This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

Controllable-domain-based fuzzy rule extraction for copper removal process control

Bin Zhang, Chunhua Yang, *Senior Member*, IEEE, Hongqiu Zhu, Peng Shi, *Fellow, IEEE*, Weihua Gui, *Senior Member*, *IEEE*

***Abstract*—In copper removal process control, the commonly used technique is the so called rule based control, which is largely dependent upon the operators' experience, likely leading to unstable process production due to each individual's characters and favors. In this paper, to enhance the effectiveness of process control, a controllable-domain-based fuzzy rule extraction strategy is proposed. New definitions of representative controlled samples are introduced, by which the input variable space is divided several controllable domains by applying positive and unlabeled learning algorithm. Also, the unreasonable removed and the controllable domains are accordingly determined. Then, support vector machine method is employed to extract fuzzy control rules for different domains. Finally an industrial experiment is presented to demonstrate the effectiveness and advantages of the developed new design scheme.**

***Index Terms*—fuzzy logic, support vector machine (SVM), rule extraction, positive and unlabeled learning (PU learning), copper removal**

I. Introduction

COPPER removal process is the first stage of zinc purification to remove most of the impurity copper from leaching sulfate solution and reserve an appropriate amount of copper ions as activator used in the next removal stage [1]. During this process, the adjustment of zinc powder determines the stability of the removal efficiency and production quality [2]. An excessive amount results insufficient activator for the next stage lowering the successive removal efficiency; whereas, an inadequate amount leads excess copper ions left in the solution decreasing the electrolyte efficiency [3]. In practical, the amounts of zinc powder are determined manually according with individual experiences and rules [4]. Due to the complexity and uncertainty of the process, the manual adjustments are sometimes improperly and usually come in late [5]. The poor control easily makes the outlet impurity concentration beyond the required range and the whole process fluctuant [6]. Therefore, it is essential to find a set of effective control rules to improve the production quality and stability.

This work is supported by the National Natural Science Foundation of China (61273185), the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (61621062), the innovation-driven plan in Central South University, the 111 project of China (B17048) and Australian Research Council (DP140102180, LP140100471, LE150100079).

B. Zhang is now with Guangdong University of Technology, 510006, China; and was with Central South University, Changsha 410083, China, (e-mail: zhangbin_csu309@163.com).

C. Yang (the corresponding author) is with Central South University, Changsha 410083, China (e-mail: ychh@csu.edu.cn).

H. Zhu is with Central South University, Changsha 410083, China (e-mail: hqcsu@mail.csu.edu.cn).

P. Shi is with the College of Automation, Harbin Engineering University, Harbin 150001, China, with the College of Engineering and Science, Victoria University, Melbourne, Vic. 8001, Australia, and with the School of Electri-cal and Electronic Engineering, University of Adelaide, Adelaide, S.A. 5005, Australia (e-mail: peng.shi@adelaide.edu.au).

W. Gui is with Central South University, Changsha 410083, China (e-mail: gwh@ csu.edu.cn).

There have been many techniques to express control rules, such as expert systems [7], decision rules [8], fuzzy rules [9] and model predictive control [10]. Among these techniques, fuzzy rules supplied attract lots of researchers for that fuzzy rules are easily interpreted and comprehended. But setting up rules is a time-consuming and expensive process [11]. To overcome this shortage, various approaches are proposed to extract fuzzy rules automatically from numerous data [12]-[26]. These approaches are basically divided into three classes: the first family of approaches applies partition methods, and the second implements clustering, and the last uses hybrid methods [15]. The first family [16]-[18] mostly classifies the input space into several subspaces, and then translates the subspaces into fuzzy rules. In the second family [19]-[20], the instances are gathered into homogeneous groups and the rule are then associated into fuzzy rules. The last family [21]-[22] mainly combines different intelligent algorithms to extract rules. For example, genetic algorithm is accomplished with CHC evolutionary model to select rules for dealing with complex search spaces [23]. For the effectiveness of rule extraction, it has been successfully used in different fields. Gao et al. presented a fuzzy-based SVM classification algorithm to determine the controllable bound for the blast furnace [24]. Prado et al. developed a knowledge acquisition with a swarm-intelligence approach for fuzzy-rule evolution [25]. Hua et al. proposed a memoryless control scheme based on T-F fuzzy approach for the control of a networked interconnected chemical reactor system [26]. Previous studies provide exciting advices for describing the control of the copper removal process with a series of dynamical and interpretable rules.

Notably, the effects of industrial data qualities on the effectiveness of the extracted rules are rarely considered in the previous approaches. In a continuous chemical process, due to the cumulativeness and inertia of big reactors, the control effect mainly relies on process conditions and inlet parameters rather than solely depends on the control variables. Accordingly, a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

proper control rules should take consideration of industrial data from different classes of process condition and inlet characteristics. Multi-classes classification might be a good choice in this field, e.g. multi-classes SVM [27], multi-class fuzzy classifier [28], and decision tree based multi-class classifier [29]. In these methods, samples should be labelled before classification, but labelling sample manually is labor-consuming and time-costing. Additionally, for the variation and complexity in industrial process, only typical samples could be identified during practical production; while the number of the rest unidentified samples is considerable.

To deal with this setting, PU learning has been introduced in that it could incorporate information from both labeled positive and unlabeled samples [30]. Most PU learnings are based on two-step strategy: identify reliable negative samples from unlabeled set and classify labeled positive and negative samples [31]. Some PU learnings apply classifiers iteratively to increase negative samples [32-34]. S-EM algorithm selects reliable negative samples by using positive samples as spies and iteratively training classification [32]. But the computation cost of iterative classification is expensive when dealing with huge number of industrial data. Some algorithms identify negative samples by the probability estimation over positive and unlabeled samples [35-37]. In the Naïve Bayesian classifier based PU learnings, the negative example model is built by statistically removing the effect of positive examples from the unlabeled example model [35, 36]. In a PU learning proposed in Ref. [37], the classifier is built based on the probability estimation of labeling positive sample. But these required prior probabilities are usually unavailable in practical industrial process. Other PU learniags select negative samples by clustering [38, 39]. Ienco and Pensa introduced a distance learning method to calculate the similarities between categorical samples for reliable negative samples detection [39]. Clustering based PU learnings becomes a good choice for our problem because of their simplicity, understandability and less computation. Most PU learnings are improved to solve problems in the fields as text classification [40, 41], web work [33], which are different from the rule extraction for a chemical process. In the copper removal processes, the irrational-operated samples are unavoidable mixed in the learning set decreasing the accuracy of extracted rules. To learn rules reasonably for copper removal, PU learning should be improved for both labeling undefined samples and rejecting irrational-operated samples effectively.

In short, to improve the production stability, a set of appropriate rules extracted from operation data, rather than set by individual experiences, are urgent needs in copper removal control. But due to the variety and abundance of industrial data, they are hard to be labeled manually during rule extraction; meanwhile, the irrational-operated data reduce the quality of the extracted rules. To solve these problems, we proposed a control rule extraction strategy for copper removal process, which can label and classify industrial samples with their control effect; and meanwhile reject unreasonable samples to avoid the misleading operation; at last extract rules to adjust zinc amount automatically instead of experiences based operation.

II. Related Work

Rule extraction explains massive data in an understandable and explainable form with series rules[42]. Classification is typical rule exaction algorithm. However, the labeling is usually expensive and unavailable for all the sample during classification, semi-supervisor classification (e.g. PU learning) becomes a good choice during this field.

Numerous PU learning algorithms have proposed recent years. One focus of the algorithm is to extract negative samples from unlabeled set reliably. Usually, self-training technique is adapt to extract labeled samples iteratively with base learner [43]. Zhang and Zuo select negative samples from unlabeled dataset by ranking their similarity to positive samples based on k-NN algorithm [44]. The method is simple and effective but lack of approximation method to determine the value of selected numbers. To avoid enlarging a special labeled set with a lower confidence level, PU learnings with stopping criteria have come out [45]. González et al. use the graphical analysis based minimum distances as stopping criteria in each iteration produced by k-NN algorithm [46], with those criteria the precision of PU learning is improved. But the outliers and noise are inevitable in natural data set, which decreases the running precision.

Traditionally, the noise and outliers are considered only existed in the unlabeled set. The noise tolerance is improved by tuning inner parameters but without structure changes [39, 47]. More generally, the contaminated samples existed in different sets. PU learning is improved to apply in outlier detection [48], where the reliable negative samples are firstly extracted by k-NN algorithm, and the outliers are identified from both positive and negative sets by fuzzy clustering. To solve the potential contamination in positive samples in PU learning, Claesen et al. proposed the robust ensemble of SVMs [49]. This approach built an ensemble model by applying a bagging strategy where both positive and unlabeled data are resampled to construct base model training set. The experiments show it is more robust against false positive than other classification methods but with a weakness of having lots of hyperparameters needed to be set. In these works, the contaminated samples are removed, usually at one time, according with their similarities between other samples. But in the studied process, the irrational-operated samples should be identified by compared with the good performed samples whose definitions rely on production requirement. And the identifying process might be carried out iteratively based on different control effect levels. Jain et al. also considered the learning problem with noisy positive and unlabeled samples [50]. They applied prior class estimation to model the noise in the positive labels for high-dimensional data, which shows good performances with real-life data. Probability estimation is useful for eliminating noise in samples. But the occurrence and definition of irrational-operated samples rely on the individual characters and experience diversity, which limits the application of these PU learnings in the specific industrial process.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

III. Preliminaries

A. *Support vector machine*

SVM, a supervised learning model for classification, is developed by the principle of structural risk minimization [38]. SVM model is achieved by finding an optimal hyperplane in a high dimensional feature space that separates different classes by the greatest extent [39].

Given a training set $\mathbf{S}=\{\mathbf{s}_i, y_i\}$, $i=1,\dots,n$, $\mathbf{s}_i \in \Re^m$, $y_i \in \{-1,1\}$ .The SVM finds the optimal separating hyperplane with the largest margin. The separating hyperplane in the linearly separable case is represented as Eq. (1):

$$\begin{cases} \mathbf{ws}_i + \mathbf{b} \geq 1 & \text{for} \quad y_i = 1 \\ \mathbf{ws}_i + \mathbf{b} \leq 1 & \text{for} \quad y_i = -1 \end{cases} \tag{1}$$

Finding a maximum separating margin in the linearly separable case could be treated as a constrained optimization problem, as Eq. (2).

$$\begin{aligned} &\min \quad \frac{1}{2}\|\mathbf{w}\|^2 \\ &s.t. \quad \forall \mathbf{s}_i \in \mathbf{S} \quad y_i(\mathbf{wx}_i + \mathbf{b}) \geq 1 \end{aligned} \tag{2}$$

The hyperplane decision function is thus written as Eq.(3).

$$f(\mathbf{s}) = sign\left(\sum_{j=1}^{SV} a_j y_j (\mathbf{s}_j, \mathbf{s}) + \mathbf{b}\right) \tag{3}$$

where $(\mathbf{s}_j, \mathbf{s})$ is the inner product of $\mathbf{s}_j$ and $\mathbf{s}$.

The non-linear classification problem is solved linearly in the high-dimensional feature space based on kernel functions. The decision function of SVM is changed to Eq. (4).

$$f(\mathbf{s}) = sign\left(\sum_{j=1}^{SV} a_j y_j \,\mathrm{K}(\mathbf{s}_j, \mathbf{s}) + \mathbf{b}\right) \tag{4}$$

where $\mathrm{K}(\bullet,\bullet)$ denotes the kernel function.

The above binary SVM classifier can be extended to deal with a multi-class classification problem by two typical methods. The first method is to integrate several binary SVM into a multi-class classifier [40]; and the last is to design a multi-class classifier as a constrained optimization problem [41].

B. *PU learning based on kNN and SVM*

A PU learning algorithm based on kNN and SVM is as shown in applied Algorithm 1. It applies kNN algorithm to search k nearest samples and k nearest positive samples for each unlabeled sample, and then the average similarities between those neighbors and the unlabeled sample are calculated. After comparing those similarities, the unlabeled sample can be identified whether it is more close to positive set. The similarity function used in this algorithm is chosen according with the requirements in actual application.

Algorithm 1 PU learning based on kNN and SVM

Input $\mathbf{P}^0$, the original positive samples set
$\mathbf{U}$, unlabeled samples set
N_K, the number of the nearest neighbors
N_p, the number of the nearest positive samples
θ, threshold

Output $\mathbf{N}$, the negative samples set
$\mathbf{P}$, the positive samples set

1. **For** each unlabeled sample u_i
   Select k nearest neighbors of u_i
   If there are any positive samples in these neighbors
   Compute the similarities between u_i and the neighboring positive samples: $sim(u_i, v_{i,j}^0)$
   Compute the similarities between u_i and the neighboring negative samples: $sim(u_i, u_{i,j}^0)$
   Else
   Select N_p nearest positive samples of u_i
   Compute the similarities between u_i and the neighboring samples: $sim(u_i, u_{i,j}^1)$
   Compute the similarities between u_i and the neighboring positive samples: $sim(u_i, v_{i,j}^1)$
   End if
   End for
2. Calculate the average similarities for u_i: $\overline{sim}(u_i, \mathbf{v}_\mathbf{i}^\mathbf{0})$, $\overline{sim}(u_i, \mathbf{u}_\mathbf{i}^\mathbf{0})$, $\overline{sim}(u_i, \mathbf{v}_\mathbf{i}^1)$, $\overline{sim}(u_i, \mathbf{v}_\mathbf{i}^1)$
3. **If** $\overline{sim}(u_i, \mathbf{v}_\mathbf{i}^\mathbf{0}) / \overline{sim}(u_i, \mathbf{u}_\mathbf{i}^\mathbf{0}) \geq \theta$ or $\overline{sim}(u_i, \mathbf{v}_\mathbf{i}^1) / \overline{sim}(u_i, \mathbf{u}_\mathbf{i}^1) \geq \theta$
   $u_i \in \mathbf{P}$
   Else
   $u_i \in \mathbf{N}$
   End if
4. Apply SVM with the $\mathbf{P}$ and $\mathbf{N}$ to train a classifier.

IV. Controllable-domain-based fuzzy rule extraction for copper removal

In this section, the copper removal process and the preparation of the industrial data are firstly introduced, and then the proposed controllable-domain-based fuzzy rule extraction for copper removal process is described. The organization of this section is shown as Fig. 1.

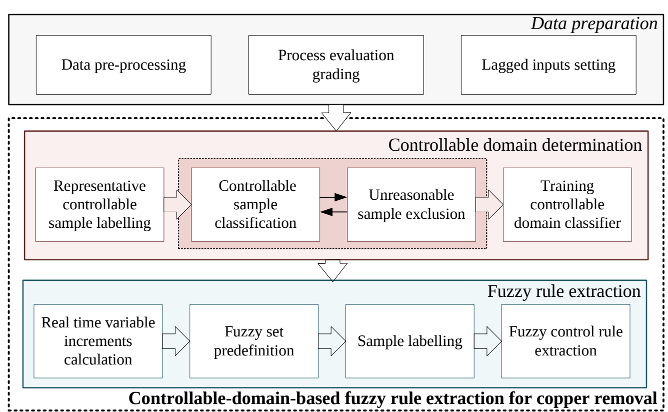

Fig. 1. Organization of controllable-domain-based fuzzy rule extraction for copper removal process.

During the data preparation, pre-processed industrial parameters, lag steps and process evaluation results are provided. The controllable-domain-based fuzzy rule extraction strategy is divided into two parts. One part is controllable domain determination where an improved learning algorithm is applied to classify industrial data space into several controllable domains with the proposed definitions of representative control effect. The other part is fuzzy rule extraction where a series of fuzzy rules are extracted for each domain based on a fuzzified SVM for copper removal process control.

A. Process description and data preparation

In this paper, the experimental data are collected from an actual copper removal process of zinc hydrometallurgy in China, as shown in Fig. 2. The leached zinc sulfate solution flows into a series of continuous stirred tank reactors, and zinc powder is continuously added in the reactors to deposit copper

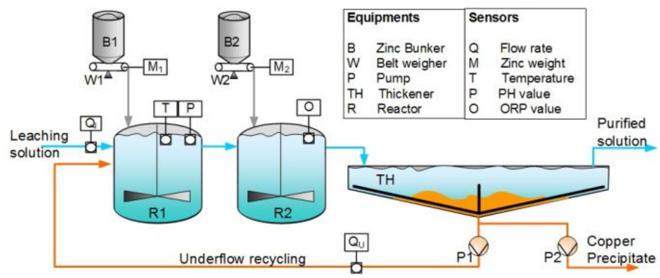

Fig. 2. Flow diagram of copper removal in the zinc purification process.

ions. The purified solution is sent to a thickener for solid-liquid separation. The clean solution is pumped to the cobalt removal stage, and a portion of the precipitate is returned to the first reactor. In the copper removal process, the process variables are mostly measured, stored, and monitored online. But the impurity concentrations are sampled manually and measured in lab every two hours. The inlet and outlet characteristics of the copper removal process are presented in Table I.

Due to the cumulative effect brought from the big chemical reactor, there are large delays in the copper removal process. This delay time should be considered when assessing the control effect and extracting the control rules, which can be roughly calculated as $T_D = \alpha v / x_1$, where v and x_1 denote the volume of the reactor and the outlet flow rate, respectively; and α is a coefficient. Then the times of lag taps of the industrial variables could be calculated as $\Gamma = T_D / T_S$, where, T_S is the sample intervals of the real time industrial variables. Table II presents the candidate inputs for the rule extraction strategy which consist of the off-line measurements, the current and lagged values of the real time variables. These lagged inputs are expressed by the backward shift operator d^{-i}, which is defined as $d^{-i}x_j(\mathrm{t}) = x_j(\mathrm{t-i})$. The symbol $x_j(\mathrm{t})$ denotes the input variables, and d^0 indicates $d^0 x_j(\mathrm{t}) = x_j(\mathrm{t})$. Because the temperature, pH value, solid content and underflow rate are mostly kept constant; to simplify the rule extraction strategy, these variables are removed from the list of the candidate inputs.

In the copper removal process, oxidation-reduction potential (ORP) is usually used to monitor process condition and control effect on line, it reflects qualitatively the movements of copper concentration in reactors. To eliminate the uncertainty of ORP,

TABLE I
THE INLET AND OUTLET CHARACTERISTIC OF THE COPPER REMOVAL PROCESS (OVER 300 DAYS)

Variable name	Inlet	Outlet
Flow rate of Cu2SO4 solution [m3/h]	100-310	100-310
Flow rate of underflow [m3/h]	28-31	--
Temperature [℃]	62-66	62-66
Solid content [g/m3]	1.4-1.5	--
pH [--]	3.9-4.2	--
ORP [mV]	--	30-150
Concentration of Cu2+ [g/L]	0.7-2.1	0.1-0.5

TABLE II
LIST OF CANDIDATE INPUT VARIABLES FROM THE COPPER REMOVAL PROCESS

Variable name	Measure way	Symbol	Input variables
Flow rate [m3/h]	On line	x_1	$d^0, d^{-1}, ..., d^{-\Gamma}$
ORP [mV]	On line	x_2	$d^0, d^{-1}, ..., d^{-\Gamma}$
Zinc powder amount (1#) [Kg/h]	On line	x_3	$d^0, d^{-1}, ..., d^{-\Gamma}$
Zinc powder amount (2#) [Kg/h]	On line	x_4	$d^0, d^{-1}, ..., d^{-\Gamma}$
Inlet concentration of Cu^{2+} [g/L]	Off line	x_5	d^{-1}
Outlet concentration of Cu^{2+} [g/L]	Off line	x_6	d^{-1}

it should be translated into process evaluation grade by using fuzzy logic technique [3]: $g = f_{EVAL}(x_2, \mathrm{d}x_2)$, where x_2 and $\mathrm{d}x_2$ denote ORP and its first derivative, respectively; g and $f_{EVAL}(\bullet)$ denote the evaluation grade and a fuzzy function translating ORP into g, respectively.

B. Controllable domain determination

In the copper removal process, due to the differences between individual experiences, distinct operations are performed under similar conditions resulting good or bad outlet copper concentrations. If the data produced by all of these operations are used to extract fuzzy rules, the control based on the extracted rules would be fail to reach an expected control effect. Hence, a reasonable classification is needed before rule extraction for copper removal.

1) Representative controllable sample labelling

Denote the training industrial data set as $\mathbf{I} = \{\mathbf{I}_R, \mathbf{I}_M\}$, $\mathbf{I}_R \in \mathbf{R}^{n\times 4(\Gamma+1)}$, $\mathbf{I}_M \in \mathbf{R}^{n\times 4}$ which consists of two parts: off-line measured concentrations $\mathbf{I}_M = \{\mathbf{x}_5, \mathbf{x}_5\mathbf{d}^{-1}, \mathbf{x}_6, \mathbf{x}_6\mathbf{d}^{-1}\}$ and real time delay variables $\mathbf{I}_R = \{\mathbf{x}_1\mathbf{d}^0, ..., \mathbf{x}_1\mathbf{d}^{-\Gamma}, ..., \mathbf{x}_4\mathbf{d}^0, ..., \mathbf{x}_4\mathbf{d}^{-\Gamma}\}$. Firstly, typical samples should be identified from the set $\mathbf{I}$. In a continuous process, the movements of control results during a proper time are useful to evaluate control operation compared with the result at a certain time. Accordingly, by observing the variation of the evaluation grade g (calculated by the time series of x_2) during the observation time, several definitions of control effect in different level are proposed for copper removal.

Definition 1: If the process evaluation grade $g(t)$ keeps

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

relatively stable towards the optimal range from t_0 to $t_0 + T_D$, which could be described as Eq. (5):

$$\begin{cases} g(t) \in [g_O^L, g_O^U],\ t \in [t_0 - \mathrm{T}_D, t_0] \\ |g'(t)| < (g_O^L + g_O^U)/\mathrm{T}_D,\ t \in [t_0 - \mathrm{T}_D, t_0] \end{cases} \tag{5}$$

where g_O^L and g_O^U denote the lower and upper limits of the optimal ranges of the evaluation grades, respectively; and $g'(t)$ denotes the first derivate of the grade, or when $g(t)$ reach the optimal range from the other ranges during the period $[t_0 - \mathrm{T}_D, t_0]$ without fluctuations and derivate to another un-optimal range, which could be described as Eq. (6):

$$\begin{cases} g(t_0 - \mathrm{T}_D) \notin [g_O^L, g_O^U] \\ g(t_0) \in [g_O^L, g_O^U] \\ g'(t) \bullet \left[x_2(t_0 - \mathrm{T}_D) - 0.5(g_O^L + g_O^U) \right] < 0,\ t \in [t_0 - \mathrm{T}_D, t_0] \end{cases} \tag{6}$$

the corresponding industrial sample series during the period $[t_0 - \mathrm{T}_D, t_0]$ is labelled as excellent control effect sample series.

Definition 2: If $g(t)$ reaches the optimal range from the other ranges during the period $[t_0, t_0 + T_D]$ without exceeding the acceptable range, which could be described as Eq. (7):

$$\begin{cases} g(t_0 - \mathrm{T}_D) \notin [g_O^L, g_O^U] \\ g(t) \in [g_A^L, g_A^U],\ t \in (t_0 - \mathrm{T}_S, t_0] \\ g(t_0) \in [g_O^L, g_O^U] \end{cases} \tag{7}$$

where g_A^L and g_A^U denote the lower and upper limits of the acceptable ranges of the evaluation grades ($[g_O^L, g_O^U] \in [g_A^L, g_A^U]$), respectively; T_S denotes the time that $g(t)$ reaches the limits of the acceptable range at the first time; and then the corresponding industrial samples during the period $[t_0 - \mathrm{T}_D, t_0]$ are labelled as fine control effect samples.

Definition 3: If $g(\mathrm{t})$ reaches the acceptable range from the unacceptable range during the period $[t_0, t_0 + T_D]$ without exceeding the acceptable range, as shown in Eq. (8):

$$\begin{cases} g(t_0 - \mathrm{T}_D) \notin [g_A^L, g_A^U] \\ |g'(t)| < (g_A^L + g_A^U)/\mathrm{T}_D,\ t \in [t_0 - \mathrm{T}_D, t_0] \\ g(t_0) \in [g_A^L, g_A^U] \cap \overline{[g_O^L, g_O^U]} \end{cases} \tag{8}$$

And then the corresponding industrial samples during the period $[t_0 - \mathrm{T}_D, t_0]$ are labelled as good control effect samples.

Typical movements of the evaluation grade of the control effects are illustrated as Fig. 3.

2) Controllable sample classification and exclusion

Practically, the reaction environments are complex and the individuals' operations are diverse, thereby resulting in the variety of the evaluation grade movements. By using the proposed definitions, only a part of samples could be identified, while a considerable portion of samples are left with no labels. Meanwhile, due to the irrational operation, unreasonable samples should be rejected during this process. Therefore, an improved PU learning is proposed to classify samples and exclude bad samples at the same time for several classes.

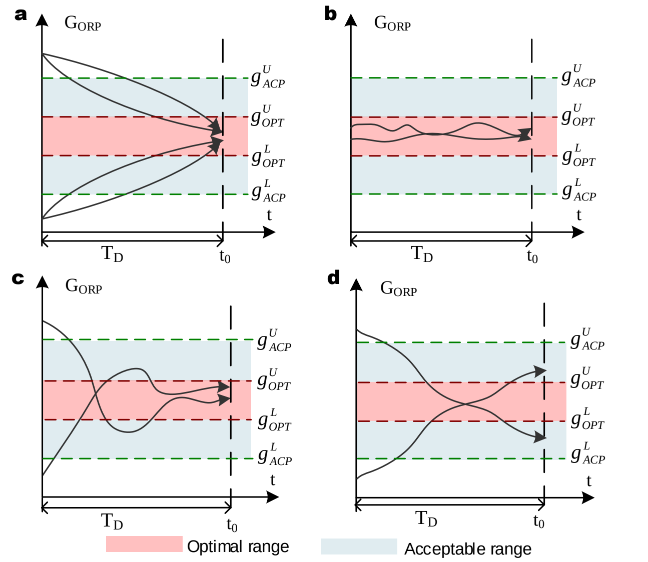

Fig. 3. Typical movements of evaluation grads for the excellent control effect (a) and (b), the fine control effect (c), and the good control effect (c).

kNN-PU algorithm could deal with two-class problems, but cannot be used directly to classify the industrial data into several controllable domains. The control effect classification is thus transferred to a classification method including a series of two-class problems. Firstly, parts of the excellent control effect samples are labelled with the above definitions, and kNN-PU algorithm is used to identify all the excellent control effect samples from the data set. These samples construct the excellent controllable domain. In the left samples, there are some bad data whose input characters are similar to excellent samples but have a different control effect. These data should be rejected before another round of classification. Similarly, all the fine control effect samples could be selected and the fine controllable domain is determined. And then, the good controllable domain can be found out in the same way. And the last left samples construct the marginal controllable domain. The procedure of the controllable domain classification is illustrated as Fig. 4.

Due to the diversity of the individual experiences, the adjustment amounts of zinc powder are distinct from each other even though the process conditions are similar. The corresponding control effects would be thus different. To extract proper and reasonable control rules, the samples produced based on bad individual experiences should be rejected during the classification. The detailed pseudo code of the controllable domain determination based on PU learning is illustrated as Algorithm 2.

The process condition similarity (PCS) used in algorithm 2 is presented as Eq. (9).

$$Sim_{PC}(\mathbf{u}_i, \mathbf{y}_j) = \frac{\sum_{k=1}^{m} w_k^2 \times \mathbf{u}_{i,k} \times \mathbf{y}_{j,k} - \alpha}{\sqrt{\sum_{k=1}^{m} \left(w_k \mathbf{u}_{i,k}\right)^2 - \beta^{\mathbf{x}}} \sqrt{\sum_{k=1}^{m} \left(w_k \mathbf{y}_{i,k}\right)^2 - \beta^{\mathbf{y}}}} \tag{9}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

where,

$$\alpha = \sum_{a=0}^{\Gamma} w_{m1,a}^{2} \times \left(m_{1,i}^{\mathbf{x}} d^{-a} \times m_{1,j}^{\mathbf{x}} d^{-a} + m_{2,i}^{\mathbf{x}} d^{-a} \times m_{2,j}^{\mathbf{x}} d^{-a} \right) \quad (10)$$

$$\beta^{\mathbf{x}} = \sum_{a=0}^{\Gamma} \left[w_{m1,a}^{2} \left(m_{1,i}^{\mathbf{x}} d^{-a} \right)^2 + \left(m_{2,i}^{\mathbf{x}} d^{-a} \right)^2 \right] \quad (11)$$

$$\beta^{\mathbf{y}} = \sum_{a=0}^{\Gamma} \left[w_{m1,a}^{2} \left(m_{1,i}^{\mathbf{y}} d^{-a} \right)^2 + \left(m_{2,i}^{\mathbf{y}} d^{-a} \right)^2 \right] \quad (12)$$

where, $\mathbf{u}_i$ denotes i th sample of the excellent, or fine or good control effect sample set ($\mathbf{s}_i^1$, $\mathbf{s}_i^2$, $\mathbf{s}_i^3$); $\mathbf{y}_j$ is j th sample of the first, or second or third left sample set ($\mathbf{s}_i^{'}$, $\mathbf{s}_i^{''}$, $\mathbf{s}_i^{'''}$); $\mathbf{u}_{i,k}$ and $\mathbf{y}_{j,k}$ are the k th variables of $\mathbf{u}_i$ and $\mathbf{y}_j$, respectively; w_k is the weight of the k th variable; $m_{1,i}^{\mathbf{x}}$ and $m_{2,i}^{\mathbf{x}}$ are the zinc amounts of the first and second reactors of $\mathbf{u}_i$, respectively; $m_{1,i}^{\mathbf{y}}$ and $m_{2,i}^{\mathbf{y}}$ are the zinc amounts of the first and second

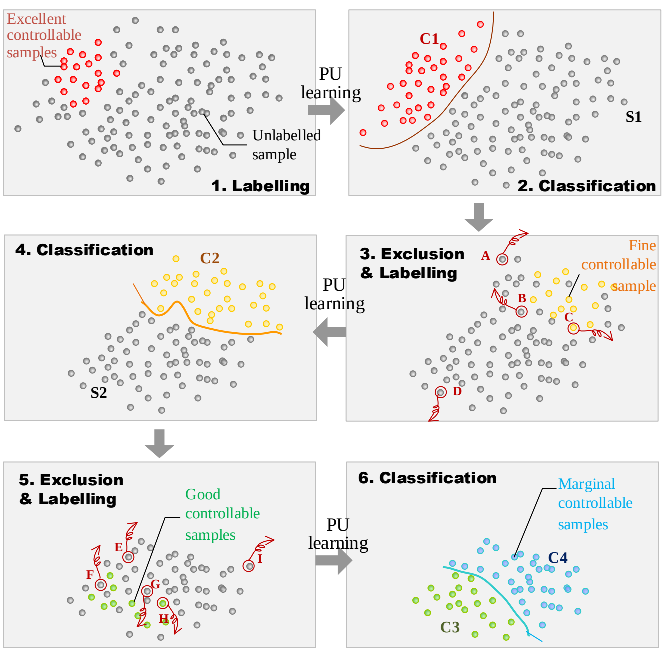

Fig. 4. The procedure of controllable domain classification based on kNN-PU learning: A~I are the rejected irrational-operated samples; S1 and S2 are the first and second left sample set, respectively; C1, C2, C3 and C4 are the excellent, fine, good and marginal controllable sample set, respectively.

reactors of $\mathbf{y}_j$, respectively; $w_{m1,a}$ ($w_{m1,a} \in w = \{w_1, \ldots, w_n\}$) and $w_{m2,a}$ (, $w_{m2,a} \in w = \{w_1, \ldots, w_n\}$) denote the weights of $m_1 d^{-a}$ and $m_2 d^{-a}$, respectively.

3) Controllable domain classifier

By applying algorithm 2, the industrial data are labelled and classified into different controllable domains. The classifiers produced in this process are not suitable for identifying new samples online during process control. Because the samples in the controllable domains differ from the samples on which these classifiers are trained. Accordingly, by multi-class SVM, a new controllable-domain classifier should be trained with the controllable samples.

Algorithm 2 PU learning based on kNN and SVM

Input $\mathbf{I} = \{\mathbf{I}_R, \mathbf{I}_M\} = \{\mathbf{s}_1, \ldots, \mathbf{s}_{m0}\}^{\mathrm{T}}$ the industrial sample set;
θ_{SIM}, threshold

Output The controllable domains

1. Label the "excellent" controllable samples from $\mathbf{I}$ according with definition 1
2. Apply kNN–PU learning algorithm to classify all the "excellent" controllable samples from $\mathbf{I}$, and denote the "excellent" controllable and the left sample sets as $\mathbf{C1} = \{\mathbf{s}_1^1, \ldots, \mathbf{s}_{m1}^1\}^{\mathrm{T}}$ and $\mathbf{S1} = \{\mathbf{s}_1^{L1}, \ldots, \mathbf{s}_{m2}^{L1}\}^{\mathrm{T}}$, ($m1 + m2 = m0$), respectively
3. **For** each "excellent" controllable samples $\mathbf{s}_i^1$
   For each left sample $\mathbf{s}_j^{L1}$
   Calculate PCS value $Sim_{PC}(\mathbf{s}_i^1, \mathbf{s}_j^{L1})$ $\mathbf{s}_j^{L1}$
   If $Sim_{PC}(\mathbf{s}_i^1, \mathbf{s}_j^{L1}) > \theta_{SIM}$, reject $\mathbf{s}_j^{L1}$ from $\mathbf{S1}$, **End if**
   End for
   End for
4. Obtain a new left sample set $\mathbf{S1}^* = \{\mathbf{s}_1^{LN1}, \ldots, \mathbf{s}_{m3}^{LN1}\}^{\mathrm{T}}$, ($m3 \le m2$)
5. Label the "fine" controllable samples from $\mathbf{S1}^*$, according with definition 2
6. Apply kNN–PU learning algorithm to classify all the "fine" controllable samples from $\mathbf{S1}^*$, and denote the "fine" controllable and the left sample sets as $\mathbf{C2} = \{\mathbf{s}_1^2, \ldots, \mathbf{s}_{m4}^2\}^{\mathrm{T}}$ and $\mathbf{S2} = \{\mathbf{s}_1^{L2}, \ldots, \mathbf{s}_{m5}^{L2}\}^{\mathrm{T}}$, ($m4 + m5 = m3$), respectively
7. **For** each "fine" controllable samples $\mathbf{s}_i^2$
   For each left sample $\mathbf{s}_j^{L2}$
   Calculate PCS value $Sim_{PC}(\mathbf{s}_i^2, \mathbf{s}_1^{L2})$
   If $Sim_{PC}(\mathbf{s}_i^2, \mathbf{s}_1^{L2}) > \theta_{SIM}$, reject $\mathbf{s}_j^{L2}$ from $\mathbf{S2}$, **End if**
   End for
   End for
8. Obtain a new left sample set $\mathbf{S2}^* = \{\mathbf{s}_1^{LN2}, \ldots, \mathbf{s}_{m6}^{LN2}\}^{\mathrm{T}}$, ($m6 \le m5$)
9. Label the "good" controllable samples from $\mathbf{S2}^*$, according with definition 3
10. Apply kNN–PU learning algorithm to classify all the "good" controllable samples from $\mathbf{S2}^*$, and denote the "good" controllable and the left sample sets as $\mathbf{C3} = \{\mathbf{s}_1^3, \ldots, \mathbf{s}_{m7}^3\}^{\mathrm{T}}$ and $\mathbf{S3} = \{\mathbf{s}_1^{L3}, \ldots, \mathbf{s}_{m8}^{L3}\}^{\mathrm{T}}$, ($m7 + m8 = m6$), respectively
11. **For** each "good" controllable samples $\mathbf{s}_i^3$
    For each left sample $\mathbf{s}_j^{L3}$
    Calculate PCS value $Sim_{PC}(\mathbf{s}_i^3, \mathbf{s}_j^{L3})$
    If $Sim_{PC}(\mathbf{s}_i^3, \mathbf{s}_j^{L3}) > \theta_{SIM}$, reject $\mathbf{s}_j^{L3}$ from $\mathbf{S3}$, **End if**
    End for
    End for
12. Obtain a new third left industrial set, and treat this set as the "marginal" controllable sample set $\mathbf{C4} = \{\mathbf{s}_1^4, \ldots, \mathbf{s}_{m9}^4\}^{\mathrm{T}}$, ($m9 \le m8$)

C. Rule extraction from copper removal fuzzified SVM classifier

Usually, the control rules are different for different controllable domains. In this case, reasonable rules should be extracted for each domain. The fuzzy rule extraction strategy is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

divided in four steps: 1) variable increments calculation, 2) fuzzy set predefinition, 3) sample labelling and 4) fuzzy rule extraction based on classification.

1) Variable increments calculation

In the practical process, the major amounts of zinc powder for copper removal are determined with the inlet flow rate and copper concentration, while the adjusted amounts are set according with increments of the online variables. So the variable increments should be calculated for rule extraction, as $\Delta x_i = x_i - x_i d^{-1}$, where x_i, ($i=1,...,4$) denotes the on-line industrial variables. In this step, to reduce computing complexity during the rule extraction, the historical values of industrial variables are removed from samples, while the variable increments are added. In this way, the i th industrial sample in the j th controllable domain ($j=1,...4$), $\mathrm{s}_i^j = \left\{x_{i,1}^j,...,x_{i,1}^j d^{-\Gamma},...,x_{i,4}^j,...,x_{i,4}^j d^{-\Gamma},x_{i,5}^j,x_{i,5}^j d^{-1},x_{i,6}^j,x_{i,6}^j d^{-1}\right\}$, is transferred to the sample with the current values of industrial variables and the increments of on-line variables, $\mathbf{s}_j^c = \left\{x_{1,j},...,x_{6,j},\Delta x_{1,j},...,\Delta x_{4,j}\right\}$. When $\mathbf{s}_j^c$ belongs to the excellent controllable domain, $c=E$, $j=1,...,m_E$; When $\mathbf{s}_j^c$ belongs to the fine controllable domain, $c=F$, $j=1,...,m_F$; When $\mathbf{s}_j^c$ belongs to the good controllable domain $c=G$, $j=1,...,m_G$; When $\mathbf{s}_j^c$ belongs to the marginal controllable domain $c=M$, $j=1,...,m_M$.

2) Fuzzy set predefinition

Mostly, fuzzy sets should be predefined for the rule' antecedents in most works related to fuzzy rule extraction methods. In this approach, the gaussian functions are used to fuzzify industrial variables, as $\mu(x)=\exp(-\frac{(x-o)}{2\sigma^2})$, where o locates the distance from the origin and σ indicates the width of the function curve.

We define 4 fuzzy sets for real time industrial variables and the measured copper concentrations, including very small (VS), little small (LS), little large (LL) and very large (VL); and 5 fuzzy sets for the increments of the industrial variables, including obviously decrease (OD), slightly decrease (SD), roughly stable (RS), slightly increase (SI) and obviously increase (OI). The membership degrees of x_i ($i=1,...,6$) in the j th sample of for VS, LS, LL, and VL are denoted as $\mu_{VS,i}(j)$, $\mu_{LS,i}(j)$, $\mu_{\mathrm{LL},i}(j)$ and $\mu_{V\mathrm{L},i}(j)$, respectively; and the membership degrees of Δx_i ($i=1,...,4$) in the j th sample of for OD, SD, RS, SI and OI are denoted as $\mu_{OD,i}(j)$, $\mu_{SD,i}(j)$, $\mu_{RS,i}(j)$, $\mu_{SI,i}(j)$ and $\mu_{OI,i}(j)$, respectively. To obtain these fuzzy sets, the parameters in these membership functions are optimized based on fuzzy Shannon entropy, as shown in Eq. (13).

$$\begin{aligned}&\max\ -\sum_{j=1}^{N_i}\left\{\mu_{\mathrm{FS},i}(j)\log\mu_{\mathrm{FS},i}(j)+\left(1-\mu_{\mathrm{FS},i}(j)\right)\log\left(1-\mu_{\mathrm{FS},i}(j)\right)\right\}\\ &s.t.\ \ \mu_{\mathrm{FS},i}(j)\in[0,1],\forall xi,j\\ &\qquad o_{FS,i}\in[o_{FS,i}^{\mathrm{L}},o_{FS,i}^{U}]\\ &\qquad \sigma_{FS,i}\in[\sigma_{FS,i}^{\mathrm{L}},\sigma_{FS,i}^{U}]\end{aligned} \tag{13}$$

where, FS denotes the name of fuzzy set which could be VS, LS, LL, VL, OD, SD, RS, SI and OI; $o_{FS,i}$ and $\sigma_{FS,i}$ are the parameters used in the membership function $\mu_{\mathrm{FS},i}(j)$; $o_{FS,i}^{\mathrm{L}}$ and $o_{FS,i}^{U}$ denote the lower and upper limits of the parameter $o_{FS,i}$, respectively; $\sigma_{FS,i}^{\mathrm{L}}$ and $\sigma_{FS,i}^{U}$ denote the lower and upper limits of the parameter $\sigma_{FS,i}$, respectively. Practically, the limits of $\sigma_{FS,i}$ are set according with the distribution of the training industrial samples, as shown in Eq. (14).

$$\begin{cases}o_{VS}^i\in\left[\min\left\{x_i\right\},\mathrm{Q}_{\alpha1}(x_i)\right]\\ o_{LS}^i\in\left[\mathrm{Q}_{\alpha1}(x_i),\mathrm{Q}_{\alpha2}(x_i)\right]\\ o_{LL}^i\in\left[\mathrm{Q}_{\alpha2}(x_i),\mathrm{Q}_{\alpha3}(x_i)\right]\\ o_{VL}^i\in\left[\mathrm{Q}_{\alpha3}(x_i),\max\left\{x_i\right\}\right]\end{cases} \tag{14}$$

where o_{VS}^i, o_{LS}^i, o_{LL}^i and o_{VL}^i denote the centre of the VS, LS, LL, and VL membership function of x_i, respectively; $\mathrm{Q}_{\alpha i}$, ($i=1,2,3$) is the quartile of the distribution of x_i. Usually, $\mathrm{Q}_{\alpha1}$, $\mathrm{Q}_{\alpha2}$ and $\mathrm{Q}_{\alpha3}$ are the first, second and third quartiles, respectively. The parameters constrains of the membership function for Δx_i are determined by Eq. (15).

$$\begin{cases}o_{OD}^i\in\left[\min(x_i)-\mathrm{Q}_{3/4}(x_i)/\Delta T_D,\mathrm{Q}_{1/4}(x_i)-\mathrm{Q}_{3/4}(x_i)/\Delta T_D\right]\\ o_{SD}^i\in\left[\mathrm{Q}_{1/4}(x_i)-\mathrm{Q}_{1/2}(x_i)/\Delta T_D,\mathrm{Q}_{1/4}(x_i)-\mathrm{Q}_{3/4}(x_i)/\Delta T_D\right]\\ o_{RS}^i\in\left[-\mathrm{Q}_{1/4}(x_i)/\Delta T_D,\mathrm{Q}_{1/4}(x_i)/\Delta T_D\right]\\ o_{SI}^i\in\left[\mathrm{Q}_{1/2}(x_i)-\mathrm{Q}_{1/4}(x_i)/\Delta T_D,\mathrm{Q}_{3/4}(x_i)-\mathrm{Q}_{1/4}(x_i)/\Delta T_D\right]\\ o_{OI}^i\in\left[\mathrm{Q}_{3/4}(x_i)-\mathrm{Q}_{1/4}(x_i)/\Delta T_D,\max(x_i)-\mathrm{Q}_{1/4}(x_i)/\Delta T_D\right]\end{cases} \tag{15}$$

And the limits of $\sigma_{FS,i}$ are determined by the limited range of $o_{FS,i}$, as shown in Eq. (16).

$$\sigma_{FS}^i\in\left[0.05(o_{FS,i}^U-o_{FS,i}^{\mathrm{L}}),0.3(o_{FS,i}^U-o_{FS,i}^{\mathrm{L}})\right] \tag{16}$$

3) Sample labelling

Before the classification for rule extraction performed, the industrial samples, $\mathbf{s}_j^c$, are labelled according with the memberships of zinc amount increments in the next control step ($\Delta x_3 p^1$ and $\Delta x_4 p^1$, where p^1 is the forward shift operator, $\Delta x_i p^1 = \Delta x_i(t+1)-\Delta x_i(t)$).

Firstly, the fuzzy sets which $\Delta x_3 p^1$ and $\Delta x_4 p^1$ belong to are determined by calculating their maximum membership degree, as Eq. (17).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

$$C_{\Delta x_i,j}=\begin{cases} OD, & f(\mu)=\mu_{OD,i}(j) \\ SD, & f(\mu)=\mu_{ID,i}(j) \\ RS, & f(\mu)=\mu_{RS,i}(j) \\ SI, & f(\mu)=\mu_{SI,i}(j) \\ OI, & f(\mu)=\mu_{OI,i}(j) \end{cases} \tag{17}$$

where

$f(\mu)=\max\left\{\mu_{OD,i}(j),\mu_{SD,i}(j),\mu_{\mathrm{RS},i}(j),\mu_{\mathrm{SI},i}(j),\mu_{OI,i}(j)\right\}$.

And then the industrial samples are labelled by evaluating the total variant amounts of zinc powder. The larger the total amount of zinc powder is decreased per unit time, the more negative the sample label is; whereas, the larger the total amount of zinc powder is increased per unit time, the more positive the sample label is. Notably, some operations, such as reducing one of zinc amounts but increasing the other, are unreasonable in the practical process control. Then the corresponding unreasonable combinations of Δx_3 and Δx_4 (e.g. OI & OD, SD & SI, etc.) are removed. After removing the repeating combinations, the sample labels are simplified to 11 kinds, as shown in Fig. 5.

ΔZn1 / ΔZn2	LD	SD	O	SI	LI
LD	-5	-4	-3	—	—
SD	-4	-2	-1	—	—
O	-3	-1	0	1	3
SI	—	—	1	2	4
LI	—	—	3	4	5

Fig. 5. The classes determined by the combination of adjustment amounts of zinc powder added in the first and second reactors: - denotes the unreasonable combination of Δx_3 and Δx_4 ..

For example, if $\Delta x_3 p^1$ and $\Delta x_4 p^1$ both belong to OD, the corresponding industrial sample $\mathbf{s}_j$ is labelled as -5. In fact, it is nearly impossible to classify 11 sets for each controllable domain. The industrial samples from a special controllable domain could only be divided into a few groups. For instance, in the excellent controllable domain, the amounts of zinc powder are mostly adjusted smoothly and timely. $\Delta x_3 p^1$ and $\Delta x_4 p^1$ in this controllable domain hardly belong to the fuzzy sets with extreme changes, such as OD and OI. Consequently, there are few samples with more negative or positive labels (e.g. -5, 5, -4, 4, etc.) in the excellent controllable domain.

4) Fuzzy rule extraction

By using the rules in Fig. 5, the industrial samples in four controllable domains can be labelled with the amounts of zinc powder. And then multi-class SVM is applied in each controllable domain. The support vectors obtained by these classifiers determine the fuzzy rules of each domain. The procedure of this step is described as follows:

1) Label the sample $\mathbf{s}_j^c=\left\{x_{1,j},...,x_{6,j},\Delta x_{1,j},...,\Delta x_{4,j}\right\}$ with $y=i$, $i=-5,...,5$, by using the definition afforded in the last section;
2) Train multi-class SVM, f_{SVM}^c, for each data set of controllable domain $\mathbf{S^c}=\left\{\mathbf{s}_1^c,....,\mathbf{s}_{m_c}^c\right\}$;
3) Select support vectors, $\mathbf{SV}_i^c$, ($i=1,...,m_{SV}^c$), from controllable domains, and project $\mathbf{SV}_i^c$ in the coordinate axes to obtain $\mathbf{s}_{SV,i}^c$;
4) Divide each real time variable attribute of the input space into 4 fuzzy sets, and each variable increment attribute of the input space into 5 fuzzy sets by using the memberships defined by Eq.(13)-(17);
5) Select the fuzzy set, $C_{\mathrm{i,jmax}}^c$, which provides maximum membership degree, $\max\limits_j\left\{\mu_{\mathrm{i,j}}^{c,SV}\right\}$, for each attribute of $\mathbf{s}_{SV,i}^c$
6) Generate fuzzy rules by these support vectors as follows. Denote $C_{\mathrm{x}_i,\mathrm{jmax}}^c \in$ {VS, LS, LL,VL} and $C_{\Delta\mathrm{x}_i,\mathrm{jmax}}^c\in\left\{OD,SD,RS,SI,OI\right\}$ as the fuzzy sets with highest membership degree of x_i ($i=1,..,6$) and Δx_i ($i=1,..,4$) of $\mathbf{s}_{SV,i}^c$, respectively. Let $C_{\mathbf{sv},j}$ be the class of the corresponding support vector $\mathbf{SV}_i^c$ classified by the multi-class SVM. And then the rule R generated by $\mathbf{SV}_i^c$ (or the projected sample $\mathbf{s}_{SV,i}^c$) could be represented as:

   If x_1 is $C_{\mathrm{x}_1,\mathrm{jmax}}^c$,…, x_6 is $C_{\mathrm{x}_6,\mathrm{jmax}}^c$, Δx_1 is $C_{\Delta\mathrm{x}_1,\mathrm{jmax}}^c$,…, Δx_4 is $C_{\Delta\mathrm{x}_4,\mathrm{jmax}}^c$; Then $\mathbf{s}_{SV,i}^c$ belongs to $C_{\mathbf{sv},j}$.

   The formation rule could be translated into another form where the label of $\mathbf{sv}_j^c$ is changed to the corresponding fuzzy classes of $\Delta x_3 p^1$ and $\Delta x_4 p^1$.

   If x_1 is $C_{\mathrm{x}_1,\mathrm{jmax}}^c$,…, x_6 is $C_{\mathrm{x}_6,\mathrm{jmax}}^c$, Δx_1 is $C_{\Delta\mathrm{x}_1,\mathrm{jmax}}^c$,…, Δx_4 is $C_{\Delta\mathrm{x}_4,\mathrm{jmax}}^c$; Then $\Delta x_3 p^1$ is $C_{\Delta\mathrm{x}_3 p^1,j}$, and $\Delta x_4 p^1$ is $C_{\Delta\mathrm{x}_4 p^1,j}$.

V. Experiment Results

To verify the proposed strategy, we sampled data per 5 minutes from an actual copper removal process. After rejecting abnormal samples, 23106 samples are selected during two months. And then the controllable domain classification is applied. Firstly, 4201 and 1000 excellent controllable samples are selected by the proposed definition and rule extraction, respectively. After removing redundant and error-operated samples, 4311 fine controllable samples are defined and 2141 samples are selected. Similarly, good controllable sample set is generated comprising 2481 defined samples and 4362 picked ones. The last left 4012 samples constitute the marginal controllable sample set after sample rejection. Accordingly, the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

samples are divided into four sets, and the controllable domain classifier is built by applying multi-class SVM. The numbers and variable fluctuation ranges of these domains are provided in Table III, and the variable distributions are obtained, as shown from Fig. 6 and Fig. 7.

higher control effect. From Fig. 6 and Fig. 7, it is known that the distribution center of most variables are mostly consistent in all the controllable domains, while the distributions of ORP are distinguished with each other in these controllable domains. The ORP distribution centers in the excellent, fine, good and marginal controllable domains are located towards 60 mV, 78 mV, 78 mV and 89 mV, respectively. So it is easy to obtain an expected control effect if ORP does not depart from the acceptable range. From Fig. 7 (d), it is noted that the distribution center of the outlet copper concentration (0.15 g/L) is far less than the optimal concentration (0.3g/L) in the

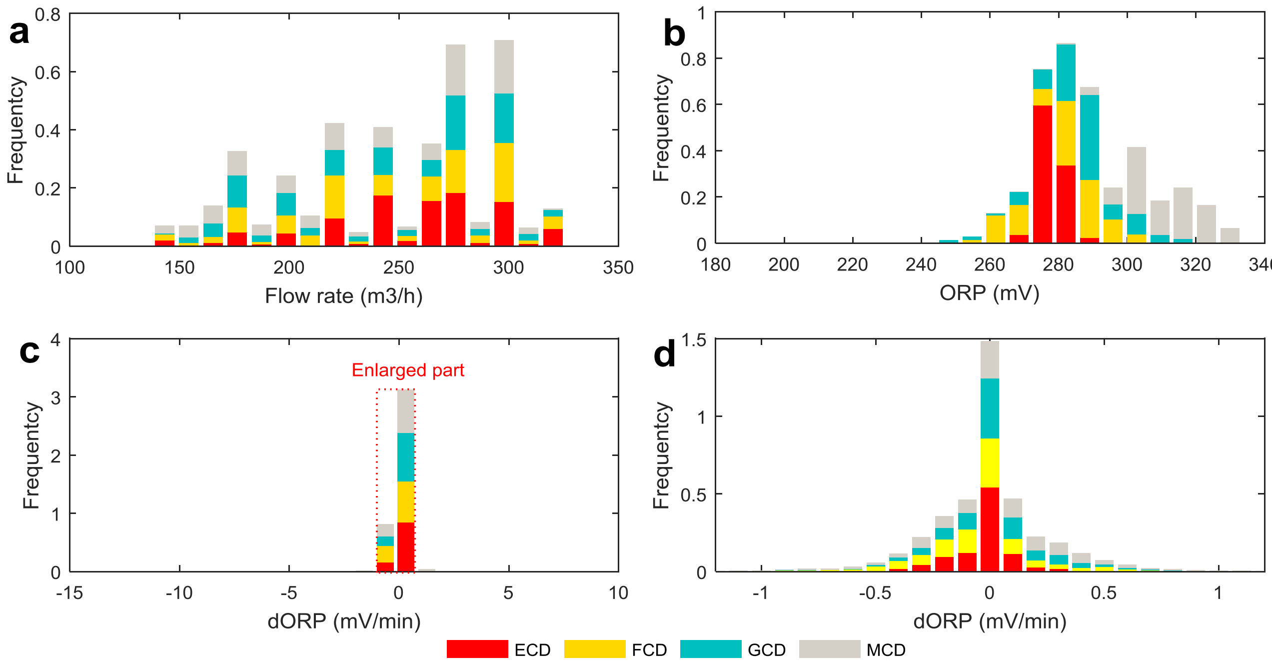

Fig. 6. The distributions of inlet flow rate (a), ORP (b), ORP increments (c) and enlarged part of ORP increments between range [-0.8 0.8] (d) for four controllable domains.

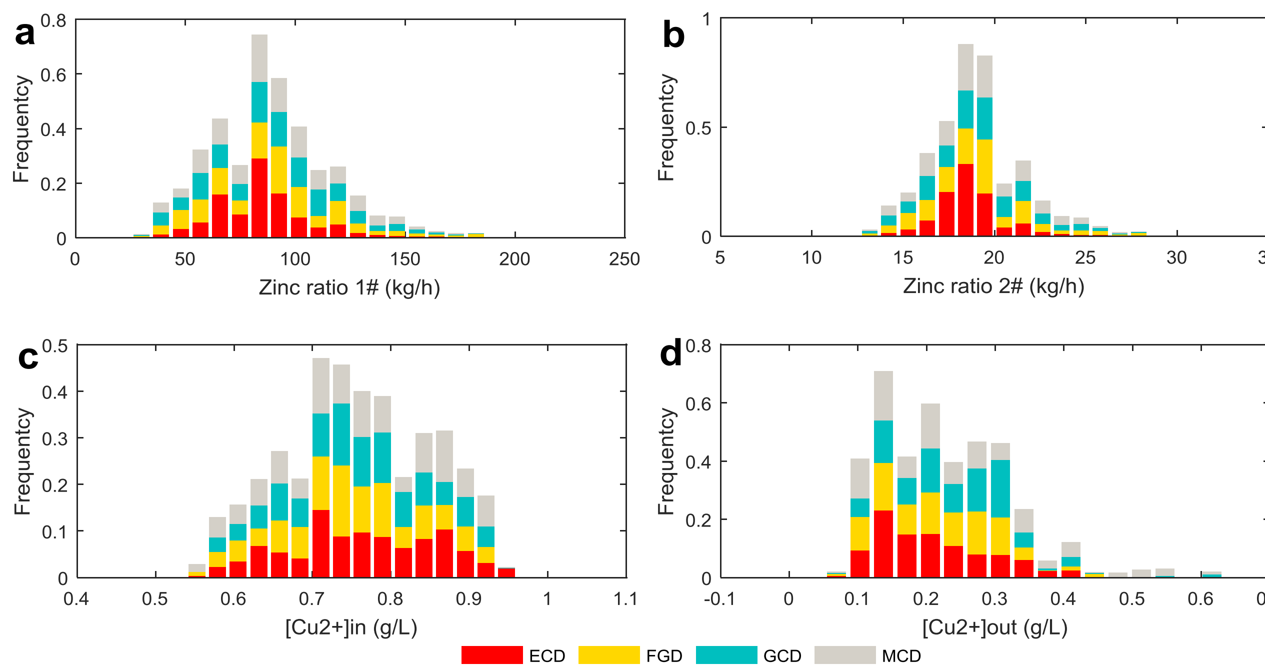

Fig. 7. The distributions of zinc amounts added in the first reactor (a) and the second reactor (b), the distributions of the inlet (c) and outlet (d) copper concentrations for four controllable domains.

From Table III, it is observed that the fluctuation ranges of the flow rate and inlet copper concentration are the same in all the controllable domains. But the fluctuation ranges of the other variables, generally, increase in the sequence of the excellent, fine, good and marginal controllable domains. The copper removal process condition becomes strict when requiring a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

TABLE III
THE NUMBERS AND BOUNDS OF CONTROLLABLE DOMAINS FOR THE COPPER REMOVAL PROCESS

Domains	Excellent	Fine	Good	Marginal
N_D [a]	5012	6452	6843	4012
x_1	[100, 310]	[100, 310]	[100, 310]	[100, 310]
x_2	[39, 86]	[39, 95]	[19, 110]	[19, 130]
dx_2	[-1.25, 2]	[-3.4, 2]	[-5, 4.8]	[-11.2, 2]
x_3	[31, 162]	[31, 181]	[31, 181]	[13, 29]
x_4	[13, 28]	[13, 27]	[13, 28]	[3.9, 4.2]
x_5	[0.55, 0.94]	[0.55, 0.94]	[0.55, 0.94]	[0.55, 0.94]
x_6	[0.1, 0.4]	[0.1, 0.6]	[0.1, 0.6]	[0.1, 0.6]

[a] N_D is the number of the samples in the controllable domains.

excellent controllable domain. This phenomenon shows the preference for adding excessive zinc amounts in human operations, the easiness to increase a lower outlet concentration and the difficulty to pull down a higher concentration.

After domain classification, a set of fuzzy rules are extracted for each domain. To verify the effectiveness of the proposed controllable-domain based fuzzy rule extraction strategy (CD-FRE), manual operation (MO) and two general rule extraction methods, including of conventional fuzzy rule extraction method based on SVM (CFRE) and conventional PU learning based fuzzy rule extraction method (PU-FRE) are introduced for comparison. In CFRE algorithm, the rules are extracted for the whole sample set without classification and rejection. In PU-FRE algorithm, the industrial samples are divided into four domains without sample rejection, and four sets of rules are extracted for the controllable domains. In this part, the extracted rules are analysed by two criteria: accuracy and coverage [20].

1) Accuracy

The accuracy of rule R associated with the class C_s, $C_s \in \{-5,-4,-3,-2,-1,0,1,2,3,4,5\}$, is defined as $A_{CS}^R = \sum_{i=1}^{m_{CS}} \mu_R(\mathbf{x_i}) \Big/ \sum_{j=1}^{m} \mu_R(\mathbf{x_j})$, where m_{CS} is the number of the industrial sample belongs to a special class C_s; m is the total number of the industrial samples; μ_R is the product of the membership degrees of the industrial sample to each fuzzy set that belongs to the antecedent of rule R.

And μ_R is calculated as $\mu_R(\mathbf{x_i}) = \prod_{j=1}^{n_{input}} \mu_{C_{x_{j,i}}}(x_{j,i})$, where n_{input} is the number of the input variables; $x_{j,i}$ is the j th variable of $\mathbf{x_i}$; $\mu_{C_{x_{j,i}}}$ is the membership degree of $x_{j,i}$ belongs to $C_{x_{j,i}}$ which is the fuzzy class appears in rule R.

2) Coverage

The fuzzy coverage measures the ability of rule R affecting the patterns, it is defined as $CO_{CS}^R = \sum_{j=1}^{m} \mu_R(\mathbf{x_j}) \Big/ m$.

Table IV presents the rule prediction results of different rule extraction methods. It is observed that the fuzzy rules extracted by the proposed method play well in the excellent and fine controllable domains, and the prediction precision and coverage decrease in other two domains. Because the reaction conditions in the last two domains become more complex and diverse, limiting the rule precision. The fuzzy rules extracted by PU-FRE play as well as those extracted by CD-FRE in the excellent domains. But both the accuracy and overage of these rules by applying PU-FRE are decreased in the other domains compared with CD-FRE. The unreasonable data cannot be removed by CD-FRE, which affects the rule effectiveness.

TABLE IV
RESULTS OF FUZZY RULE EXTRACTION FOR COPPER REMOVAL PROCESS

Rule extraction methods		Numbers of rules	Accuracy (%)	Coverage (%)
Excellent	PU-FRE	10	79	96
	CD-FRE	10	79	96
Fine	PU-FRE	12	73	90
	CD-FRE	13	77	92
Good	PU-FRE	14	71	84
	CD-FRE	13	73	89
Marginal	PU-FRE	9	70	74
	CD-FRE	8	71	79
CFRE		14	73	78

Notably, the variable distributions in the original sample set are similar to that in marginal controllable domain. Accordingly, the performances of the rules extracted by CFRE are near to those extracted by CD-FRE for marginal controllable domain.

To observe the behavior of the extracted fuzzy rules in copper removal process control, these rules extracted by three methods and MO method are performed to adjust the amount of zinc powder. For the experiments, more than 2800 representative samples selected from an actual copper removal process in China. Fig. 8 (a) and (b) show the values of the flow rate, ORP and the inlet copper concentrations. These data are sampled from four typical inlet conditions. During the first 10 hours, the flow rate changes several times but the inlet copper concentration is kept relative constant. The flow rate becomes stable while the concentration fluctuates in the second 10 hours. In the third 10 hours, both of the concentration and flow rate are in a steady state, while both of them are fluctuant in the left hours. The fluctuation frequencies also prove the condition movements.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

We present the adjustment amounts of zinc powders set by these methods in Fig. 8 (c) and (d). It is noticed that the adjustment frequency of MO is below that of the others, and the adjustment amounts are casual and sometimes change sharply. The adjusted amounts set by the CFRE and PU-FRE change frequently. But by applying these two methods, the adjusted amounts for the first and second reactor are, sometimes, unreasonable and conflicting. For example, the adjusted amount is increased for the first reactor but decreased for the second reactor at the 24th hour. Because the conflicting operations are common in the original set, and cannot be rejected by CFRE and PU-FRE. The data produced by the unreasonable operations reduce the expected control efficiency of the extracted rules. The operations based on the proposed method are both timely and proper.

The total amounts of zinc powder are shown in Fig. 9 (a) and (b). Clearly, in Fig. 9 (a), the amounts added in the first reactor set by CFRE, PU-FRE and CD-FRE are changed more smoothly than MO. However, Fig. 9 (b) shows that the operations for the second reactor by applying the proposed strategy are more moderate and reasonable compared with the others. A gentle and timely operation might be more benefit for

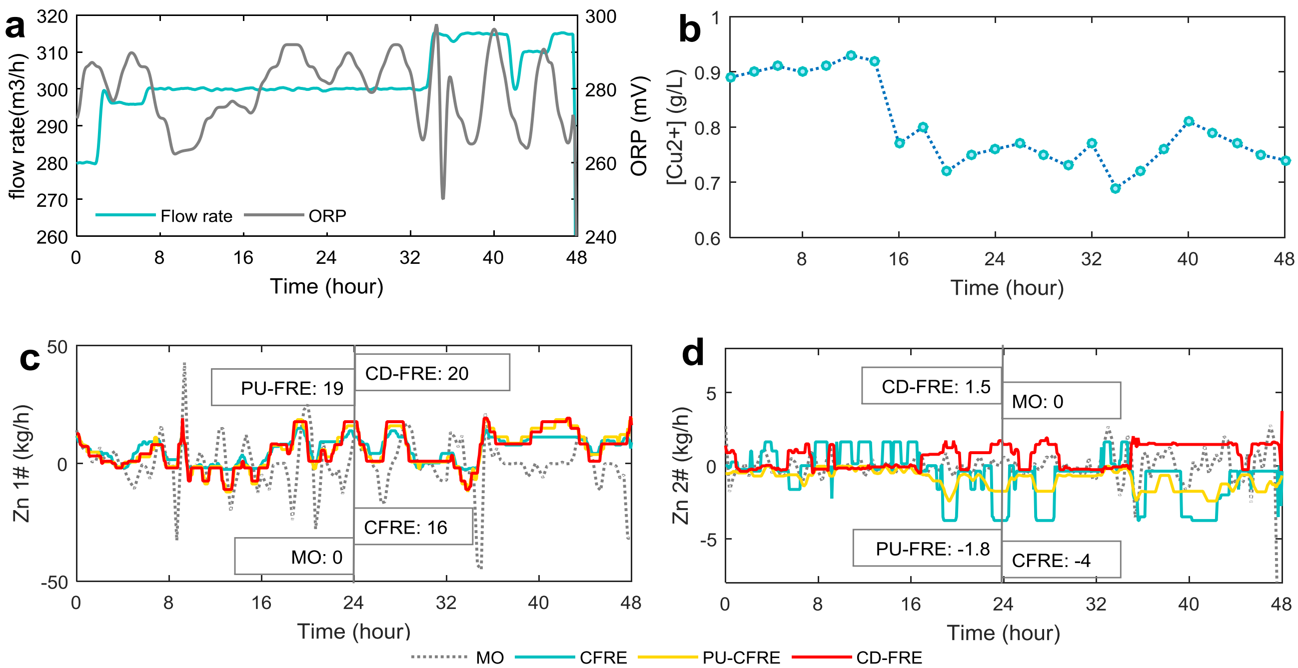

Fig. 8. The condition variables and the zinc amounts adjustments of copper removal process: (a) for the inlet flow rate and ORP, (b) for the inlet copper concentration, (c) and (d) for the zinc amounts adjustments for the first and second reactor, respectively.

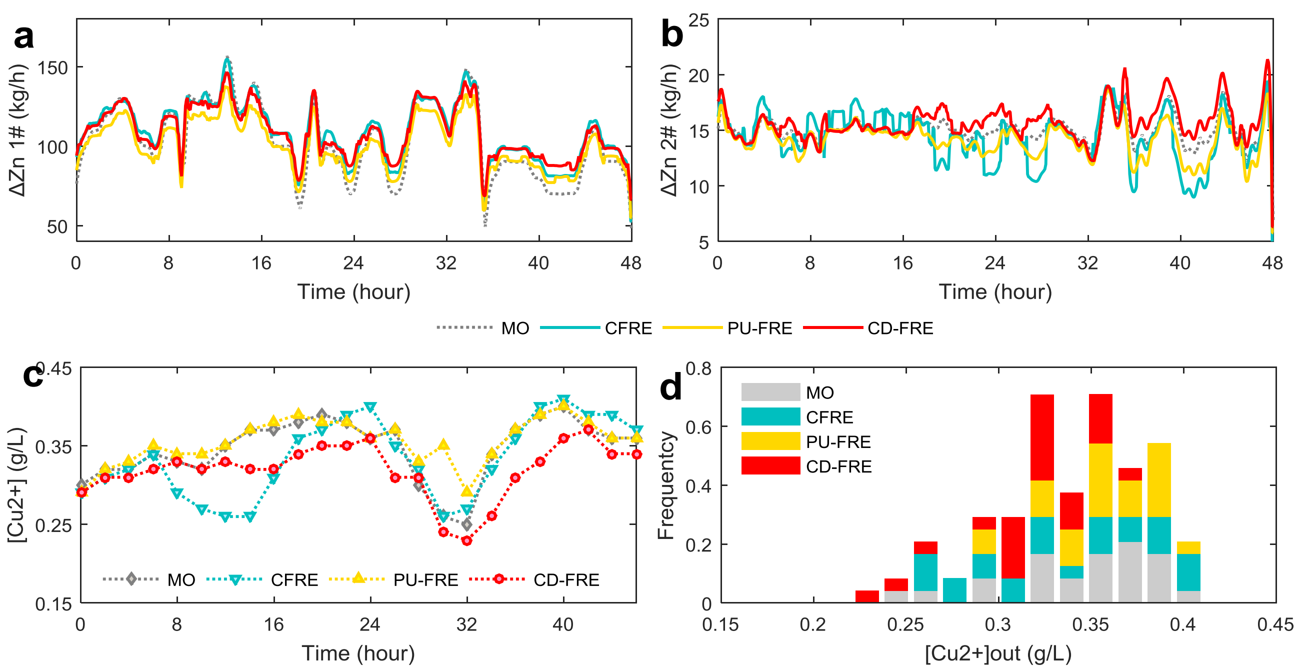

Fig. 9. The control results under different methods: (a) and (b) for the zinc amounts added in the first and second reactors, respectively; (c) and (d) for the curves of the outlet copper concentrations and the distributions of the concentrations, respectively.

process stabilization.

The control results of the three methods are confirmed in Fig. 9 (c) and (d), where the red lines are the upper and lower limits set by the production index. By applying CD-FRE, all of the outlet concentrations fall into the required ranges, while those obtained under the other methods occasionally exceeds the limitations, as shown in Fig. 9 (c). Additionally, the outlet concentrations are closer to the center of the required range (0.3 g/L) by using CD-FRE rather than others. The concentration distribution plot, Fig. 9 (d), is also proved the effectiveness of the proposed strategy. The distribution range of CD-FRE is the narrowest in three methods; and by applying CD-FRE the number of the concentrations around 0.3 g/L is the largest.

To accurately evaluate the control results, four indexes are introduced in this paper: oscillation range, qualification rate, mean value and median (Table V). From median and mean values, it is noticed that the concentrations obtain by MO are mostly higher than the range center. And not all of the concentrations are qualified. By applying CFRE, the unqualified concentrations are slightly increased, but more concentrations move to the range center. When performing PU-FRE, the fluctuation range of the outlet concentrations is narrowed, but most of the concentrations seem far beyond the optimal concentration. Compared with these methods, by using CD-FRE, the qualities of the outlet copper concentrations are greatly improved on all the aspects (including oscillation range, qualification rate, mean and median values).From the simulation results, it is suggested that the performance of CD-FRE is superior to manual operation, CFRE and PU-FRE. It could not only remove the copper impurity effectively, but also reserve an appropriate amount of copper ions as activator for the next removal stage.

TABLE V
EVALUATION INDEXES FOR THE CONTROL RESULTS

Methods	Oscillation range (g/L)	Q_r (g/L)	Mean value (g/L)	Median (g/L)
MO	[0.25, 0.41]	95%	0.35	0.36
CFRE	[0.26, 041]	91%	0.33	0.33
PU-FRE	[0.28, 0.405]	95%	0.36	0.38
CD-FRE	[0.23, 0.37]	100%	0.31	0.32

[a] Q_r is Qualification rate.

VI. CONCLUSION

This paper has developed a controllable-domain based fuzzy rule extraction strategy to draw reasonable adjustment laws from abundant industrial samples. Firstly, new definitions are proposed to label samples by their representative controlled effects. Afterwards, with these labeled samples, PU learning is introduced to classify all samples into four controllable domains. At last, fuzzy rules are extracted by using fuzzy-SVM for each controllable domain. The extracted rules are applied to adjust the amounts of zinc powder in the copper removal process. The encouraging control results reveal that the rules could effectively improve the production qualification rate and process stability more than the rules extracted by conventional method and practical manual operations. However, some problems are still existed, which needs us to study, such as how to take energy consumption in the consideration of the controllable domain classification and rule extraction. Overall, the simulation results show the proposed rule extraction strategy has promising applications for industrial-scale impurity removal process.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (61273185), the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (61621062), the innovation-driven plan in Central South University, the 111 project of China (B17048) and Australian Research Council (DP140102180, LP140100471, LE150100079).

REFERENCES

[1] B. Zhang, C. H. Yang, H. Q. Zhu, Y. G. Li, and W. H. Gui, “Kinetic modeling and parameter estimation for competing reactions in copper removal process from zinc sulfate solution,” *Ind. Eng. Chem. Res.*, vol.52, no. 48, pp.17074-17086, Nov. 2013.

[2] B. Zhang, C. H. Yang, Y. G. Li, X. L. Wang, H. Q. Zhu, and W. H. Gui, “Additive requirement ratio prediction using trend distribution features for hydrometallurgical purification processes”. *Control. Eng. Pract.*, vol. 46, pp.10-25, Jan. 2016.

[3] B. Zhang, C. H. Yang, H. Q. Zhu, Y. G. Li, and W. H. Gui, “Evaluation strategy for the control of the copper removal process based on oxidation-reduction potential.” *Chem. Eng. J.*, vol.284, no. 15, pp.294–304, Jan. 2016.

[4] Y. G. Li, W. H. Gui, K. L. Teo, H. Q. Zhu, and Q. Q. Chai, “Optimal control for zinc solution purification based on interacting CSTR models,” *J. Process. Contr.*, vol. 22, no. 10, pp. 1878-1889, Dec. 2012.

[5] B. Sun, W. H. Gui, Y. L. Wang, and C. H. Yang, “Intelligent optimal setting control of a cobalt removal process,” *J. Process. Contr.*, vol. 24, no. 5, pp. 586-599, May. 2014.

[6] B. Sun, W. H. Gui, Y. L. Wang, and C. H. Yang, “An integrated prediction model of cobalt ion concentration based on oxidation-reduction potential,” *Hydrometallurgy.*, vol. 140, pp. 102-110, Nov. 2013.

[7] B. Gholami, J. M. Bailey, W. M. Haddad, and A. R. Tannenbaum, “Clinical decision support and closed-loop control for cardiopulmonary management and intensive care unit sedation using expert systems,” *IEEE Trans. Contr. Syst. T.*, vol. 20, no. 5, pp. 1343-1350, Sep. 2012.

[8] R. R. Yager, “Perception-based granular probabilities in risk modeling and decision making,” *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 2, pp. 329-339, Apr. 2006.

[9] D. P. Pancho, J. M. Alonso, O., Quirin, A. Cordón, and L. Magdalena, “FINGRAMS: visual representations of fuzzy rule-based inference for expert analysis of comprehensibility,” *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 6, pp. 1133-1149, Dec. 2013.

[10] H. Han, and J. Qiao. “Nonlinear model-predictive control for industrial processes: An application to wastewater

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

treatment process." *IEEE Trans. Ind. Electron.*, vol. 61, no. 4, pp. 1970-1982, Apr. 2014.

[11] J. J. Castro-Schez, J. M. Murillo, R. Miguel, and X. Luo, "Knowledge acquisition based on learning of maximal structure fuzzy rules," *Knowl-Based. Syst.*, vol. 44, pp. 112-120, May. 2013.

[12] Y. Han, P. Shi and S. Chen, "Bipolar-valued rough fuzzy set and its applications to decision information system", *IEEE Trans. Fuzzy Syst.*, vol.23, no.6, pp.2358-2370, 2015.

[13] Q. Lu, P. Shi, H. K. Lam and Y. Zhao, "Interval Type-2 fuzzy model predictive control of nonlinear networked control systems", *IEEE Trans. Fuzzy Syst.*, vol.23, no.6, pp.2317-2328, 2015.

[14] Y. Han, P. Shi and S. Chen, "Bipolar-valued rough fuzzy set and its applications to decision information system", *IEEE Trans. Fuzzy Syst.*, vol.23, no.6, pp.2358-2370, 2015.

[15] Y. C. Chen, N. R. Pal, and I. Chung, "An integrated mechanism for feature selection and fuzzy rule extraction for classification," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 4, pp. 683-698, Aug. 2012.

[16] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Selecting fuzzy if-then rules for classification problems using genetic algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 3, no. 3, pp. 260–270, Aug. 1995.

[17] R. Alcala, M. J. Gacto, and F. Herrera, "A fast and scalable multiobjective genetic fuzzy system for linguistic fuzzy modeling in high-dimensional regression problems," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 4, pp. 666–681, Aug. 2011.

[18] R. Alcalá, Y. Nojima, F. Herrera, and H. Ishibuchi. "Multiobjective genetic fuzzy rule selection of single granularity-based fuzzy classification rules and its interaction with the lateral tuning of membership functions." *Soft. Comput.*, vol. 15, no.12, pp. 2303-2318, Dec. 2011.

[19] M. Delgado, A. F. Gomez-Skarmeta, and F. Martin, "A fuzzy clustering based rapid-prototyping for fuzzy modeling," *IEEE Trans. Fuzzy Syst.*, vol. 5, no. 2, pp. 223–233, May. 1997.

[20] E. Lughofer and S. Kindermann, "SparseFIS: Data-driven learning of fuzzy systems with sparsity constraints," *IEEE Trans. Fuzzy Syst.*, vol. 18, no. 2, pp. 396–411, Apr. 2010.

[21] S. Kulluk, L. Özbakır, and A. Baykasoğlu. "Fuzzy DIFACONN-miner: A novel approach for fuzzy rule extraction from neural networks." *Expert. Syst. with Appl.*, Vol. 40, Issue 3, 15 February 2013, Pages 938-946. vol. 40, no. 3, pp. 938-946, Feb. 2013.

[22] C. J. Carmona, P. González, M. J. del Jesus, and M. Navío-Acosta, L. Jiménez-Trevino. "Evolutionary fuzzy rule extraction for subgroup discovery in a psychiatric emergency department." *Soft .Comput.*, vol. 15, no.12, pp. 2435-2448, Dec. 2011.

[23] R. Alcalá, J. Alcalá-Fdez, and F. Herrera, "A proposal for the genetic lateral tuning of linguistic fuzzy systems and its interaction with rule selection," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 4, pp. 616-635, Aug. 2007.

[24] C. Gao, Q. Ge, and L. Jian. "Rule extraction from fuzzy-based blast furnace SVM multiclassifier for decision-making." I*EEE Trans. Fuzzy Syst.*, vol. 22, no. 3, pp. 586-596, June 2014.

[25] R. P. Prado, S. García-Galán, J. E. M. Expósito, and A. J. Yuste. "Knowledge acquisition in fuzzy-rule-based systems with particle-swarm optimization." *IEEE Trans. Fuzzy Syst.*, vol. 18, no. 6, pp. 1083-1097 , Dec. 2010.

[26] C. Hua, and S. Ding. "Decentralized networked control system design using T-S fuzzy approach." *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 1, pp. 9-21, Feb. 2012.

[27] Y. Rahulamathavan, R. C. W. Phan, S. Veluru, K. Cumanan, and M. Rajarajan, "Privacy-preserving multi-class support vector machine for outsourcing the data classification in cloud," *IEEE Trans. Depend. Secure.*, 11(5), 467-479. Sep.-Oct. 2014.

[28] M. Elkano, M. Galar, J. A. Sanz, A. Fernandez, E. Barrenechea, F. Herrera, and H. Bustince, "Enhancing Multiclass Classification in FARC-HD Fuzzy Classifier: On the Synergy Between-Dimensional Overlap Functions and Decomposition Strategies," *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 5, pp. 1562-1580, Oct. 2015.

[29] K. Polat, and S. Güneş, "A novel hybrid intelligent method based on C4. 5 decision tree classifier and one-against-all approach for multi-class classification problems," *Expert. Syst. Appl.*, vol. 36, no. 2, pp. 1587-1592, Mar. 2009.

[30] K. Nigam, A.K. McCallum, S. Thrun, and T.M. Mitchell, "Text classification from labeled and unlabeled documents using EM," *Mach. Learn.*, vol. 39, no. 2, pp. 103-134, May. 2011

[31] B. Zhang, and W. Zuo, "Learning from positive and unlabeled examples: A survey," *In Proc. ISIP*, Moscow, Russia, 2008, pp. 650-654.

[32] B. Liu, W. Lee, P. Yu, and X. Li, "Partially supervised classification of text documents," *in Proc. ICML*, Sydney, Australia, 2002, pp. 387-394.

[33] H. Yu, J. Han, and K. Chang, "PEBL: Positive example based learning for web page classification using SVM," *in Proc. ACM SIGKDD*, Alberta, Canada, 2002, pp: 239-248.

[34] X. Li, and B. Liu, "Learning to classify texts using positive and unlabeled data," *in Proc. IJCAI*, Acapulco, Mexico, 2003, pp: 587-594.

[35] J. He, Y. Zhang, X. Li, and P. Shi. "Learning naive Bayes classifiers from positive and unlabelled examples with uncertainty." *Int. J. Syst. Sci.* vol. 43, no. 10, pp. 1805-1825, Oct. 2012.

[36] M. Li, S. Pan, Y Zhang, and X. Cai. "Classifying networked text data with positive and unlabeled examples." *Pattern. Recogn. Lett.*, vol. 77, pp. 77: 1-7, Jul. 2016.

[37] W. Li, Q. Guo, and C. Elkan. "A positive and unlabeled learning algorithm for one-class classification of remote-sensing data." *IEEE Trans Geosci. Remote.*, vol. 49, no. 2, pp: 717-725, Aug. 2011.

[38] M. González, C. Bergmeir, I. Triguero, Y. Rodríguez, and J. M. Benítez, "On the stopping criteria for k-Nearest Neighbor in positive unlabeled time series classification problems," *Inform. Sciences.*, vol. 328, pp. 42-59, Jan. 2016.

[39] D. Ienco, and R. Pensa. "Positive and unlabeled learning in categorical data". *Neurocomputing.*, vol. 196, pp: 113-124, Jul. 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2017.2751000, IEEE Transactions on Fuzzy Systems

[40] X. Li, B. Liu, and S. Ng, “Negative training data can be harmful to text classification,” *in Proc. EMNLP*, Massachusetts, USA, 2010, pp. 218-228.

[41] L. Shi, X. Ma, L. Xi, Q. Duan, and J. Zhao, “Rough set and ensemble learning based semi-supervised algorithm for text classification,” *Expert Syst. Applicat.*, vol. 38, no. 5, pp. 6300-6306, 2011

[42] B. Nahla, and A. Bradley, "Rule extraction from support vector machines: a review." *Neurocomputing*. vol. 74, no. 1-3, pp. 178-190, Dec. 2010.

[43] T. Jafar, M. Someren, and H. Afsarmanesh, "Semi-supervised self-training for decision tree classifiers." *Int. J. Mach. Learn. & Cyber.*, vol. 8, pp. 355-370, 2017.

[44] B. Zhang, and W. Zuo, "Reliable negative extracting based on kNN for learning from positive and unlabeled examples". *JCP*., vol. 4, no. 1, pp. 94-101, 2009.

[45] M. Nguyen, X. Li, and S. Ng, “Positive unlabeled leaning for time series classification”. *in Proc IJCAI*, Barcelona, Spain, 2011, pp. 1421-1426.

[46] M. González, C. Bergmeir, I. Triguero, Y. Rodríguez, and J. Benítez, "On the stopping criteria for k-nearest neighbor in positive unlabeled time series classification problems." *Inform. Sciences.*, vol. 328, pp. 42-59, Jan. 2016.

[47] Y. Shao, W. Chen, L. Liu, and N. Deng, "Laplacian unit-hyperplane learning from positive and unlabeled examples." *Inform. Sciences.*, vol. 314, pp. 152-168, Sep. 2015.

[48] A. Daneshpazhouh, and A. Sami, "Semi-supervised outlier detection with only positive and unlabeled data based on fuzzy clustering." *Int. J. Artif. Intell. T.*, vol. 24, no. 3. pp, 344-348, Jun. 2015.

[49] M. Claesen, F. Smet, J. Suykens, and B. Moo, "A robust ensemble approach to learn from positive and unlabeled data using SVM base models." *Neurocomputing.*, vol. 160. pp.73-84, Jul. 2014.

[50] S. Jain, M. White, and P Radivojac, "Estimating the class prior and posterior from noisy positives and unlabeled data." *in Porc. NIPS*, Barcelona, Spain, 2016. pp. 2685-2693.

Bin Zhang received the B. Eng. Degree in Electronical and Information Engineering from Xiangtan University, China; the M. Eng and Ph.D in Control Science and Engineering from Central South University, China in 2010 and 2016, respectively. She is currently a lecture in Guangdong University of Technology, China. Her main research interests include modeling and optimal control of complex industrial process, and computational intelligence.

Chunhua Yang received her B. Eng. and M.Eng. in Automatic Control Engineering and her Ph.D. in Control Science and Engineering from Central South University, China in 1985, 1988 and 2002 respectively, and was with the Electrical Engineering Department, Katholieke Universiteit Leuven, Belgium from 1999 to 2001. She is currently a full professor in the School of Information Science & Engineering, Central South University. Her research interests include modeling and optimal control of complex industrial process, intelligent control system, and fault-tolerant computing of real-time systems.

Hongqiu Zhu obtained his MSc and PHD degree in Control Science and Engineering from from Central South University, China, in 2002 and 2010 respectively. He worked in Central South University from 2006. Presently he is working at Central South University as an associate professor in the School of Information Science and Engineering. His research work includes information processing and optimal control of complex industrial process.

Peng Shi received the PhD degree in Electrical Engineering from the University of Newcastle, Australia in 1994; the PhD degree in Mathematics from the University of South Australia in 1998. He was awarded the Doctor of Science degree from the University of Glamorgan, Wales in 2006; and the Doctor of Engineering degree from the University of Adelaide, Australia in 2015.

Dr Shi is now a professor at the University of Adelaide. His research interests include system and control theory, computational intelligence, and operational research. He is a co-recipient of the Andrew Sage Best Transactions Paper Award from IEEE Systems, Man and Cybernetics Society in 2016. He has served in the editorial board of a number of journals, including Automatica, IEEE Transactions on Automatic Control; IEEE Transactions on Fuzzy Systems; IEEE Transactions on Cybernetics; IEEE Transactions on Circuits and Systems; and IEEE Access. He now serves as an IEEE Distinguished Lecturer. He is a Fellow of the Institute of Electrical and Electronic Engineers; and Institution of Engineering and Technology.

Weihua Gui received the degree of the B.Eng. in Electrical Engineering and the M.Eng. in Automatic Control Engineering from Central South University, China in 1976 and 1981 respectively. From 1986 to 1988 he was a visiting scholar at Universität-GH-Duisburg, Germany. He has been a full professor in the School of Information Science & Engineering, Central South University since 1991. His main research interests are in modeling and optimal control of complex industrial process, distributed robust control, and fault diagnoses.