IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 10, 2019, accepted October 10, 2019, date of publication October 15, 2019, date of current version October 28, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2947499

Abnormal Condition Identification Modeling Method Based on Bayesian Network Parameters Transfer Learning for the Electro-Fused Magnesia Smelting Process

PING YUAN[1], YUFENG SUN[1], HUI LI[1], FULI WANG[1,2], AND HONGRU LI[1]
[1]Information Science and Engineering, Northeastern University, Shenyang 110819, China
[2]State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China

Corresponding author: Hui Li (lihui_neu@163.com)

This work was supported in part by the National Nature Science Foundation of China under Grant 61533007, Grant 61873049, and Grant 61973057, in part by the Foundation for Innovative Research Groups of the National Natural Science Foundation of China under Grant 61621004, and in part by the National Key Research and Development Program of China under Grant 2017YFB0304205.

ABSTRACT When the data of target domain are scarce, the established model will not be accurate enough to analyze the target problem. For the abnormal condition identification modeling problem of electro-fused magnesia smelting process, this paper proposes the new Bayesian network (BN) parameters transfer learning method based on the expert knowledge from target domain to increase the accuracy of abnormal condition identification. First of all, the electro-fused magnesia smelting process is introduced and the existing research results on the abnormal condition identification are analyzed. The problem to solve in this paper is described. Furthermore, the constraints from expert knowledge for the target model are shown in two forms. The new BN parameters transfer learning method is proposed. Finally, the proposed method is verified by the Asia network, and it is applied to establish the abnormal exhausting condition identification model for the electro-fused magnesia smelting process. The simulation results demonstrate the effectiveness of proposed method which owns the better performances.

INDEX TERMS Bayesian network, transfer learning, fused magnesium furnace, abnormal conditions identification, expert knowledge.

I. INTRODUCTION

In the electro-fused magnesia smelting process, the raw materials are dumped into the fused magnesium furnace (FMF) every ten to fifteen minutes. Because the raw materials come from the different areas, the qualities of dumped raw materials at the different instants of time may have a big difference. The FMF completes the smelting process by the control system tracking the different current setpoints based on the different conditions. When the raw material granule sizes and impurity constituent change, and the setpoints of electrode currents are not adjusted properly, the abnormal conditions will happen. The process monitoring and abnormal condition identification problems have obtained broad attention [1]–[3]. The occurrence of abnormal conditions will result in high energy consumption, performance deterioration or even safety threat. Therefore, it is necessary to analyze and identify the abnormal conditions, and avoid the serious consequences. The occurrence of abnormal conditions shows a certain randomness, which depends on the temporal quality of raw material and operation condition. As the improvement of control algorithm, the occurrence rate of abnormal conditions is reduced. This condition is very good for the production of fused magnesium, but it will result in that the collection of abnormal conditions data becomes very difficult. Especially for the single FMF, the abnormal condition data are very limited. At the same time, it is impractical to collect the abnormal conditions data by producing the abnormal conditions deliberately. The shortage of abnormal condition data has a strong impact on the accuracy of established model.

Some research results on the abnormal condition identification and safety control have been proposed for the

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Wang.

electro-fused magnesia smelting process [4]–[8]. The papers [7], [8] proposed the abnormal conditions identification method based on Bayesian network (BN) by fusing the related multi-source information including current, image and sound. The degree of abnormal condition was distinguished effectively. The existing research results are based on the sufficient data information. However, when the acquired information is not sufficient or it is hard to collect information, it is difficult to establish the BN model effectively. Under this situation, the abnormal conditions data from the other FMFs in the same factory and/or other factories which produce fused magnesium provide the available information. If the data are applied to establish the abnormal condition model for the target FMF directly, the established model might be inaccurate, because the underlying parameters distributions for the different FMFs might be different. To establish the effective abnormal condition identification model for the target FMF, the transfer learning provides us a new way.

Transfer learning is also known as domain adaptation, which aims to significantly reduce data requirements by leveraging data from the related sources. Transfer learning has been successfully applied in a variety of fields, such as, classification [9], [10], filtering [11], recognition [12], [13], fault diagnosis [14], prediction [15], [16] and optimal control [17]. Transfer learning methods using computational intelligence have been shown as a good survey by [18]. Among the computational intelligence methods, BN is an important knowledge expression and inference method, which has attracted more and more attention. The methods mainly focus on the structure learning [19]–[22] and parameters learning [23]–[25]. The paper [19] elicited expert knowledge by the interpretive structural model method to enhance the BN structure learning method (K2 algorithm). The paper [21] proposed the approximate structure learning algorithms for BNs and discussed the preparation of the scores and structure optimization. The paper [22] proposed a hybrid framework to combine a local structure learning algorithm (either constraint-based or score-based) with a data-driven symmetry correction method of the other type. The paper [23] developed the MiniMax Fitness method to avoid overfitting and improve the fitness between parameters and data. The paper [25] proposed the weighted super parameters of Dirichlet distribution algorithm to perform probability learning. However, in the BN context, the transfer learning studies are limited relatively [26]–[33]. For the BN structure transfer learning, the paper [26] presented a novel weighted sum method based on the conditional independence measures. The papers [27], [28] proposed the BN structure learning algorithm for the related tasks based on the score-based method. For the BN parameters transfer learning, the paper [29] introduced the BN parameters transfer learning algorithm based on both network and fragment (sub-graph) relatedness. The paper [26] introduced the distance based linear pooling and local linear pooling probability aggregation methods. However, the proposed method only considered the influences of conditional probability tables (CPTs) entry size and dataset size. When measuring the weights of different sources, the fitness to the target domain was ignored. In addition, the expert knowledge plays the important role in the process of BN parameters learning. For the evaluation of task-relatedness in multitask BN structure learning, the paper [30] incorporated the domain knowledge to relax assumption condition. By integrating the knowledge transfer and expert constraints, the paper [31] presented a new BN parameters transfer learning method. The paper [34] provided the constrained maximum a posteriori (CMAP) method to learn parameters by incorporating convex constraints derived from expert judgments. To further improve this method, a type of constrained Bayesian Dirichlet priors was presented, which was compatible with the given constraints. Furthermore, for the incomplete data, an improved expectation maximum algorithm was proposed by combining with the CMAP method. The paper [35] proposed a constrained Bayesian estimation approach to learn parameters by incorporating constraints deduced from expert judgments and Dirichlet priors.

Inspired by the expert knowledge and transfer learning, the motivation of this paper is to establish the abnormal condition identification BN model by applying the transfer learning thought for the target FMF with scarce abnormal condition data. Based on the research results in the papers [7], [8], the modeling problem is transformed into the BN parameters transfer learning problem with the expert knowledge constrains for the electro-fused magnesia smelting process. By extracting the expert knowledge of the target domain as the constraints, the alternative source information is chosen and the similarity of source is distinguished. The new BN parameters transfer learning method is proposed. To evaluate the proposed method, the Asia network is used to show and compare the effects before and after transfer learning. Finally, taking the abnormal exhausting condition as the example, the proposed method is applied to establish the abnormal condition identification model for the electro-fused magnesia smelting process. The simulation results demonstrate that the proposed method owns the better identification performance than the modeling method using the scarce abnormal condition data of single target FMF.

The contributions of this paper reflect two aspects. On the one hand, this paper proposes the new BN parameters transfer learning method by integrating the expert knowledge and data information. On the other hand, when the target FMF is with scarce abnormal condition data, the proposed method is applied to solve the abnormal condition identification modeling problem for the electro-fused magnesia smelting process.

The remaining sections of this article are organized as follows. The problem that needs to be solved in this paper is described in the Section II. The new BN parameters transfer learning method is proposed in the Section III. In the Section IV, a set of simulation results are shown to evaluate the proposed algorithm for the Asia network. Furthermore, the proposed method is applied to establish the abnormal exhausting condition identification model in the electro-fused

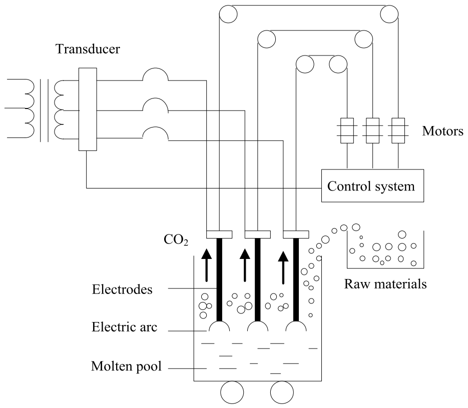


FIGURE 1. **The electro-fused magnesia furnace smelting process.**

magnesia smelting process. Finally, the Section V concludes the paper.

II. PROBLEM DESCRIPTION

The simplified schematic diagram of electro-fused magnesia smelting process is shown in the Figure 1. The FMF completes the smelting process by the control system tracking the different current setpoints based on the different conditions. BN is a valuable tool to model the uncertainty problem and obtain decisions in practice. For the abnormal condition identification modeling of electro-fused magnesia smelting process, BN is an effective method to solve the target problem. The expert knowledge can be used to define variables and determine the network structure in the absence of data. Nevertheless, the parameters need to be learned from the abnormal condition data. In the existing research results [7], [8], the abnormal conditions have been analyzed deeply and the corresponding BN model has been established. However, these models are established under the condition that the abnormal conditions data are enough. When there are limited abnormal condition data for the single target FMF, we attempt to learn the CPTs of target FMF using the abnormal condition information from other FMFs as sources. To improve the performance of parameters learning, the expert can specify the qualitative and quantitative relationships on the CPTs as constraints for the target domain. Therefore, in this paper we focus on the BN parameters transfer learning under the constraints on CPTs from expert knowledge for the target domain.

The problem to solve can be described as the following form. The target domain is the single FMF with scarce abnormal condition data. The source domains include the abnormal condition data from the other FMFs in the same factory and/or other factories which produce fused magnesium. In this background, the tasks of target and sources are the same, so the target and sources own the same variables and the BN structures own the same graphs. Only the parameters have different distribution relationships. Therefore, the abnormal condition identification modeling problem is transformed into the BN parameters transfer learning problem. We concentrate on learning the target network parameters from multiple related sources. We need to judge whether to transfer the parameter or not, evaluate the similarity of parameters from multiple sources and determine the weight to avoid negative transfer. We need to propose the transfer learning scheme to fuse the parameters of sources and target to improve the performance of parameters learning.

In the BN parameters learning setting, a domain $D = \{\mathrm{V}, \mathrm{G}, \mathrm{Da}\}$ includes three components: the variables $\mathrm{V} = \{\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3, \ldots, \mathrm{X}_n\}$ represent the BN nodes, Da represents the associated data, and G represents a directed acyclic graph which encodes the statistical dependencies among the variables. The CPTs specify the probability $p(\mathrm{X}_i|\mathrm{pa}(X_i))$ of every variable given its parents as defined by graph G. The goal of parameters learning is to determine all $p(\mathrm{X}_i|\mathrm{pa}(X_i))$. The parameters can be obtained by maximum likelihood estimation (MLE). In this paper, we have one target domain D^t, and a set of sources $\{D_1^s, D_2^s, \ldots, D_L^s\}$ $(L \geq 1)$. The target domain and each source domain have training data $\mathrm{Da}^t = \{d_1^t, d_2^t, \ldots, d_N^t\}$ and $\mathrm{Da}_i^s = \{d_1^s, d_2^s, \ldots, d_M^s\}$ $i \geq 1$. For transfer learning, we are interested in the case where target domain data are relatively scarce $0 < N \ll M$, and/or N is small relative to the dimensionality of target problem. BN parameters transfer learning aims to improve the parameters learning accuracy of BN in D^t using the information in $\{D_1^s, D_2^s, \ldots, D_L^s\}$ $(L \geq 1)$ [29]. Therefore, BN parameters transfer learning is defined as:

Given a set of source domains $\{D_1^s, D_2^s, \ldots, D_L^s\}$ $(L \geq 1)$ and a target domain D^t, the target domain parameters θ^t are estimated by

$$\hat{\theta}^t = \arg\max_{\theta^t} p(D^t, \{D_1^s, D_2^s, \ldots, D_L^s\}|\theta^t) \tag{1}$$

The following conditions are assumed: $\mathrm{V}^t = \mathrm{V}^s, \mathrm{G}^t = \mathrm{G}^s$, $\{D_1^s, D_2^s, \ldots, D_L^s\}$ $(L \geq 1)$ and D^t own different distribution properties.

III. A NEW BN PARAMETERS TRANSFER LEARNING METHOD

A. THE PARAMETERS CONSTRAINTS FROM EXPERT KNOWLEDGE

Two forms of constraints from expert knowledge are used to express the parameters information of target domain. One kind is the qualitative form which is the relationship expression between the different parameters. The expert knowledge of this form is comparatively accurate, which can be used to judge whether the parameter information is transferred or not. Another kind is the quantitative form which is the range value of parameter. This kind of expert knowledge is suitable to determine the role of source in the transfer learning. The specific forms can be expressed as follows.

Form one: The probabilistic relationship of the same node at the different states can be expressed as the following form

$$\theta_{ijk'}^t > \theta_{ijk''}^t \tag{2}$$

where $1 \leq i \leq n$, n represents the number of nodes in BN model; $1 \leq j \leq q_i$, q_i represents the number of candidate combination states of father nodes of ith node;k' $(1 \leq k' \leq r_i)$ and $k''(1 \leq k'' \leq r_i)$ are the different states of ith node and r_i represents the number of candidate states of ith node; $\theta^t_{ijk'}$ represents the real parameter of target domain, which satisfies the condition that the ith node is in the k'th state and its father node set is in the jth state; $\theta^t_{ijk''}$ represents the real parameter of target domain, which satisfies the condition that the ith node is in the k''th state and its father node set is in the jth state.

Form two: The value range of one node at single state can be expressed as the following form

$$a_k < \theta^t_{ijk} < b_k \tag{3}$$

where θ^t_{ijk} represents the real parameter of target domain, which satisfies the condition that the ith node is in the kth state and its father node set is in the jth state; a_k $(0 \leq a_k \leq 1)$ and b_k $(0 \leq b_k \leq 1)$ are the constants, which need to be set in advance. They can be obtained by the expert knowledge, operation experiences and statistical analysis of historical dataset.

B. THE PROPOSED PARAMETERS TRANSFER LEARNING STRATEGY

Central two challenges for the transfer learning are how to measure the similarity between the multiple sources and target to distinguish the relatedness and how to transfer the information from the multiple sources to learn the target model. In this section, the new BN parameters transfer learning strategy is proposed to solve the above problems. The specific process is depicted in the Figure 2, and the steps are shown as follows:

Step 1: The expert knowledge Form one is used to decide that the parameter information of which source is applied to transfer learning.

If the parameter in one source conforms to the expert knowledge Form one in the target domain, this parameter is regarded as the alternative information for the transfer learning. Otherwise, it is not considered as the alternative information for the transfer learning. Above operation is implemented on each group of parameters for ith node whose father node set is in the jth state. Each group of parameters of the same source is evaluated separately to determine whether the information is used to transfer learning or not. Therefore, it is possible that some parameters are used to transfer and other parameters are not used to transfer for the same source.

Step 2: The similarities of alternative sources are distinguished based on the expert knowledge Form two.

In the process of transfer learning, the negative transfer problem needs to be focused. The motivation of distinguishing the similarities of multiple sources is to avoid the 'negative transfer' risk and ensure the effectiveness of transfer learning. The similarity of alternative source is calculated based on the expert knowledge Form two. If the source satisfies more constraints from expert knowledge, it is more likely to own the similar probability distribution to the target domain. When the corresponding parameter of alternative source satisfies one constraint from the expert knowledge Form two, the similarity score for this source is added one. For the more important constraints, the bigger score is added to the similarity score. By this kind of scoring way, the similarities of multiple sources are distinguished.

Step 3: The weights of alternative sources for the fusion algorithm are calculated.

The weights of alternative sources for the fusion algorithm are calculated by the following equality:

$$\omega^l_{ijk} = score_l \Big/ \sum_{l=1}^{L'} score_l \tag{4}$$

where ω^l_{ijk} represents the weight of lth source; $score_l$ represents the similarity score of lth source; $L'(L' \leq L)$ represents the number of alternative sources.

Step 4: Before transfer learning, the parameters $\tilde{\theta}^t$ of target domain are obtained by the target data Da^t based on MLE.

$$\tilde{\theta}^t = \arg\max_{\theta^t} p(D^t|\theta^t) \tag{5}$$

Step 5: Judge whether the parameters $\tilde{\theta}^t$ conform to the expert knowledge Form one or not; If the answer is 'Yes', shift to Step 6; If the answer is 'No', shift to Step 7.

Step 6: The final parameters of target domain are calculated by the fusion algorithm (6).

The partial matched parameters $\tilde{\theta}^t$ are used to obtain the final target parameters. The fusion function is shown in the following form:

$$\hat{\theta}^t_{ijk} = (1-\eta)\tilde{\theta}^t_{ijk} + \eta \sum_{l=1}^{L'} \omega^l_{ijk}\theta^l_{ijk} \tag{6}$$

where η $(0 \leq \eta < 1)$ represents the weight of parameters obtained by the source data, which is determined by the expert knowledge; θ^l_{ijk} represents the obtained parameters by the lth alternative source; $\hat{\theta}^t_{ijk}$ represents the final obtained parameters of target by the alternative sources and target information.

Step 7: The final parameters of target domain are calculated by the weighted sum of obtained parameters by the alternative sources.

The fusion function is shown in the following form:

$$\hat{\theta}^t_{ijk} = \sum_{l=1}^{L'} \omega^l_{ijk}\theta^l_{ijk} \tag{7}$$

Remark 1: In the proposed BN parameters transfer learning strategy, only the available expert knowledge in two forms are used to determine whether the parameter is transferred or not and the weights of alternative sources for transfer learning. It is not the necessary condition that all the constraints for all the parameters of target domain are all known. Every parameter of target domain is calculated in the proposed strategy separately, because the alternative sources for different parameters may be different. In the proposed BN parameters

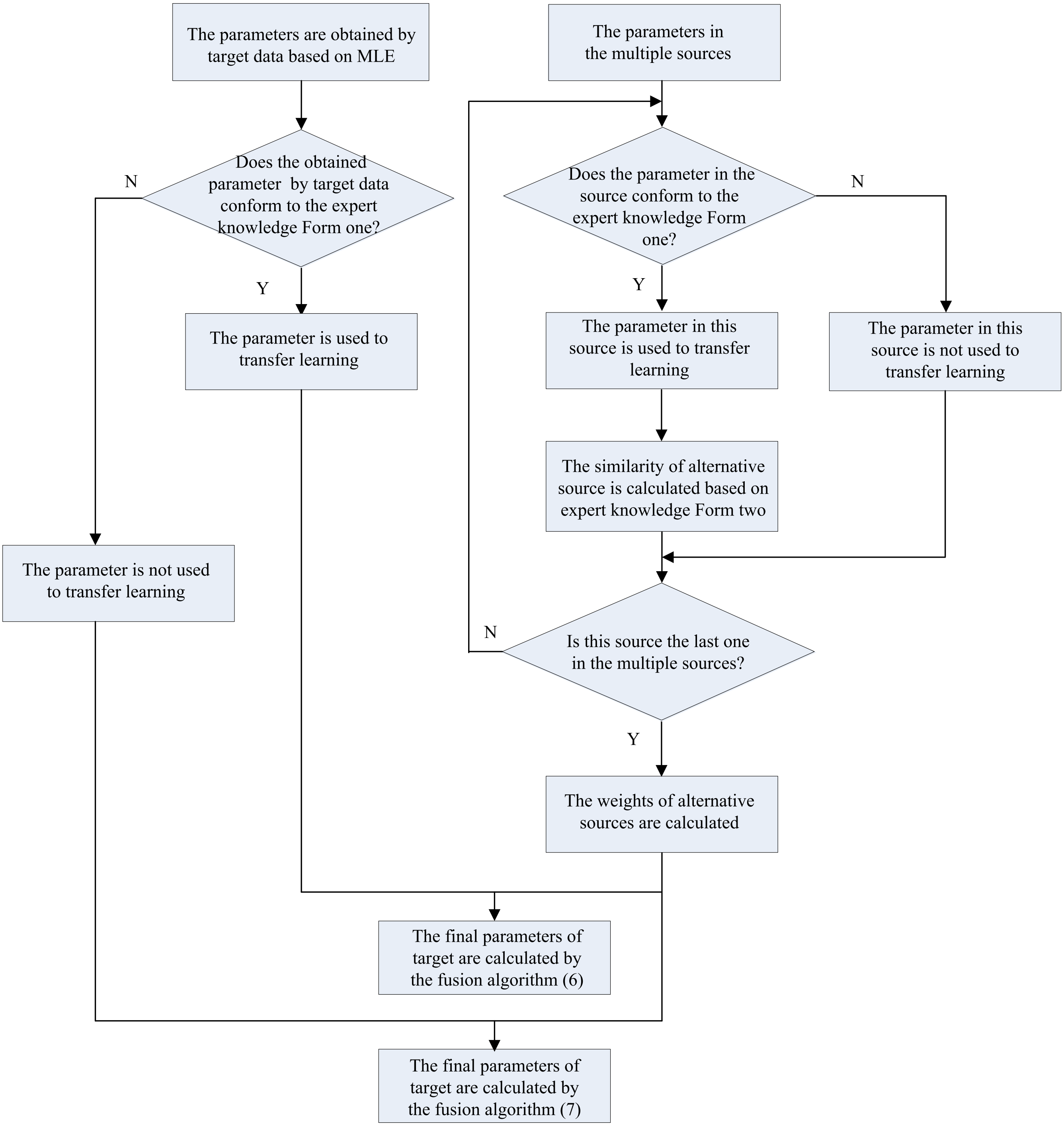


FIGURE 2. The BN parameters transfer learning strategy.

transfer learning strategy, although all source domains are executed sequentially, this algorithm is not related with the sequence. If the sequence of source domains is changed, the final result does not be affected.

Remark 2: Compared with the latest related research result [31], the advantages of proposed method reflect in the following aspects:

(a) In the proposed method, the constrains from the expert knowledge are expressed in two forms: qualitative form and quantitative form. This expression way is simpler and more intuitionistic. Although there is a degree of similarity for the expert constraints in form between the paper [31] and the proposed method, the specific meaning and function are different.

In this paper, the expert knowledge is divided into two types. When the experts only know the qualitative relationship among the variables or parameters, the knowledge is expressed in the qualitative form. When the experts know the quantitative values or range for the variables or parameters, the knowledge is expressed in the quantitative form. The qualitative expert knowledge is used to judge which source is applied to transfer learning. The quantitative expert knowledge is used to distinguish the similarities of alternative sources and determine the weights of sources.

(b) In this paper, the expert constraints are used to judge which source parameter is chosen to transfer learning and the weights of alternative sources. In the paper [31], the expert constraints only are used as the nodes in the extended multinomial parameter learning model.

(c) In this paper, the information of alternative sources which satisfy the expert constraints are all used to transfer learning. After distinguishing the similarities of alternative sources, the parameters of target are obtained by the fusion function (6) or (7). The method is easy to operate. In the paper [31], the transfer learning is exploited by measuring the relatedness of every source fragment. The best source fragment is chosen as parameter priors in the target model. Therefore, only the information of one source is applied in the transfer learning. The final parameters of target are determined by inference based on the given target data, target constraints and source networks.

(d) In this paper, to guarantee the effectiveness of transfer learning for the electro-fused magnesia smelting process, we regard other FMFs in the same factory and/or other factories which produce fused magnesium as the sources. The backgrounds and researched problems are the same, therefore, $\mathrm{V}^t = \mathrm{V}^s$, and $\mathrm{G}^t = \mathrm{G}^s$. $\{D_1^s, D_2^s, \ldots, D_L^s\}$ $(L \geq 1)$ and D^t may own different distribution properties. These chosen sources avoid the negative transfer in some degree. In the paper [31], in the process of transfer learning, the method allows the heterogeneity $V^t \neq V^s$ and $G^t \neq G^s$, and the fragment needs to satisfy the fragment compatibility. The possibility of negative transfer is bigger relatively.

IV. EXPERIMENTAL RESULTS

At first, we evaluate the proposed BN parameters transfer learning method on the well-known Asia network [36], [37]. The Asia network is a diagnostic demonstrative BN, which is used to represent the relationships among the relevant variables for medical knowledge related to the shortness of breath (Chest Clinic). Furthermore, the proposed method is used to establish the abnormal exhausting condition model in the electro-fused magnesia smelting process to identify the abnormal condition, which is compared with the parameters learning method only using limited abnormal condition data of target FMF to verify the superiority.

A. SIMULATION RESULTS ON THE ASIA NETWORK

The network structure of Asia network is shown in the Figure 3. The CPTs of Asia network are shown in the Tables 1-7.

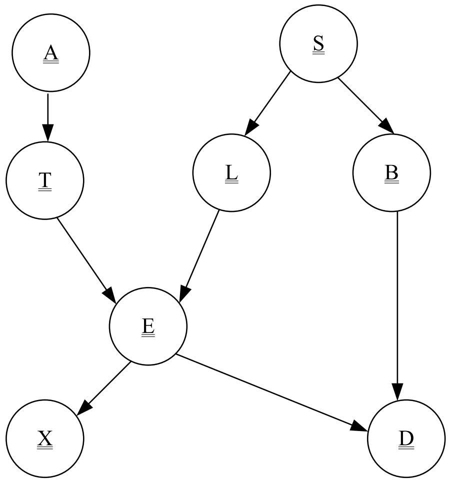


FIGURE 3. The structure of Asia network.

TABLE 1. The CPTs of nodes A and S.

	A(i=1)	S(i=2)
0(k=1)	0.99	0.5
1(k=2)	0.01	0.5

TABLE 2. The CPT of node T.

		A	
		0 (j=1)	1 (j=2)
T(i=3)	0(k=1)	0.99	0.95
	1(k=2)	0.01	0.05

TABLE 3. The CPT of node L.

		S	
		0 (j=1)	1 (j=2)
L(i=4)	0(k=1)	0.99	0.9
	1(k=2)	0.01	0.1

TABLE 4. The CPT of node B.

		S	
		0 (j=1)	1 (j=2)
B(i=5)	0(k=1)	0.7	0.4
	1(k=2)	0.3	0.6

The dataset with 300 samples from the true Asia network is used as the target data information. Four sources networks with the same structure and different probability distribution characteristics are constructed. The expert knowledge Form

TABLE 5. The CPT of node E.

	T	0		1	
	L	0 (j=1)	1 (j=2)	0 (j=3)	1 (j=4)
E(i=6)	0(k=1)	1	0	0	0
	1(k=2)	0	1	1	1

TABLE 6. The CPT of node X.

		E	
		0 (j=1)	1 (j=2)
X(i=7)	0(k=1)	0.95	0.02
	1(k=2)	0.05	0.98

TABLE 7. The CPT of node D.

	E	0		1	
	B	0 (j=1)	1 (j=2)	0 (j=3)	1 (j=4)
D(i=8)	0(k=1)	0.9	0.2	0.3	0.1
	1(k=2)	0.1	0.8	0.7	0.9

TABLE 8. The expert knowledge form two for Asia network.

Number	Constraint	Score
1	$0 < \lvert \theta^t_{2\_1} - \theta^t_{2\_2} \rvert < 0.02$	1
2	$0.38 < \theta^t_{521} < 0.42$	1
3	$0.01 < \theta^t_{721} < 0.03$	2
4	$0.18 < \theta^t_{821} < 0.22$	2

one includes $\theta^t_{1\_1} > \theta^t_{1\_2}$ and $\theta^t_{641} < \theta^t_{642}$. The expert knowledge Form two is shown in the Table 8.

For the BN parameters learning, the Kullback-Leibler (KL) divergence is used to measure how closely the learned parameters with the true parameters and evaluate the performance of parameters learning. The specific form of KL divergence is shown as follow.

$$D_{KL}(\theta^t_{ijk} || \hat{\theta}^t_{ijk}) = \sum_{i=1}^{n} \sum_{j=1}^{q_i} \sum_{k=1}^{r_i} \theta^t_{ijk} * (\log \theta^t_{ijk} - \log \hat{\theta}^t_{ijk})$$

The smaller the value of KL divergence is, the better the learned parameter is.

First of all, the transfer learning method is not applied, and the parameters of target are learned only using target data. The KL divergence value is 41.4487. This value is very big and it is obvious that the performance of parameters learning is very bad when only using the target data. What's more, to show the performance of proposed method in the Section III, the parameters of target are learned by the proposed BN parameters transfer learning method. The values of KL divergence under the different weights η are shown in the Figure 4. From the Figure 4, it can obtain that as the weight η increases, the role of alternative sources for transfer learning increases, the value of KL divergence is smaller and the learned parameters are closer to the true parameters. Therefore, the proposed BN parameters transfer learning method is feasible to obtain good parameters and the learned parameters are better than the parameters learning only using target data.

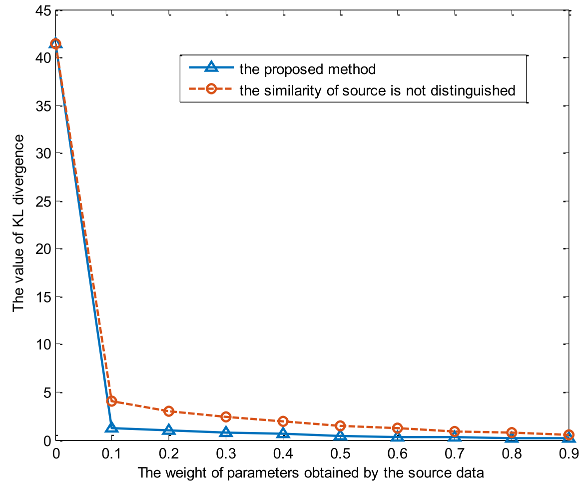


FIGURE 4. The values of KL divergence under the different weights η.

To verify the necessity of distinguishing the similarities of sources, the values of KL divergence under the different weights η are shown in the Figure 4 when the similarities of sources are not distinguished.

Comparing the results in the Figure 4, when the similarities of sources are not distinguished, the obtained values of KL divergence are all bigger than the proposed method for all the weights η. It can conclude that the proposed similarity measure method of sources is effective and can avoid the influence of bad information for transfer learning.

The numerical experimental results on the Asia network show that our proposed BN parameters transfer learning method is effective to learn the parameters of target and owns the better performance than the method only using target data. The proposed similarity measure method of sources avoids the influence of negative transfer.

B. SIMULATION RESULTS ON THE ELECTRO-FUSED MAGNESIA SMELTING PROCESS

The previous section demonstrates the effectiveness of our proposed BN parameters transfer learning method. In this section we explore its application on the abnormal exhausting condition modeling for the electro-fused magnesia smelting process. The simulation is carried out by the hardware in the loop simulation platform for the electro-fused magnesia smelting process. The simulation platform is designed and

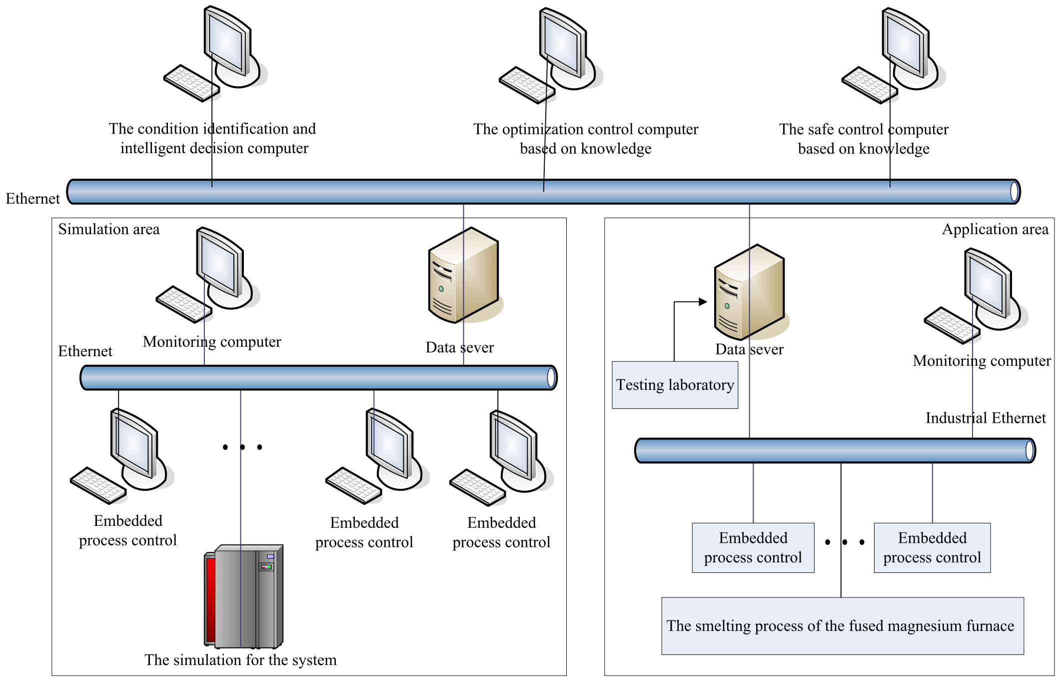


FIGURE 5. The hardware in the loop simulation platform for the electro-fused magnesia smelting process.

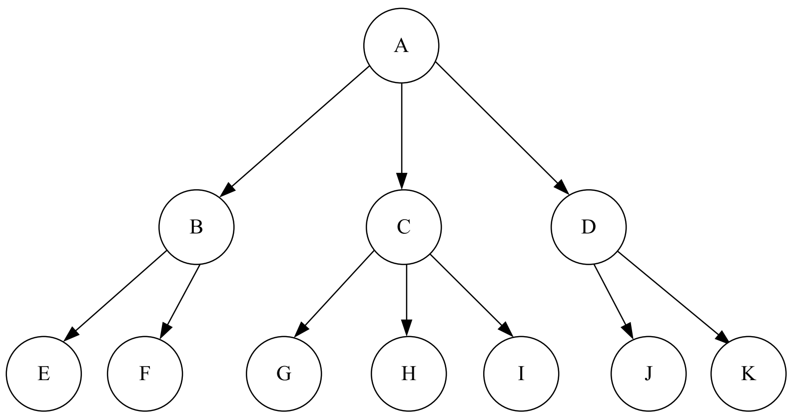


FIGURE 6. The established BN structure for the abnormal exhausting condition.

constructed in my research team during the past few years, which is shown in the Figure 5.

We can research and validate the proposed algorithm on this platform. The simulation platform can simulate the electro-fused magnesia smelting process based on mechanism analysis and actual data, which can complete optimal control, abnormal condition identification and safety control for this process. The platform includes the following equipment: a number of computers with different functions, embedded process control systems, data sever, sensing devices and transferring devices. The sensing devices include current measurement instrument, image measurement instrument and sound measurement instrument. On site, the multi-source information is collected by the sensing devices and stored in the data sever. The transferring devices transfer the information to the simulation area by the Ethernet in a certain time interval. Because the data which are used to verify the proposed method are based on the simulation of practical production process, if the proposed method is verified, it could be considered to apply in the practical production process.

Based on the research results on the abnormal condition identification for the abnormal exhausting condition [7], [8], the established BN structure for the abnormal exhausting condition is shown in the Figure 6. The practical meanings of nodes are shown in the Table 9. The numbers of nodes "A-K" in the Figure 6 are "1-11". In the established BN model, the current, image and sound characteristics are used as the nodes. These characteristics are related with the abnormal exhausting condition. As the further research, if other kinds of sensor source information are discovered that they

TABLE 9. The characteristics for the abnormal exhausting condition.

Node	Meaning
A	The abnormal exhausting condition
B	The abnormal sound signal
C	The abnormal image signal
D	The abnormal current signal
E	The energy in a short time for the splattering characteristic frequency
F	The amplitude for the splattering characteristic frequency
G	The change of the average gradation
H	The change of the gradation variance
I	The abundance of the gradation
J	The current change rates
K	The current tracking errors

TABLE 10. The parameters expression of node A.

	A(i=1)
1(k=1)	$\theta^t_{1\_1}$
2(k=2)	$\theta^t_{1\_2}$
3(k=3)	$\theta^t_{1\_3}$
4(k=4)	$\theta^t_{1\_4}$

TABLE 11. The parameters expression of node B.

		A(i=1)			
		1 (j=1)	2 (j=2)	3 (j=3)	4 (j=4)
B(i=2)	1(k=1)	θ^t_{211}	θ^t_{221}	θ^t_{231}	θ^t_{241}
	2(k=2)	θ^t_{212}	θ^t_{222}	θ^t_{232}	θ^t_{242}
	3(k=3)	θ^t_{213}	θ^t_{223}	θ^t_{233}	θ^t_{243}

TABLE 12. The expert knowledge Form one for the abnormal exhausting condition model.

Number	Constraint
1	$\theta^t_{232} < \theta^t_{233}$
2	$\theta^t_{333} < \theta^t_{332}$
3	$\theta^t_{434} < \theta^t_{433}$
4	$\theta^t_{613} < \theta^t_{612}$
5	$\theta^t_{721} < \theta^t_{723}$
6	$\theta^t_{1031} < \theta^t_{1032}$

are also related with the abnormal exhausting condition, the new characteristics could be extracted as the new BN nodes. To express the constraints from the expert knowledge for the target FMF, the examples of parameters expression for nodes A and B are shown in the Tables 10 and 11. The parameters expression for other nodes is similar to nodes A and B, which is omitted. The expert knowledge Form one and two for the abnormal exhausting condition model are shown in the Tables 12 and 13 respectively.

By analyzing the practical situations of the abnormal exhausting condition, some possible abnormal scenarios are

TABLE 13. The expert knowledge Form two for the abnormal exhausting condition model.

Number	Constraint	Score
1	$0.04 < \theta^t_{341} < 0.06$	1
2	$0.00 < \theta^t_{313} < 0.02$	1
3	$0.02 < \theta^t_{521} < 0.04$	1
4	$0.95 < \theta^t_{711} < 0.97$	1
5	$0.84 < \theta^t_{733} < 0.86$	1
6	$0.18 < \theta^t_{432} < 0.22$	2
7	$0.13 < \theta^t_{932} < 0.17$	2
8	$0.03 < \theta^t_{1021} < 0.07$	2

extracted and shown in the Table 14. In the Table 14, every abnormal scenario includes seven characteristics with different degrees. The characteristics are divided into three (or four) degrees. The numbers 1-3 (or 1-4) are used to represent the different degrees. The characteristics E-J are divided into three degrees: small, medium and large; the characteristic K is divided into four degrees: very small, small, large and very large. Taking the abnormal scenario 18 as the example, the physical meaning is that the states of characteristics E-F are small; the states of characteristics G-I are medium; the states of characteristics J-K are very large. Not all scenarios are included in the Table 14. Only typical scenarios are considered and other similar scenarios can be analyzed in the same way.

Firstly, the parameters of abnormal exhausting condition model are learned by the scarce data of target FMF, and the model is represented as Model one. Furthermore, we collect the abnormal exhausting condition data from other four groups FMFs in the same factory. The parameters of abnormal exhausting condition model are learned by the proposed BN parameters transfer learning method in the Section III based on the constraints from the expert knowledge in the Tables 12 and 13, and the model is represented as Model two.

The possible abnormal scenarios in the Table 14 are used as the evidences to infer by the established Models one and two. The exhausting condition node A is divided into four degrees: normal, sight abnormal, moderate abnormal and serious abnormal, which are represented by numbers 1-4. The reasoning results are shown in the Tables 15 and 16 respectively, which owns the largest posterior probability is regarded as the identification result. The corresponding identification results are in bold. Based on the practical operational experience, by comparing the identification results in the Tables 15 and 16, it can obtain that the accuracy rate of identification for Model one is 50% and the accuracy rate of identification for Model two is 100%. The proposed transfer learning method can make the abnormal exhausting condition model own the higher accuracy rate when the data of target FMF is very limited.

TABLE 14. Some possible abnormal scenarios for the abnormal exhausting condition.

Scenario number		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Sound information	E	1	2	2	3	3	3	3	2	1	1	1	1	1	1	1	1	1	1
	F	1	2	3	2	3	3	2	3	1	1	1	1	1	1	1	1	1	1
Image information	G	1	1	1	1	1	1	1	1	3	3	3	3	3	2	2	2	2	2
	H	1	1	1	1	1	1	1	1	3	3	3	3	3	2	2	2	2	2
	I	1	1	1	1	1	1	1	1	3	3	3	3	3	2	2	2	2	2
Current information	J	1	1	1	1	1	2	2	2	3	3	2	3	2	3	2	2	3	3
	K	1	1	1	1	1	2	2	2	2	4	3	3	4	2	3	4	3	4

TABLE 15. The identification results of abnormal scenarios in the TABLE 14 for model one.

Scenario number		1	2	3	4	5	6
The state of node A	1	**0.6432**	**0.5274**	**0.5215**	**0.5258**	**0.5210**	0.4148
	2	0.1317	0.2066	0.2449	0.2608	0.3042	0.1450
	3	0.2232	0.2646	0.2325	0.2124	0.1738	**0.4318**
	4	0.0019	0.0014	0.0011	0.0010	0.0008	0.0084
Scenario number		7	8	9	10	11	12
The state of node A	1	0.3871	0.3704	0.0036	0.0024	0.0012	0.0012
	2	0.1150	0.1042	0.0105	0.0159	0.0266	0.0183
	3	**0.4881**	**0.5154**	**0.5594**	**0.5431**	**0.5766**	**0.4960**
	4	0.0098	0.0100	0.4265	0.4386	0.3956	0.4845
Scenario number		13	14	15	16	17	18
The state of node A	1	0.0056	0.1596	0.0594	0.2899	0.0553	0.1099
	2	0.0159	0.0160	0.0466	0.0282	0.0284	0.0250
	3	**0.7334**	0.1508	0.1781	0.2305	0.1363	0.1510
	4	0.2451	**0.6736**	**0.7159**	**0.4514**	**0.7800**	**0.7141**

TABLE 16. The identification results of abnormal scenarios in the TABLE 14 for model two.

Scenario number		1	2	3	4	5	6
The state of node A	1	**0.8553**	0.1693	0.1824	0.1787	0.1860	0.0355
	2	0.1334	**0.7894**	**0.7212**	**0.7205**	**0.6515**	0.1559
	3	0.0088	0.0394	0.0950	0.0995	0.1618	**0.8043**
	4	0.0025	0.0019	0.0014	0.0013	0.0007	0.0043
Scenario number		7	8	9	10	11	12
The state of node A	1	0.0481	0.0505	0.0059	0.0013	0.0013	0.0008
	2	0.2431	0.2507	0.0072	0.0040	0.0032	0.0035
	3	**0.6975**	**0.6867**	0.0613	0.0313	0.0358	0.0291
	4	0.0113	0.0121	**0.9256**	**0.9634**	**0.9597**	**0.9666**
Scenario number		13	14	15	16	17	18
The state of node A	1	0.0027	0.1055	0.0275	0.0536	0.0170	0.0269
	2	0.0045	0.0317	0.0159	0.0218	0.0180	0.0203
	3	0.0425	0.1418	0.0954	0.1084	0.0792	0.0838
	4	**0.9503**	**0.7210**	**0.8612**	**0.8162**	**0.8858**	**0.8690**

TABLE 17. The identification results of abnormal scenarios in the TABLE 14 for model three.

Scenario number		1	2	3	4	5	6
The state of node A	1	**0.8090**	0.2993	0.2971	0.2954	0.2964	0.0065
	2	0.1836	**0.6795**	**0.5608**	**0.5405**	**0.4777**	0.1678
	3	0.0055	0.0172	0.1401	0.1624	0.2253	**0.8201**
	4	0.0019	0.0040	0.0020	0.0017	0.0006	0.0056
Scenario number		7	8	9	10	11	12
The state of node A	1	0.0081	0.0089	0.0000	0.0000	0.0000	0.0000
	2	0.2365	0.2691	0.0897	0.0843	0.0828	0.0881
	3	**0.7364**	**0.6966**	0.0500	0.0495	0.0632	0.0487
	4	0.0190	0.0254	**0.8603**	**0.8662**	**0.8540**	**0.8632**
Scenario number		13	14	15	16	17	18
The state of node A	1	0.0000	0.0000	0.0002	0.0002	0.0000	0.0000
	2	0.1035	0.3996	0.3725	0.4260	0.3957	0.3839
	3	0.0722	0.0948	0.1207	0.1262	0.0929	0.0959
	4	**0.8243**	**0.5056**	**0.5066**	**0.4476**	**0.5114**	**0.5202**

TABLE 18. The identification results of abnormal scenarios in the TABLE 14 for model four.

Scenario number		1	2	3	4	5	6
The state of node A	1	**0.7938**	0.1017	0.1165	0.1195	0.1296	0.0049
	2	0.1967	**0.8631**	**0.7697**	**0.7619**	**0.6632**	0.1710
	3	0.0069	0.0324	0.1118	0.1167	0.2062	**0.8176**
	4	0.0026	0.0028	0.0020	0.0019	0.0010	0.0065
Scenario number		7	8	9	10	11	12
The state of node A	1	0.0067	0.0067	0.0005	0.0003	0.0004	0.0002
	2	0.2905	0.3013	0.0184	0.0110	0.0156	0.0106
	3	**0.6843**	**0.6725**	0.0333	0.0244	0.0289	0.0237
	4	0.0185	0.0195	**0.9478**	**0.9643**	**0.9551**	**0.9655**
Scenario number		13	14	15	16	17	18
The state of node A	1	0.0008	0.0043	0.0040	0.0071	0.0023	0.0027
	2	0.0195	0.1322	0.1162	0.1384	0.0843	0.0874
	3	0.0328	0.1897	0.1719	0.1858	0.1502	0.1534
	4	**0.9469**	**0.6738**	**0.7079**	**0.6687**	**0.7632**	**0.7565**

To show the influence of data size of target FMF on transfer learning, the abnormal exhausting condition data of target FMF are increased to establish the Model three, but the used data of target FMF are still less than the data in the source domain. Furthermore, based on the increased target FMF data, the established model is represented as Model four by the proposed BN parameters transfer learning method. The reasoning results are shown in the Tables 17 and 18 respectively. By comparing the identification results in the Tables 17 and 18, it can obtain that as the data size increases in the target FMF, two methods all can obtain true identification results which conform to the practical operational experience. However, for the most abnormal scenarios in the Table 14, the identification results by the proposed BN parameters transfer learning method own the better discrimination, that is to say, the true degree of identification result owns the bigger probability. The difference of probability between the true degree and the wrong degree is bigger. Therefore, by the above analysis and comparison, it can conclude that the obtained abnormal exhausting condition model by the proposed BN parameters transfer learning method owns the better performance than the method only using the data of target FMF.

V. CONCLUSION

This paper develops the new BN parameters transfer learning method for the electro-fused magnesia smelting process. First of all, the electro-fused magnesia smelting process is analyzed. Based on the existing research results, the problem to solve in this paper is described. Furthermore, the new BN parameters transfer learning method is proposed based on the constraints from expert knowledge in two forms.

Finally, the experimental results demonstrate that the proposed BN parameters transfer learning method is effective and owns the better performances than the method only using the scarce data from target. It is beneficial to the establishment of abnormal condition identification model in the electro-fused magnesia smelting process when the abnormal condition data are scarce for the target FMF. The proposed similarity measure method avoids the influence of negative transfer in some degree.

REFERENCES

[1] C. Zhao and B. Huang, "A full-condition monitoring method for nonstationary dynamic chemical processes with cointegration and slow feature analysis," *AIChE J.*, vol. 64, pp. 1662–1681, Jan. 2018.

[2] C. Zhao and H. Sun, "Dynamic distributed monitoring strategy for large-scale nonstationary processes subject to frequently varying conditions under closed-loop control," *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4749–4758, Jun. 2019.

[3] S. Zhang and C. Zhao, "Slow-feature-analysis-based batch process monitoring with comprehensive interpretation of operation condition deviation and dynamic anomaly," *IEEE Trans. Ind. Electron.*, vol. 66, no. 5, pp. 3773–3783, May 2019.

[4] Y. Fu, N. Wang, Z. Wang, Z. Wang, B. Ji, and X. Wang, "Smelting condition identification for a fused magnesium furnace based on an acoustic signal," *J. Mater. Process. Technol.*, vol. 244, pp. 231–239, Jun. 2017.

[5] Y. Fu, Z. Wang, Z. Wang, N. Wang, and X. Wang, "Splattering suppression for a three-phase AC electric arc furnace in fused magnesia production based on acoustic signal," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4772–4780, Jun. 2017.

[6] Z. Wu, Y. Wu, T. Chai, and J. Sun, "Data-driven abnormal condition identification and self-healing control system for fused magnesium furnace," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1703–1715, Mar. 2015.

[7] H. Li, F. Wang, and H. Li, "Abnormal condition identification and safe control scheme for the electro-fused magnesia smelting process," *ISA Trans.*, vol. 76, pp. 178–187, May 2018.

[8] H. Li, F. Wang, and H. Li, "Abnormal condition identification for the electro-fused magnesia smelting process," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 720–725, 2018.

[9] I. Chaturvedi, Y.-S. Ong, and R. V. Arumugam, "Deep transfer learning for classification of time-delayed Gaussian networks," *Signal Process.*, vol. 110, pp. 250–262, May 2015.

[10] M. Talo, U. B. Baloglu, Ö. Yıldırım, and U. R. Acharya, "Application of deep transfer learning for automated brain abnormality classification using MR images," *Cogn. Syst. Res.*, vol. 54, pp. 176–188, May 2019.

[11] W. Cai, J. Zheng, W. Pan, J. Lin, L. Li, L. Chen, X. Peng, and Z. Ming, "Neighborhood-enhanced transfer learning for one-class collaborative filtering," *Neurocomputing*, vol. 341, pp. 80–87, May 2019.

[12] D. Malmgren-Hansen, A. Kusk, J. Dall, A. A. Nielsen, R. Engholm, and H. Skriver, "Improving SAR automatic target recognition models with transfer learning from simulated data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 9, pp. 1484–1488, Sep. 2017.

[13] L. Zhang, Y. Liu, and P. Deng, "Odor recognition in multiple e-nose systems with cross-domain discriminative subspace learning," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1679–1692, Jul. 2017.

[14] B. Yang, Y. Lei, F. Jia, and S. Xing, "An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings," *Mech. Syst. Signal Process.*, vol. 122, pp. 692–706, May 2019.

[15] E. Zhong, W. Fan, and Q. Yang, "User behavior learning and transfer in composite social networks," *ACM Trans. Knowl. Discovery from Data*, vol. 8, Feb. 2014, Art. no. 6.

[16] X. Liu, Y. Li, and G. Chen, "Multimode tool tip dynamics prediction based on transfer learning," *Robot. Comput.-Integr. Manuf.*, vol. 57, pp. 146–154, Jun. 2019.

[17] F. Chu, X. Zhao, Y. Yao, T. Chen, and F. Wang, "Transfer learning for batch process optimal control using LV-PTM and adaptive control strategy," *J. Process Control*, vol. 81, pp. 197–208, Sep. 2019.

[18] J. Lu, V. Behbood, P. Hao, H. Zuo, S. Xue, and G. Zhang, "Transfer learning using computational intelligence: A survey," *Knowl.-Based Syst.*, vol. 80, pp. 14–23, May 2015.

[19] L. Huang, G. Cai, H. Yuan, and J. Chen, "A hybrid approach for identifying the structure of a Bayesian network model," *Expert Syst. Appl.*, vol. 131, pp. 308–320, Oct. 2019.

[20] G. Luo, B. Zhao, and S. Du, "Causal inference and Bayesian network structure learning from nominal data," *Appl. Intell.*, vol. 49, pp. 253–264, Jan. 2019.

[21] M. Scanagatta, G. Corani, C. P. de Campos, and M. Zaffalon, "Approximate structure learning for large Bayesian networks," *Mach. Learn.*, vol. 107, pp. 1209–1227, Sep. 2018.

[22] J. Zhao and S.-S. Ho, "Improving Bayesian network local structure learning via data-driven symmetry correction methods," *Int. J. Approx. Reasoning*, vol. 107, pp. 101–121, Apr. 2019.

[23] X.-G. Gao, Z.-G. Guo, H. Ren, Y. Yang, D.-Q. Chen, and C.-C. He, "Learning Bayesian network parameters via minimax algorithm," *Int. J. Approx. Reasoning*, vol. 108, pp. 62–75, May 2019.

[24] K. Masmoudi, L. Abid, and A. Masmoudi, "Credit risk modeling using Bayesian network with a latent variable," *Expert Syst. Appl.*, vol. 127, pp. 157–166, Aug. 2019.

[25] Y. Guo, W. Chen, Y.-X. Zhu, and Y.-Q. Guo, "Research on the integrated system of case-based reasoning and Bayesian network," *ISA Trans.*, vol. 90, pp. 213–225, Jul. 2019.

[26] R. Luis, L. E. Sucar, and E. F. Morales, "Inductive transfer for learning Bayesian networks," *Mach. Learn.*, vol. 79, pp. 227–255, May 2009.

[27] A. Niculescu-Mizil and R. Caruana, "Inductive transfer for Bayesian network structure learning," in *Proc. 11th Int. Conf. Artif. Intell. Statist. (AISTATS)*, San Juan, Puerto Rico, Mar. 2007, pp. 339–346.

[28] D. Oyen and T. Lane, "Transfer learning for Bayesian discovery of multiple Bayesian networks," *Knowl. Inf. Syst.*, vol. 43, pp. 1–28, Apr. 2015.

[29] Y. Zhou, T. M. Hospedales, and N. Fenton, "When and where to transfer for Bayes net parameter learning," *Expert Syst. Appl.*, vol. 55, pp. 361–373, Aug. 2016.

[30] D. Oyen and T. Lane, "Leveraging domain knowledge in multitask Bayesian network structure learning," in *Proc. 26th AAAI Conf. Artif. Intell.*, Toronto, QC, Canada, Jul. 2012, pp. 1091–1097.

[31] Y. Zhou, N. Fenton, T. M. Hospedales, and M. Neil, "Probabilistic graphical models parameter learning with transferred prior and constraints," in *Proc. 31st Conf. Uncertainty Artif. Intell.*, 2015, pp. 972–981.

[32] D. Oyen and T. Lane, "Bayesian discovery of multiple Bayesian networks via transfer learning," in *Proc. IEEE 13th Int. Conf. Data Mining*, Dec. 2013, pp. 577–586.

[33] L. J. Fiedler, L. E. Sucar, and E. F. Morales, "Transfer learning for temporal nodes Bayesian networks," *Appl. Intell.*, vol. 43, pp. 578–597, Oct. 2015.

[34] Y. Yang, X. Gao, Z. Guo, and D. Chen, "Learning Bayesian networks using the constrained maximum a posteriori probability method," *Pattern Recognit.*, vol. 91, pp. 123–134, Jul. 2019.

[35] X. Gao, Y. Yang, and Z. Guo, "Learning Bayesian networks by constrained Bayesian estimation," *J. Syst. Eng. Electron.*, vol. 30, pp. 511–524, Jun. 2019.

[36] S. L. Lauritzen and D. J. Spiegelhalter, "Local computations with probabilities on graphical structures and their application to expert systems," *J. Roy. Statist. Soc. Ser. B, Methodol.*, vol. 50, no. 2, pp. 157–224, 1988.

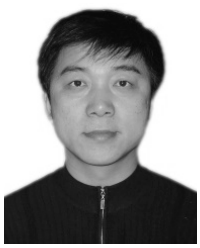
[37] D. W. Kim, S. Ko, and B. Y. Kang, "Structure learning of Bayesian networks by estimation of distribution algorithms with transpose mutation," *J. Appl. Res. Technol.*, vol. 11, pp. 586–596, Aug. 2013.

PING YUAN received the B.S., M.S., and Ph.D. degrees in measurement technology and automation instrument from Northeastern University, Shenyang, China, in 1993, 1996, and 2006, respectively, where he is currently an Associate Professor with the College of Information Science and Engineering, Institute of Automation. His research interests include control systems and machine learning.

YUFENG SUN received the B.E. degree from the University of Science and Technology Liaoning, in 2017. He is currently pursuing the M.E. degree with the College of Information Science and Engineering, Northeastern University, China. His main research interests include transfer learning, Bayesian networks, artificial intelligence, and fault diagnosis.

HUI LI received the B.E. degree from the Shenyang University of Technology, in 2013, and the M.E. degree from Northeastern University, in 2015, where she is currently pursuing the Ph.D. degree with the College of Information Science and Engineering. Her main research interests include transfer learning, Bayesian networks, artificial intelligence and its applications, health maintenance for industrial systems, and self-healing control.

FULI WANG received the B.E., M.E., and Ph.D. degrees from Northeastern University, Shenyang, China, in 1982, 1985, and 1988, respectively, where he is currently a Professor with the College of Information Science and Engineering. His main research interests include modeling, control, and optimization for complex systems, monitoring and health maintenance for industrial systems, artificial intelligence, and fault diagnosis.



HONGRU LI received the B.E., M.E., and Ph.D. degrees in control theory and control engineering from Northeastern University, Shenyang, China, in 1991, 1994, and 2001, respectively, where he is currently a Professor with the College of Information Science and Engineering. His current research interests include modeling, control, and optimization for complex systems, artificial intelligence and its applications, and monitoring and health maintenance for industrial systems.