algorithm. Given that the Euclidean distance is combined with the causal relationships among the clustering variables, the proposed clustering method is more suitable to the application problems described in this paper than the traditional one. A most reasonable solution will be generated according to the current status of the scheduling units. Then, an evaluation method is designed to quantify the effectiveness of the suggested solution. A series of comparative experiments using the real industrial data are carried out. The results indicate that the proposed method achieves much higher scheduling accuracy and improves the efficiency of the BFG utilization.

This paper is organized as follows. In Section II, a practical scheduling problem in steel industry is described. And a causality diagram-based approach is proposed in Section III to deduce the scheduling solution. The proposed method is verified in Section IV by using the real data from a steel plant in China, and the related results with some analysis are addressed. Finally, Section V draws the conclusions for this study.

II. Problem Description

In practice (the studied Shanghai Baosteel Co. Ltd, in China), a large amount of sensors acquire the energy real time data and transfer them into programmable logic control (PLC). And then, the data are stored into the real time database (iHistorianTM on-site), and are displayed with various time intervals on the supervisory control and data acquisition (SCADA) system. The running states of the energy units are monitored by the energy operators through SCADA system on-site. The BFG system in steel industry is very complex, and its balance status is mainly represented by the varying tendency of the gas tank level. Fig. 1 shows the structure of the BFG system. Regarded as a kind of very important byproduct gas, BFG is generated from four blast furnaces, and its generation amounts can be up to 1800 km^3/h on average. The transportation system contains pressure stations, mixing devices and pipeline network, through which BFG can be transported to a number of consumption units, such as coke ovens, hot rolling and cold rolling. The surplus gas is stored in the two gas tanks, which are connected to the pipeline as buffer devices, and they always keep one connected to the pipeline, and the other standby at high tank level for emergent situation. Each of the blast furnaces is fed by four hot blast stoves. Due to the production requirements, the combustion status of them should be switched frequently. As the BFG consumption of one hot blast stove can be up to about 10 km^3/h, when the combustion status switches, there will be a drastic impact of the generated BFG on the whole gas system. Although the gas tanks can be treated as the buffer devices, the capacity of about only 60 km^3/h can hardly completely stabilize the gas fluctuation.

For the stability of the system, the trend of the gas tank level has to be maintained in the safety region. When the gas tank level is becoming out of the boundaries, the operators have to change the operational mode of some adjustable units so as to balance the generation and the consumption amounts of BFG. In this study, a typical example to describe the BFG scheduling problem can be briefly addressed as follows. Fig. 2 shows the situation that the gas tank level has reached the upper boundary while the generation amount is continually higher than that of the consumption. In this case, the diffusing towers have to diffuse the excess gas if the scheduling operation is not carried out in time or correctly.

Thus, in order to ensure the efficient production, the scheduling operation must be performed ahead of time. Currently, the total scheduling amount and the corresponding scheduling units, which are referred to the tendency of the gas tank level and the difference amount between the gas generation and the consumption, are roughly determined by the operators. This may cause a large computational cost and uncompleted scheduling.

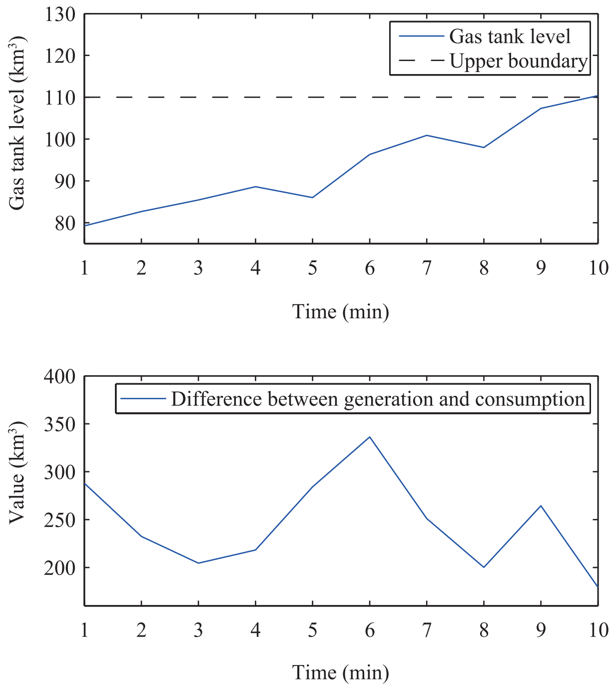


Fig. 2. Condition that gas tank level has reached the upper boundary.

III. Causality Diagram-based Approach

In this paper, a causality diagram-based approach is proposed. The algorithm flow is shown in Fig. 3.

A. Causality Diagram-based Approach

In view of the fact that the scheduling of the BFG system should take the related factors (such as the current gas tank level and the flow difference between gas generation and consumption) into account, the causality diagram of the variables is established as shown in Fig. 4.

In the causality diagram, $y(t)$ denotes the current gas tank level, and the flow difference between gas generation and consumption (FDGGC) can be denoted by $\boldsymbol{X} = [x(t-\theta_i+1),\ldots,x(t)]^T$ where t is the current time and θ_i denotes the time span before t. The gas tank level in the future is denoted by $\boldsymbol{Y} = [y(t+1),\ldots,y(t+\theta_2)]^T$ where θ_2 denotes the time span of the tank level in the future, and $\boldsymbol{Z}$ denotes the corresponding scheduling solution. The probability of the gas tank level reaching the boundary which is represented by P_{adjust} is described as

$$\begin{aligned} P_{\text{adjust}} &= P(\boldsymbol{Y} \not\subset \boldsymbol{R}_{\text{safe}} | y(t), \boldsymbol{X}) \\ &= P(\boldsymbol{Y} > \varepsilon_1 | y(t), \boldsymbol{X}) + P(\boldsymbol{Y} < \varepsilon_2 | y(t), \boldsymbol{X}) \end{aligned} \tag{1}$$

where $\boldsymbol{R}_{\text{safe}} = [\varepsilon_2, \varepsilon_1]$ denotes the safety region, ε_1 and ε_2 are the upper boundary and the lower one, respectively. The gas tank level has to be adjusted when the total probability is not 0.

B. Causal Fuzzy C-Means Algorithm

In the studied practical problem, it is highly possible that two different production conditions could lead to the same flow differences of the gas generation and the consumption, but different scheduling amounts. In contrast, the identical scheduling amounts could also correspond to somewhat various generation-consumption flow differences. In order to obtain the solutions, a causal fuzzy C-means (CFCM) clustering method is proposed, which combines the historical solutions with the corresponding production status. The minimal variance criterion is used as the objective function of the proposed clustering method. Assuming that there are n patterns, and c clusters required, the optimization problem can be formulated by

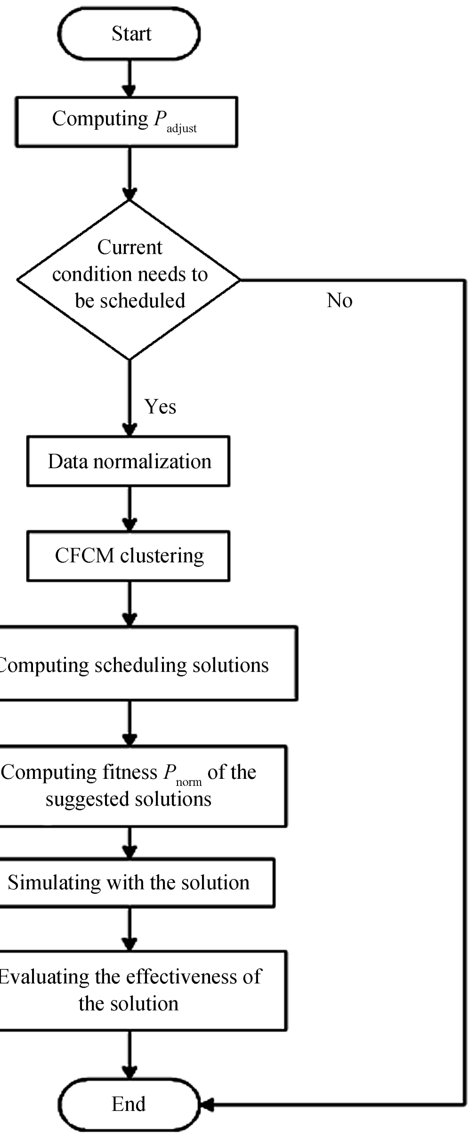


Fig. 3. Algorithm flow of the proposed approach.

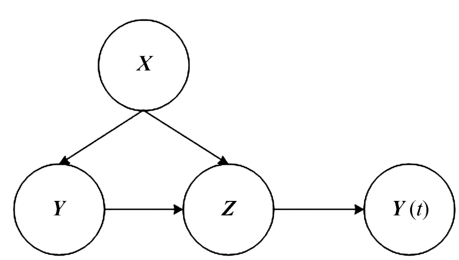


Fig. 4. Causality diagram of the variables.

$$\begin{aligned}&J_{\mathrm{FCM}}=\sum_{j=1}^{n}\sum_{i=1}^{c}u_{ij}^{m}(w_1 d_{ij}+w_2 ACE_{ij})^2\\&\text{s.t.}\qquad \sum_{i=1}^{c}u_{ij}=1,\qquad \forall j=1,2,\ldots,n\\&w_1+w_2=1, u_{ij}\in[0,1], i\in[1,c], j\in[1,n]\end{aligned}\tag{2}$$

where $d_{ij}=\|v_i-x_j\|$ denotes the Euclidean distance between the ith clustering center and the jth data sample, $m\in[1,+\infty)$ is the weighted index, w_1 and w_2 are the weights of the distances. ACE_{ij} is the average causal effect (ACE) between data samples and clustering centers which can be described by

$$\begin{aligned}ACE_{ij}&=ACE(x_j\to v_i)\\&=E(y_i|\mathrm{do}(X=x_j))-E(y_i|\mathrm{do}(X=v_i))\end{aligned}\tag{3}$$

where $\mathrm{do}(\cdot)$ is Pearl's do-operator and y_i is the gas tank level in ith cluster. Equation (3) provides the relationship between the production status of the jth historical solution and the ith clustering, meanwhile measures the expectation differences of the gas tank level after being scheduled by the historical solution and the corresponding clustering center. Let $D_{ij}=w_1 d_{ij}+w_2 ACE_{ij}$, then the clustering centers v_i and the degree of membership u_{ij} can be expressed by

$$v_i=\frac{\sum_{j=1}^{n}u_{ij}^{m}x_j}{\sum_{j=1}^{n}u_{ij}^{m}}\tag{4}$$

$$u_{ij}=\frac{1}{\sum_{k=1}^{c}\left(\frac{D_{ij}}{D_{kj}}\right)^{\frac{-2}{m-1}}}.\tag{5}$$

According to the causality principle, the historical total scheduling amount and the difference flow between gas generation and consumption are taken as the clustering variables, which are regarded as the inputs of the clustering. The outputs, i.e., the clustering centers, correspond to the different production status. The center, whose gas flow difference is close to the current difference, is regarded as the key center, and the total scheduling amount of the center will be taken as the average value of the total amount of the current solution. The historical solutions that belong to the key center are recomposed as follows.

1) If the scheduling amount of one historical solution is close to the average total scheduling amount, then the historical solution is regarded as a suggested solution.

2) If the sum of the amounts of several historical solutions is close to the total scheduling amount, then the group of them is regarded as a suggested solution.

C. Causal Probability of Suggested Solution

The causal probability of each new solution can be computed by considering the running states of the scheduling units. Denoting P_{norm} as the probability of each scheduling solution, we have

$$P_{\mathrm{norm}}=P(\boldsymbol{Y}\subset\boldsymbol{R}_{\mathrm{safe}}|\mathrm{do}(\boldsymbol{Z})),\boldsymbol{Z}\subset\boldsymbol{C}\tag{6}$$

where $\boldsymbol{C}$ represents all the new solutions. Equation (6) can be explained as the probability of the gas tank level back to the safety region with the intervention of the scheduling solution $\boldsymbol{Z}$.

D. Input Delay-based Least Square Support Vector Machine (LSSVM) for Scheduling Solution Verification

In this study, given the characteristic of the BFG system, an input delay-based LSSVM model is established to verify and evaluate the effectiveness of the proposed solution for the gas scheduling.

According to the causal relationship of the BFG system, one can designate the previous gas generation, consumption and the gas tank level as the most relevant factors of the current gas tank level. Therefore, the cumulative flow difference between gas generation and consumption in the previous moments and the previous gas tank level are taken as the inputs of the model, and the tank levels in the future are taken as the outputs. Let τ_1 and τ_2 be the input delay values of the flow difference between the gas generation and the consumption X_1 and the previous tank level X_2, respectively. N denotes the sample length, and the training samples are constructed by

$$X_1=\begin{bmatrix}x(t-\tau_1-N+1)&\cdots&x(t-N+1)\\\vdots&\ddots&\vdots\\x(t-\tau_1-1)&\cdots&x(t-1)\\x(t-\tau_1)&\cdots&x(t)\end{bmatrix}\tag{7}$$

$$X_2=\begin{bmatrix}y(t-\tau_2-N+1)&\cdots&y(t-N)\\\vdots&\ddots&\vdots\\y(t-\tau_2-1)&\cdots&y(t-2)\\y(t-\tau_2)&\cdots&y(t-1)\end{bmatrix}.\tag{8}$$

The regression model established in this paper can be formulated by

$$y=\sum_{j=1}^{m}\boldsymbol{w}_j\varphi(x_j)+b\tag{9}$$

where y is the gas tank level in the future, b is the bias, x_j denotes the jth factor of the inputs, m denotes the number of the factors. The nonlinear mapping function and the weight of the jth factor are represented by φ and $\boldsymbol{w}_j$, respectively. The solving of the input delay-based LSSVM can be regarded as an optimization problem

$$\begin{aligned}&\min\quad J\{\boldsymbol{\omega},b,\boldsymbol{e}\}=\frac{1}{2}\sum_{j=1}^{m}\boldsymbol{\omega}_j^T\boldsymbol{\omega}_j+\frac{\gamma}{2}\sum_{j=1}^{m}\sum_{i=1}^{n}e_{ij}^2\\&\text{s.t.}\quad y_i=\sum_{j=1}^{m}\boldsymbol{\omega}_j^T\varphi(x_i)+b+\sum_{j=1}^{m}e_{ij}\end{aligned}\tag{10}$$

where γ is a penalty coefficient, n denotes the length of the sample and e_{ij} is the fitting error of the jth input. One has to solve the equations set as follows

$$\begin{bmatrix}\boldsymbol{0}&\boldsymbol{1}^T\\\boldsymbol{1}&\sum_{j=1}^{m}K_j+\frac{2}{\gamma}\boldsymbol{I}\end{bmatrix}\begin{bmatrix}b\\\boldsymbol{\alpha}\end{bmatrix}=\begin{bmatrix}0\\\boldsymbol{y}\end{bmatrix}\tag{11}$$

where $K_j = \varphi(x_j)^T \varphi(x_i)$ is the kernel function (usually Gaussian kernel function) of the jth factor, $\boldsymbol{I}$ is an n order unit matrix and $\mathbf{1} = [1, 1, \ldots, 1]^T$. Then, we have the regression model

$$y = \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{m} K_j(x_j, x_{ij}) + b. \tag{12}$$

By repeating the process above, the gas tank levels in the future will be computed to verify the effectiveness of the scheduling solution.

E. Scheduling Evaluation

The effectiveness of the scheduling solution should be evaluated. The definition of the objective function is defined as follows

$$J_{\text{obj}} = \frac{G_d}{G_g} + \frac{T_s}{T_p} \tag{13}$$

where G_d refers to the BFG diffusion and G_g the generation, T_s refers to the moment needed to be scheduled and T_p the predicted length. In (13), both the rate of BFG diffusion in the next period of time and the frequency of the scheduling operation are considered. When the scheduling amount was not consumed by the scheduling units, the gas will be diffused, leading to a waste of energy and environment pollution. The smaller the value of J_{obj} is, the better the scheduling solution is. When the value of J_{obj} is 0, it means that there will be no diffusion and the gas tank level will be in the safety region in a period of time.

IV. Experiments and Analysis

To verify the performance of the proposed method, the practical data obtained in April 2016 from a steel plant in China are employed for the experiments. In the experiments, both circumstances, approaching the lower boundary and the upper one, are considered. The scheduling units include #1−#4 power plants (PP) and low pressure boiler (LPB). The evaluation index mentioned in this paper is regarded as the judgment criteria of the scheduling solutions.

A. Approaching the Lower Boundary

The lower boundary of the gas tank is 60 km^3 in this plant, and the solution has to be provided when the level is becoming lower than it. In this section, a typical practical situation is studied, in which the gas tank level is becoming lower than 60 km^3 and the gas generation amount, meanwhile, is continually lower than the consumption amount.

Initially, one can compute P_{adjust} according to the current production status. In this situation, it equals to 0.435 6, which means the scheduling is required. For the sake of calculating the total scheduling amount, the proposed CFCM is used for clustering all the 43 samples selected from the historical data. The number of clusters is 5, which is decided by expert knowledge. The clustering results based on CFCM and normal FCM are shown in Fig. 5, where X label is the flow difference between gas generation and consumption and Y label is the total scheduling amounts. The value of w_1 and w_2 are set to 0.7 and 0.3, respectively.

Comparing to the current situation with the clustering centers in Fig. 5, the average value of the total amount is -100 km^3 and -50 km^3 respectively. The running states of the scheduling units are shown in Table I where t represents the current moment.

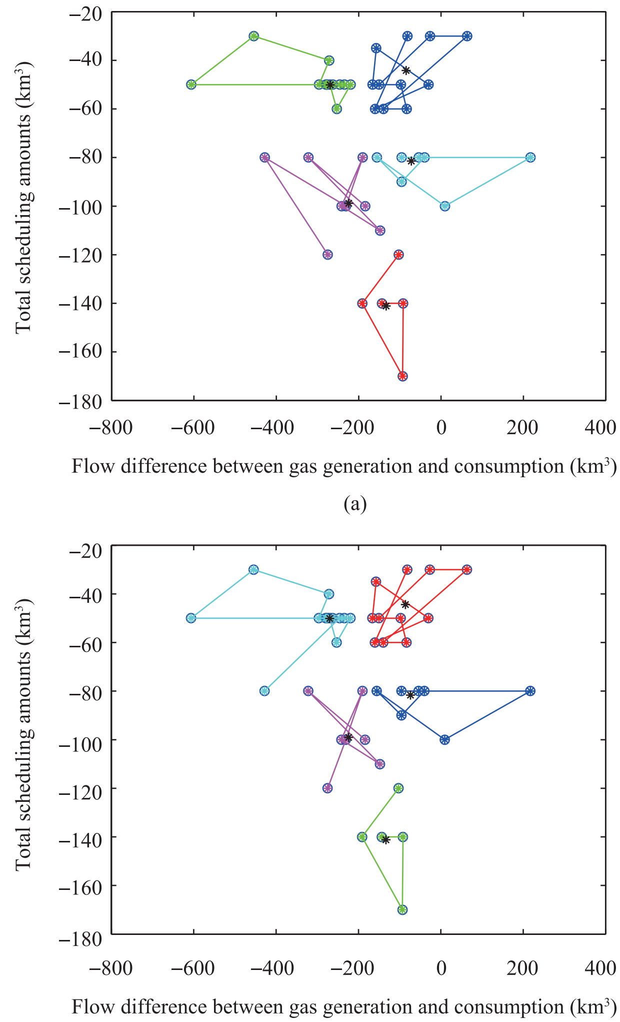


Fig. 5. Clustering results of the historical solutions. (a) Clustering result with CFCM. (b) Clustering result with normal FCM.

TABLE I
Current Running States of the Scheduling Units

Time	#1PP	#2PP	#3PP	#4PP	LPB
$t-9$	0.70	0.85	150.07	565.22	48.45
$t-8$	0.58	0.82	149.74	565.68	48.23
$t-7$	0.61	0.92	150.39	565.80	48.47
$t-6$	0.70	0.92	149.94	566.41	48.22
$t-5$	0.61	0.89	149.29	566.44	48.09
$t-4$	0.67	0.85	149.33	566.44	48.32
$t-3$	0.61	0.92	150.57	565.74	48.95
$t-2$	0.67	0.95	149.86	565.71	49.22
$t-1$	0.61	0.85	149.48	566.61	48.95
t	0.61	0.92	146.34	565.80	47.40

Table I shows that #1PP and #2PP are at stopping stage, therefore the maximal adjustment ranges of #1PP, #2PP and #3PP according to the operation load are 0 km^3, 0 km^3, and

$-150\,\mathrm{km}^3$ respectively. In order to ensure the regular running of #4PP and LPB, the maximal adjustable range of them are $-100\,\mathrm{km}^3$ and $0\,\mathrm{km}^3$, respectively. Therefore, the new solutions with special scheduling units and the corresponding amount and the causal probability can be shown in Table II and Table III.

TABLE II
SOLUTIONS OBTAINED BY CFCM

Solution number	Total amount (km^3)	Solution (amount(km^3))	Causal probabilities
1	-120	#1PP (-80), LPB (-40)	0
2	-110	#3PP (-50), #1PP (-60)	0
3	-100	#3PP (-100)	0.15
4	-100	#2PP (-100)	0
5	-80	#3PP (-50), #2PP (-30)	0
6	-80	#3PP (-80)	0.85
7	-80	#3PP (-50), #4PP (-30)	0

TABLE III
SOLUTIONS OBTAINED BY NORMAL FCM

Solution number	Total amount (km^3)	Solution (amount(km^3))	Causal probabilities
1	-80	#3PP (-50), #4PP(-30)	0
2	-60	#1PP (-60)	0
3	-50	#1PP (-50)	0
4	-50	#2PP (-50)	0
5	-50	#3PP (-50)	0.9
6	-50	#4PP (-50)	0.45
7	-40	#3PP (-40)	0.75
8	-30	#3PP (-30)	0.6

According to Table II, the sixth solution is regarded to be the best one, while the fifth solution is the best in Table III. As to the human experience, the solution is to reduce the usage of #3PP for $50\,\mathrm{km}^3$ first, then reduce the usage of #4PP for $30\,\mathrm{km}^3$ after 10 minutes. The comparison of the results for 60 minutes in the future by the three methods, i.e., the proposed one, the generic FCM and the human experiences applied in the current production is illustrated in Fig. 6. The gas tank level can be pulled back to the safety region by the proposed method and can last for at least 60 minutes. The FCM based method can also pull the gas tank level back to the safety region, but it may lack stability. The human experience method results in a second operation due to the uncompleted scheduling in the first time, and the accuracy is apparently the lowest. The scheduling evaluation of the three methods are shown in Table IV.

In Table IV, we can see the evaluation result clearly. The proposed method makes the gas tank no diffusion and no second scheduling, while the other two methods contain second, even third operation in the next 60 minutes.

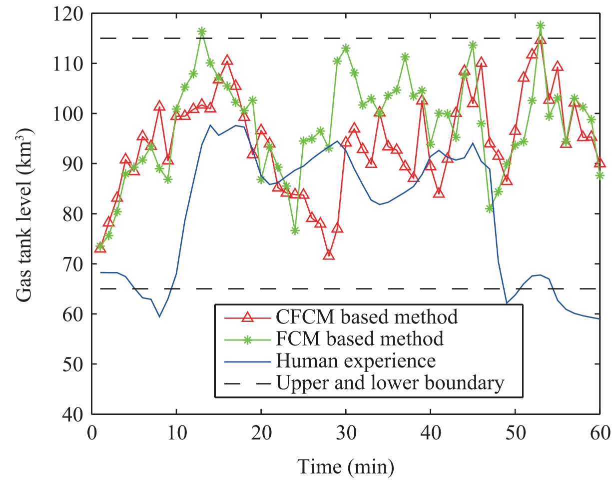


Fig. 6. Comparison of the three scheduling solutions.

TABLE IV
SCHEDULING EVALUATION OF THE THREE METHODS

Method	Objective value
Proposed method	0
FCM based method	0.03
Human experience	0.2

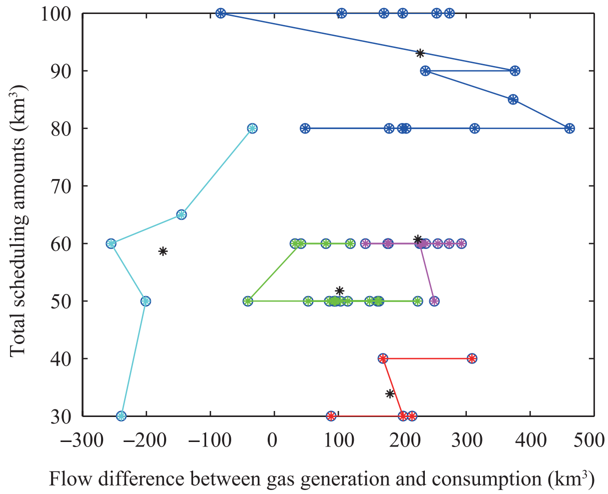
B. Approaching the Upper Boundary

The upper boundary of the gas tank is $115\,\mathrm{km}^3$. A typical situation is here studied, in which the tank level is becoming higher than the boundary and the gas generation, however, continually higher than the consumption.

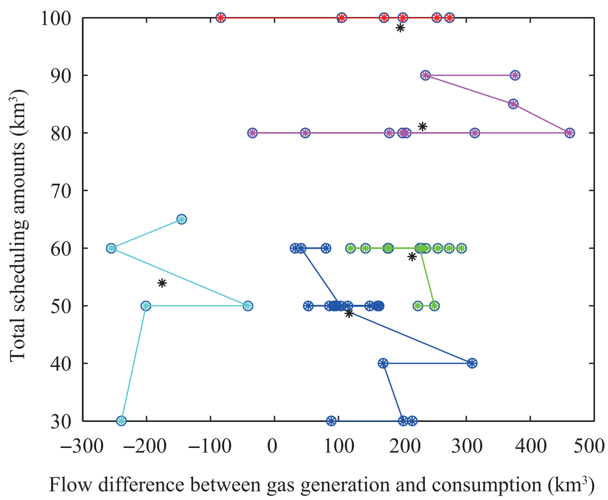
As the value of P_{adjust} is above zero, the result of the two clustering methods using 54 historical samples is shown in Fig. 7, where the average of the total amount is $95\,\mathrm{km}^3$ and the running states of the scheduling units are shown in Table V, where t represents the current moment.

Table V lists that #1PP is at stopping stage, so the maximal adjusting ranges of the first three scheduling units are $60\,\mathrm{km}^3$, $0\,\mathrm{km}^3$, $50\,\mathrm{km}^3$, respectively. In order to ensure the regular running of #4PP and LPB, the maximal adjustment range of them are $0\,\mathrm{km}^3$ and $50\,\mathrm{km}^3$ respectively. Therefore, the new solutions with special scheduling units, the corresponding amount and the causal probability are shown in Table VI and Table VII.

According to Table VI, the first solution is the best one, while in Table VII, the first solution is also the best one. The human experience is to increase the usage of #3PP for $50\,\mathrm{km}^3$. The comparison of the results for 60 minutes in the future is shown in Fig. 8 and the scheduling evaluation of the three methods is listed in Table VIII, where the proposed method makes the gas tank no diffusion and no second scheduling, while the FCM based method causes the level lower than the lower boundary. As compared to the human experience, the proposed method produces the scheduling amount in advance, which can prevent the gas tank level from going higher than the upper boundary.



(a)



(b)

Fig. 7. Clustering results of the historical solutions. (a) Clustering result with CFCM. (b) Clustering result with normal FCM.

TABLE V
CURRENT RUNNING STATES OF THE SCHEDULING UNITS

Time	#1PP	#2PP	#3PP	#4PP	LPB
$t-9$	0.61	84.08	197.59	746.67	28.28
$t-8$	0.61	84.59	198.26	745.60	24.64
$t-7$	0.61	84.78	198.49	746.47	20.72
$t-6$	0.64	85.39	200.52	746.64	17.24
$t-5$	0.64	85.57	201.09	746.30	14.73
$t-4$	0.64	84.96	199.52	745.54	13.68
$t-3$	0.61	84.35	197.96	746.61	12.93
$t-2$	0.64	84.38	198.71	747.25	12.72
$t-1$	0.58	84.63	199.38	746.47	12.26
t	0.61	84.66	200.05	746.04	12.28

TABLE VI
SOLUTIONS OBTAINED BY CFCM

Solution number	Total amount (km^3)	Solution (amount (km^3))	Causal probabilities
1	50	#1PP (+50)	0.6
2	50	#2PP (+50)	0
3	50	#3PP (+50)	0.55
4	50	#4PP (+50)	0

TABLE VII
SOLUTIONS OBTAINED BY NORMAL FCM

Solution number	Total amount (km^3)	Solution (amount (km^3))	Causal probabilities
1	60	#1PP (+60)	0.55
2	60	#2PP (+60)	0
3	60	#3PP (+60)	0
4	60	#4PP (+60)	0

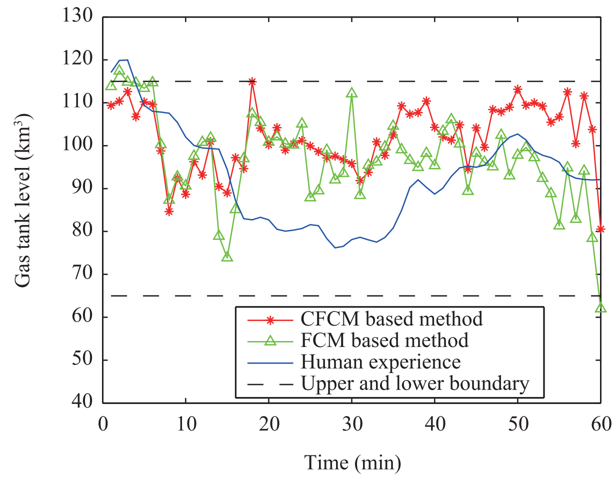


Fig. 8. Comparison of the three scheduling solutions.

TABLE VIII
SCHEDULING EVALUATION OF THE THREE METHODS

Method	Objective value
Proposed method	0
FCM based method	0.03
Human experience	0.05

V. CONCLUSION

Considering that the scheduling solution is fairly critical in BFG system, a causality diagram-based scheduling approach is proposed in this study to provide the total scheduling amount and allocate it to the reasonable scheduling units. The proposed method takes the running states of the scheduling units into account to ensure the practicability of the suggested scheduling solution. An evaluation method is also proposed to evaluate all the suggested scheduling solutions. To verify the effectiveness of the proposed method, a set of experiments are conducted based on the real-world data from an iron and steel enterprise. The experimental results illustrate that the proposed method exhibits higher accuracy in comparison with the human experience based approach and provides an effective guidance for energy balancing.

REFERENCES

[1] J. Zhao, W. Wang, K. Sun, and Y. Liu, "A Bayesian networks structure learning and reasoning-based byproduct gas scheduling in steel industry," *IEEE Trans. Automat. Sci. Eng.* vol. 11, no. 4, pp. 1149−1154, Oct. 2014.

[2] X. P. Zhang, J. Zhao, W. Wang, L. Q. Cong, and W. M. Feng, "An optimal method for prediction and adjustment on byproduct gas holder in steel industry," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 4588−4599, Apr. 2011.

[3] J. Zhao, Q. L. Liu, W. Wang, W. Pedrycz, and L. Q. Cong, "Hybrid neural prediction and optimized adjustment for coke oven gas system in steel industry," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 3, pp. 439−450, Mar. 2012.

[4] H. N. Kong, E. S. Qi, S. G. He, and G. Li, "MILP model for plant-wide optimal by-product gas scheduling in iron and steel industry," *J. Iron Steel Res. Int.*, vol. 17, no. 7, pp. 34−37, Jul. 2010.

[5] H. Ru, F. Gao, Y. F. Xu, and X. H. Guan, "Online strategy for scheduling by-product gas supply," in *Proc. 2013 IEEE International Conference on Information and Automation (ICIA)*, Yinchuan, China, pp. 1246−1251.

[6] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York, USA: Plenum Press, 1981.

[7] J. Zhao, Z. Y. Han, W. Pedrycz, and W. Wang, "Granular model of long-term prediction for energy system in steel industry," *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 388−400, Feb. 2016.

[8] H. B. Cao, H. W. Deng, and Y. P. Wang, "Segmentation of M-FISH images for improved classification of chromosomes with an adaptive fuzzy C-means clustering algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 1, pp. 1−8, Feb. 2012.

[9] X. Y. Tang, J. Zhao, C. Y. Sheng, and W. Wang, "Long term prediction for generation amount of Converter gas based on steelmaking production status estimation," in *Proc. 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Beijing, China, pp. 1088−1095.

[10] J. Pearl, *Causality: Models, Reasoning, and Inference*. Cambridge, UK: MIT Press, 2000.

[11] M. Erwig, and E. Walkingshaw, "Causal reasoning with neuron diagrams," in *Proc. 2010 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, Leganes, Philippines, 2010, pp. 101−108.

[12] F. Thoemmes and K. Mohan, "Graphical representation of missing data problems," *Struct. Equat. Model.*, vol. 22, no. 4, pp. 631−642, 2015.

[13] K. Mohan and J. Pearl, "Graphical models for recovering probabilistic and causal queries from missing data," in *Proc. 27th International Conference on Neural Information Processing Systems*, Montreal, Canada, 2014, pp. 1520−1528.

[14] H. H. Dai, S. Keble-Johnston, and M. Gan, "Reliable knowledge discovery with a minimal causal model inducer," in *Proc. 2012 IEEE 12th International Conference on Data Mining Workshops (ICDMW)*, Brussels, Belgium, pp. 629−634.

[15] F. M. Chen and D. Zhang, "Combining a causal effect criterion for evaluation of facial attractiveness models," *Neurocomputing*, vol. 177, pp. 98−109, Feb. 2016.

Feng Jin graduated from Northeasten University at Qinhuangdao, China, in 2012. He is currently pursuing the Ph.D. degree with the School of Control Science and Engineering, Dalian University of Technology, Dalian, China. His research interests include data based prediction and scheduling.

Jun Zhao (M'10) graduated from Dalian Jiaotong University, China, in 2003. He received the Ph.D. degree in engineering from the Dalian University of Technology, in 2008. He is currently a Professor with the School of Control Science and Engineering, Dalian University of Technology. His current research interests include industrial production scheduling, computer integrated manufacturing, intelligent optimization, and machine learning.

Chunyang Sheng received the Ph.D. degree in control theory and control engineering, Dalian University of Technology, in 2014. He worked as a post-doctoral at the School of Computer Science and Technology, Dalian University of Technology from 2015 to 2016 . He currently works as a Lecturer in Shandong University of Science and Technology. His research interests include machine learning, data based prediction scheduling and optimization.

Wei Wang (SM'13) received the B.S., M.S., and Ph.D. degrees from Northeastern University, Shenyang, China, in 1982, 1986, and 1988 respectively, all in industrial automation. He was a Post-Doctoral Fellow with the Division of Engineering Cybernetics, Norwegian Science and Technology University, Trondheim, Norway, from 1990 to 1992, a Professor and Vice Director of the National Engineering Research Center of Metallurgical Automation of China, Beijing, China, from 1995 to 1999, and a Research Fellow with the Department of Engineering Science, University of Oxford, Oxford, U.K., from 1998 to 1999. He is currently a Professor with the School of Control Sciences and Engineering, Dalian University of Technology, China. His current research interests include adaptive controls, computer integrated manufacturing, and computer controls of industrial processes.