Front. Chem. Sci. Eng. 2017, 11(3): 414–428
DOI 10.1007/s11705-017-1663-x

RESEARCH ARTICLE

A knowledge reasoning Fuzzy-Bayesian network for root cause analysis of abnormal aluminum electrolysis cell condition

Weichao Yue[1], Xiaofang Chen (✉)[1], Weihua Gui[1], Yongfang Xie[1], Hongliang Zhang[2]

1 School of Information Science and Engineering, Central South University, Changsha 410083, China
2 School of Metallurgy and Environment, Central South University, Changsha 410083, China

Abstract Root cause analysis (RCA) of abnormal aluminum electrolysis cell condition has long been a challenging industrial issue due to its inherent complexity in analyzing based on multi-source knowledge. In addition, accurate RCA of abnormal aluminum electrolysis cell condition is the precondition of improving current efficiency. RCA of abnormal condition is a complex work of multi-source knowledge fusion, which is difficult to ensure the RCA accuracy of abnormal cell condition because of dwindling and frequent flow of experienced technicians. In view of this, a method based on Fuzzy-Bayesian network to construct multi-source knowledge solidification reasoning model is proposed. The method can effectively fuse and solidify the knowledge, which is used to analyze the cause of abnormal condition by technicians providing a clear and intuitive framework to this complex task, and also achieve the result of root cause automatically. The proposed method was verified under 20 sets of abnormal cell conditions, and implements root cause analysis by finding the abnormal state of root node, which has a maximum posterior probability by Bayesian diagnosis reasoning. The accuracy of the test results is up to 95%, which shows that the knowledge reasoning feasibility for RCA of aluminum electrolysis cell.

Keywords abnormal aluminum electrolysis cell condition, Fuzzy-Bayesian network, multi-source knowledge solidification and reasoning, root cause analysis

1 Introduction

Aluminum electrolysis cell, hereafter referred to as cell, is a complex multivariable system, and the occurrence of abnormal cell condition is characterized by multi-source, coupling, randomness. Accurately identify the cause of the abnormal aluminum electrolysis cell condition, hereafter referred to as abnormal cell condition, and the right decisions of judgement for abnormal cell condition play an important role in promoting the high efficiency and stable operation of cell. However, it is difficult to meet the needs of root cause analysis (RCA) of cell with large coupling characteristics and blind area monitoring in the abnormal situation only by the existing detection device. Therefore, the RCA of abnormal cell condition in the production site usually requires the technicians to analyze and judge according to the detection results and the observed state of cell, and combine with the experiential knowledge and process knowledge to find out the root cause and make a correct decision [1,2]. Although aluminum electrolysis companies have operational benchmarks to handle abnormal cell conditions, different cells also have different control effect for the same baseline due to the uneven level of technicians in the process of most cases. This shows that the experience and process knowledge play an important role in the efficient, stable for operation for the cell. However, with the dwindling and frequent flow of experienced technicians, it will inevitably affect the accuracy of RCA for abnormal cell condition. In addition, the complex physical and chemical reactions take place in the environment with strong magnetic field and high corrosion. External conditions and interference of the operation determine the existence of various uncertainties in the running process of aluminum electrolysis. However, these problems are often difficult to use analytical model for accurate description [3–5]. In addition, the difficulty of judgment is also caused by the difference of the knowledge of the technician and the limitation of the testing method. At present, there is a lack of an effective method to

Received February 26, 2017; accepted April 27, 2017

E-mail: xiaofangchen@csu.edu.cn

integrate and consolidate all kinds of knowledge and form an effective mechanism of knowledge solidification under the condition of unified knowledge construction mode. Therefore, it is very important to solidify the decision-making knowledge of aluminum electrolysis process, and to construct a multi-source knowledge solidification reasoning (MSKR) model.

The RCA tool is based on the characteristics of the system to find the cause of the failure [6]. Doggett provided a framework to analyze the performance of three RCA tools: interrelation diagram, cause-and-effect diagram and reality tree [7]. Their framework provided information on performance characteristics of the tools so that decision-makers can understand the approach recommended by the RCA tool well. Demirli and Vijayakumar proposed a control causal relation network model based on fuzzy rules to resolve the uncertainties in identifying the (real) chart patterns and relating them to assignable causes, and it was related to the specified reason. They also discussed that classification the out of control states into isolated shifts, sustained shifts and gradual shifts, which could accelerate the RCA process [8]. Majid et al. developed a new framework based on multi-way principal component analysis to analyze abnormal condition of cell in real-time in the process of alumina feed and pole-exchange [5].

There are many uncertainties in the complex industrial production process, and decision-makers usually need to integrate the experiential and process knowledge. Consequently, combining the uncertainties with those knowledges are necessary for modeling analysis of complex system [2]. In the field of reasoning analysis based on uncertain modeling, the probability graph model has become a hotspot in recent years. It provides a unified frame model for the representation and probability reasoning of multiple variables with coupling, causal relations [9,10]. Bayesian network, as a common probability graph model, can express the coupling, causality and other relations among variables as well as express the multi-source knowledge explicitly to realize the rigorous knowledge reasoning under uncertain conditions [11,12]. Weidl et al. developed a method based on object-oriented Bayesian networks, which integrates decision-theoretic troubleshooting with risk assessment for industrial process control. The method could be able to present to users corrective actions with explanations of the root causes [13]. Supplier selection was a complex decision problem. Ferreira and Borenstein proposed a novel model that could rank and evaluate suppliers by Bayesian reasoning based on the integration of influence diagram and Fuzzy-Bayesian network (FBN). In addition, this model integrated the tacit knowledge of decision-makers [14]. Cai et al. presented a multi-source information fusion model based on Bayesian network to increase the diagnostic accuracy of ground-source heat pump system by Bayesian reasoning. The model was consist of fault source layer and fault feature layer, and the fault feature layer was consist of measurable data with sensors and observable state characteristics with decision-makers [15].

The RCA of abnormal cell condition is a kind of complex knowledge work with multidimensional decision. The research of RCA of abnormal cell condition mainly focus on the methods that are barely based on data-driven or mechanism analysis [5,16]. However, the characteristics of cell, with multivariate, strong coupling and uncertainty, determine that the above methods are difficult to guarantee the accuracy of RCA. The means of RCA depending on technicians are difficult to meet the development requirements of cell with intelligent and refinement in the actual aluminum electrolysis process. In this paper, a multi-source knowledge solidification reasoning model based on fuzzy Bayesian network is constructed for the RCA process of abnormal cell condition, and an effective mechanism for multi-source knowledge unified solidification under the same knowledge representation method is established. The causal relationships of the characteristic variables of aluminum electrolysis process are expressed by MSKR model, as well as can solidify the multi-source knowledge. In addition, the technicians will get the reasoning results after entering the data or state of variable to the model, and achieve the automatic reasoning mechanism of RCA of abnormal aluminum electrolysis cell condition. To our best knowledge, this paper was the first time that proposes the solidification of multi-source knowledge, and the method of RCA automatically based on knowledge reasoning for aluminum electrolysis cell.

The outline of this paper is as follows. Firstly, the knowledge used in the production of aluminum electrolysis and the basis for the technicians' decisions of abnormal cell condition are analyzed, and the difficulty in making the judgment is also analyzed. Secondly, on the basis of those analysis, the feasibility of fuzzy Bayesian network in the RCA of abnormal cell condition is analyzed. The causal framework is given prior probability, forming a fuzzy Bayesian network with the statistic of historical data and the consultation of experienced aluminum electrolysis experts, which constitutes MSKR model. Finally, the effectiveness of the proposed method is verified.

2 Aluminum electrolysis knowledge analysis

Aluminum electrolysis production is based on the characteristics parameters values and visual information, combined with domain knowledge, expert knowledge to achieve effective control of cell. Controlled cell produced aluminum liquid according to the intended target. The molten alumina-cryolite melts is used as the bath, and the anode bar is introduced with direct current. The anodes are introduced into the bath layer and the aluminum liquid layer, and participated in the electrochemical reaction, finally exported through cathode bar. Heat energy is generated between anode and cathode to make the cryolite and alumina melt into ionic state, and to maintain the normal electrolysis temperature.

The final product of the anode is a mixed gas of CO and CO_2. The liquid aluminum is precipitated on the cathode. During the electrolysis process, the cell will reduce heat energy and prevent the leak passes with the help of side ledge. The schematic diagram of cell is shown in Fig. 1.

In the aluminum electrolysis process, powdery alumina is fed into molten electrolyte from point feeders. As the electrolysis progresses, the molten aluminum will accumulate at the bottom of the cell. To ensure the stable operation of the cell, molten aluminum is necessary to tap from cell every day. The amount of the molten aluminum tapped from cell is determined by current cell condition, aluminum liquid height and bath height [17]. According to the mass balance and energy balance, the cell can get a higher current efficiency. The alumina addition is the main factor that causes the change of mass balance. Therefore, keeping alumina concentration in a narrow and reasonable range can reduce the effect coefficient and avoid bottom sediment. In order to get ideal mass balance, the changes of pseudo-resistance slope and accumulated slope are followed by feeding control system of cell to determine alumina concentration. To maintain fluctuation of alumina concentration at the optimum concentration point, underfeed and overfeed are performed alternately [18,19]. Aluminum electrolysis process occurs the primary chemical reaction:

$$\frac{1}{2}\mathrm{Al_2O_3(par)} + \frac{3}{4}\mathrm{C(sol)} \rightarrow \mathrm{Al(liq)} + \frac{3}{4}\mathrm{CO_2(gas)} \tag{1}$$

Interaction among heat balance, mass balance and stability of cell will increase the difficulty of cell management undoubtedly. Technicians often judge aluminum cell condition through the integrated means of appearance, measurement, calculation and touchment, and then determine the alumina feed interval, aluminum fluoride addition, bath level, setting voltage and aluminum tapping volume. It is worth noting that these factors are mutual influence and mutual restriction. In order to ensure a longer period of cells in a stable and efficient state, the technical parameters are required to be coordinated variation of cell management [17,19]. However, such a collaborative model is difficult to establish, which usually requires the combination of empirical knowledge and field knowledge for decision-making. Here, the measured data and visual information are able to reflect the above knowledge. On the investigation and consultation of experienced technologist and consulting a large number of relevant literatures, technicians judge cell condition based on the observed state of flame color, flame intensity, bath color, bath status and bottom sediment. In addition, the judgement also depends on measurable characteristic parameter values of voltage vibration, electrolysis temperature, superheat degree, cell voltage, ledge length, bottom voltage, side ledge thickness and effect coefficient [3,5,20,21]. If judging results are that the cell condition is abnormal, adjustment amount of aluminum tapping volume, setting voltage, aluminum fluoride addition quantity as well as alumina feeding interval will be confirmed according to the domain and experiential knowledge. The characteristics of the parameters and the corresponding effect are shown in Table 1.

Remark: In the Table 1, the main characteristic parameters of cell are analyzed. The causal relationship between the characteristic parameters depends on domain and process knowledge.

Before undertaking the causal analysis, the following assumptions are made.

Assumption 1: During the RCA process, the cell is not out of aluminum, pole, carrying bus and other operations;

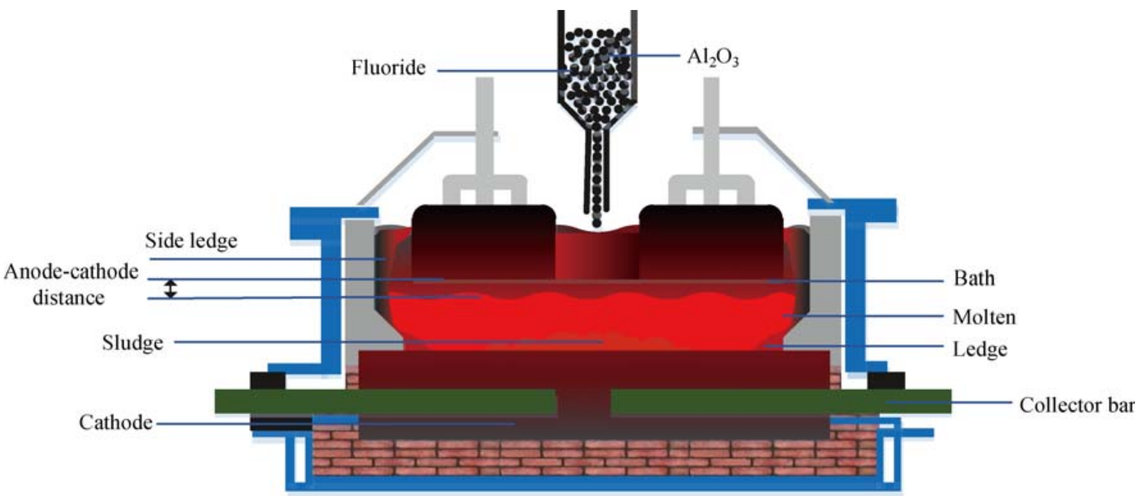

Fig. 1 Schematic diagram of aluminum electrolysis cell

Table 1 The choice of characteristic parameters and characteristics state

Parameter	Effect analysis
Aluminum molten level (*AL*)	◆ The height of aluminum molten. Suitable *AL* is able to maintain energy balance of the electrolytic cell and cell voltage stability.
Cell voltage (*CV*)	◆ The suitable voltage to maintain the normal condition of cell is the source of heat of electrolysis, the most important means to regulate the energy balance of cell.
Molecular ratio (*MR*)	◆ Affecting the solubility of alumina in the electrolyte and the primary crystal temperature is the main means to regulate the energy balance of cell.
Electrolyte level (*EL*)	◆ Maintain the thermal stability and sensitivity of cell, the energy balance is robust with suitable *EL* when energy balance is disturbed, and *EL* affects effect coefficient and alumina solubility.
Feeding interval (*NB*)	◆ The change of *NB* interval has influence on primary crystal temperature, superheat degree, furnace type and bottom sediment.
Voltage vibration (*VV*)	◆ Vibration of voltage is the performance of the abnormal stability of cell.
Bath temperature (*BT*)	◆ Affect the entire operation condition of cell.
Superheat degree (*SD*)	◆ The difference value between the electrolyte temperature and the primary crystal temperature is a comprehensive reflection of the technical indicators for cell.
Flame color (*FC*)	◆ Parameters can reflect the cell condition, and provide visual information for technologist.
Flame intensity (*FI*)	
Bath color (*BC*)	
Bath status (*BS*)	
Ledge length (*LL*)	◆ *LL* has great influence on safe and stable operation of cell. What's more, suitable *LL* is able to prevent leakage of cell and satisfied the need of superheat degree.
Side thickness (*HT*)	◆ *HT* plays the role of insulation and insulation. It has a greater impact on energy balance because of reducing energy loss, and it's the self-balancing bridge of cell.
Bottom sediment (*CS*)	◆ *CS* caused by the lower superheat degree, and it has influence on bottom voltage and stability of cell.
Effect coefficient (*EC*)	◆ Frequency of the anode effect occurs.
Bottom voltage (*BV*)	◆ *BV* can reflect status of bottom crusting and bottom temperature.

and also there is no operation of aluminum tapping, anode changing and lifting up bar.

Assumption 2: During the RCA process, there is no power failure and other similar failures for cell external conditions.

Assumption 3: The cell conditions studied in this paper are in the early and middle of abnormal cell conditions.

On the investigation and consultation of experienced technologist and consulting a large number of relevant literatures, aluminum volume level, setting voltage, aluminum fluoride addition quantity as well as alumina feeding interval are technical parameters (operating parameter) which are able to ensure stable operation of cell. Technical parameters determine the operating cell condition, the adjustment of technical parameters can make the cell to break the existing non-ideal balance state, and establish a new dynamic balance [3,5,20,21]. When the technical parameters are properly adjusted and the abnormal cell conditions are handled timely and accurately, the dynamic balance of cell will move to the ideal direction, and the cell furnace will be regular, thick, solid. What's more, cell conditions are better and indicators increase. Operation of cell will be naturally long-term stable, high efficiency, low consumption and order. Otherwise, the abnormal cell condition continues to deteriorate. For other characteristic parameters (condition parameters), the combination of these parameters can reflect the cell condition, such as the bottom sediment is usually due to low superheat which caused by unreasonable collocation of technical parameters. Combining the process knowledge with experiential knowledge, the causal relationship between operating parameters and condition parameters are shown in Table 2.

3 Proposed multi-source knowledge solidification reasoning model

3.1 FBN for RCA of abnormal cell condition

Probability theory is an effective tool for dealing with uncertainties and stochastic problems. The most common Bayesian Network (BN) is the result of the combination of probability Bayes theory and causality representation graph. Here, BN was also known as knowledge map, causal network and so on [22–24]. Pear proposed BN model in 1986, which was probabilistic knowledge representation and reasoning model. The model was a

Table 2 Analysis of cause and effect among parameters

Operating parameter	The influence of operating parameters each state
NB	◆ *NB* makes a difference to effect coefficient, the *NB* being large, likely to cause low alumina concentration, resulting in an anode effect. ◆ The alumina will not be complete fusion with small *NB*, and the alumina of no melt descend to the bottom of cell, which are easy to be sludge.
AL	◆ More energy will loss from cell bottom with high *AL*, and also the phenomena of temperature decreasing, bottom sludge being more, side ledge being thick will appear ◆ The phenomena of deep immersion in the bath of anode, bath temperature rising, the role of horizontal magnetic field increasing will cause the aluminum liquid in the tank by, prone to voltage swing phenomenon will occur, if aluminum molten level was low.
CV	◆ The phenomena of bath temperature and superheat degree reducing, flame color being blue and white, flame intensity being weak, bath color being red, boiling hard and longer ledge will appear, if *CV* was low. ◆ The phenomena of bath temperature increment, superheat degree and alumina concentration increasing, flame color being yellow, flame intensity being weak, bath color being highlight, fluidity of bath being quick, ledge being smaller, bottom sediment being more and side ledge being thinner will appear, if *CV* was high.
MR	◆ If *MR* was low, alumina solubility will decrease which will be conducive to the precipitation of aluminum from bath with higher surface tension, and the electrolysis temperature will be lower with the chance of secondary oxidation of aluminum reduce, moreover, the amount of sodium precipitation decreased. ◆ If *MR* was high, the phenomena will appear which contain primary crystal temperature increasing, superheat degree being smaller, side ledge being thicker, superheat being higher.
EL	◆ The high *EL* will result in gas discharge difficult, and the chance of anode spikes appearance will increase, what's more, bath boiling will be difficult, resistance increase, the effect coefficient increase. ◆ The low *EL* will result in energy stability being poor and it's sensitive to heat changes, and easy to generate precipitation and produce anode effects.

directed acyclic graph of statistical dependencies between random variables[1)]. Such a graph structure provides a way for the representation and reasoning of uncertain domains. It is suitable for events with uncertainties expression and multiple control factors decision-making. The inference can be made from incomplete or uncertain knowledge with BN.

Nodes of BN represent a set of random variables of the domain model, therefore, the BN could be an abstraction of any problem. The directed edges denote the probabilistic causal dependencies among the variables. The prior knowledge is usually given by data statistics as well as by experts[1)] [24]. Mathematically, BN is represented by $G = \langle\langle V,E\rangle,P\rangle$. Where V denote the set of nodes. Each node has a conditional probability table (CPT), and the CPT is used to denote the influence on child nodes quantitatively. The reasoning principle of BN is based on Bayes probability theory. The reasoning process is a posteriori probability calculation process essentially, which is based on the following the three equations [25].

The joint probability is:

$$P(V) = P(X_1,X_2,X_3,\ldots X_n) = \prod_{i=1}^{n} P\Big(X_i|pa(X_i)\Big), \quad (2)$$

where $pa(X_i)$ is parent node of X_i. Marginal probability of X_i is:

$$P(X_i) = \sum_{\text{except } X_i} P(V). \quad (3)$$

Bayesian network inference is based on node values, observations and other events according to Bayesian rules for calculation. It can perform knowledge reasoning when given some evidence by following equation:

$$P(V_i|e) = \frac{p(V,e)}{p(e)} = \frac{p(e,V)p(V)}{\Sigma_{i=1}^{n} p(e,V_i)p(V_i)}, \quad (4)$$

where e denote evidence, it can be value or observation and so on. The right side of the equation is the known prior probability, and the left side is the posterior probability.

In the aluminum electrolysis process, technologist usually uses the imprecise and uncertain fuzzy event to describe the variable state. Such as *CV*, technologist makes judgment for *CV* state high or low according to the combinationvalue of nodes with domain knowledge. The traditional BN is invalid when the judging results of technologist are fuzzy. So it is necessary to introduce fuzzy set theory into the Bayesian network, and the state of the node variables is described by fuzzy events. The traditional nodes of BN are changed to the fuzzy nodes by FBN, which is better able to deal with the fuzzification of nodes. For example, the states of root nodes *NB* are described by three fuzzy events with short, normal and long in this paper. The state of each node in the FBN represents the characteristic state of variable in the cell. The Bayesian reasoning method provides a way to obtain a posterior probability from prior probability. In the Eqs. (2–4), V_i represents the root cause of abnormal cell condition and e represents characteristic state of abnormal cell condition. $P(e|V)$ and $P(V)$ are given by statistic of historical data and expert knowledge. The posterior probability $P(V_i|e)$ is

1) Pan H, McMichael D. Fuzzy causal probabilistic networks — a new ideal and practical inference engine. Proceedings of 1st International Conference on Multisource-Multisensor Information Fusion, 1998, 6–8

calculated by the Eq. (4) and the abnormal state of corresponding node is the root cause of abnormal cell conditions when the posterior probability of the abnormal state of corresponding node is maximum. As the RCA is based on the characteristics of the system to identify the cause of problem, FBN can also find the state with maximum probability which is calculated based on the Bayesian rule according to the child node states. Therefore, FBN can be used in the RCA process. A graphical framework is built for RCA structure based on FBN according to above idea as shown in Fig. 2.

Remark: Based on the Bayesian network framework, the fusion of domain knowledge and expert knowledge can simulate the decision-making process effectively, which is more secure to enhance the accuracy of RCA. The measured data and visual information fuse effectively based on the above Bayesian network framework. The framework is able to simulate the human decision-making process with the fusion of the domain knowledge and expert knowledge. The accuracy of RCA improving is more secure under the characteristic of the framework [9].

Layer 1 contains root nodes. Layer 2 contains child nodes, and Layer 3 is same to Layer 2 in the Fig. 2. For RCA of abnormal cell condition, Layer 1 contains the root cause of abnormal cell condition, i.e., the states high, low of *AL*, *MR*, *CV* and *EL*, the states long, short of *NB*. Layer 2 contains measurable variables, i.e., the variables *VV*, *BT*, *LL*, *ST*, *EC*, *BV*. Layer 3 contains observable variables, i.e., the variables *FC*, *FI*, *BC*, *BS*, *CS*. The variables in Layer 2 and Layer 3 depend on the accurate measurement of the sensing device and the visual judgment of the experienced technicians, respectively. Layer 3 has the assistant judgment function to Layer 2, and the technicians can combine these condition parameters with their own experiential knowledge to make analysis and judgment to the cause of abnormal cell condition.

3.2 Multi-source knowledge solidification reasoning model based on FBN

MSKR model is a kind of graph model that contains experiential knowledge of knowledge workers and production of the necessary process knowledge based on FBN. Due to the introduction of fuzzy concept, it will be more reasonable to express the multi-source knowledge of aluminum electrolysis with fuzzy, and explain the results of network reasoning more persuasively. This kind of graph model not only can carry on the qualitative expression to the domain knowledge, but also is able to inference quantitatively after getting the prior knowledge [11,12,23]. How to integrate the process knowledge, prior knowledge as well as expert experiential knowledge into the framework and finally to form a multi-source knowledge curing model, which will be analyzed in this chapter. The modeling process consists of the following steps are shown in Fig. 3:

(1) combining with the aluminum electrolysis knowledge, the characteristic variables are selected that which are able to reflect the condition of cell;

(2) the causal relationship graph reflects the relationships between variables that are constructed based on the knowledge analysis of aluminum electrolysis process;

(3) the fuzzification of each node in the causal graph is necessary to meet the process conditions, and the fuzzification process should combine with the control knowledge of cell;

(4) the prior probability of each node in the network is given based on consultation of experienced experts of

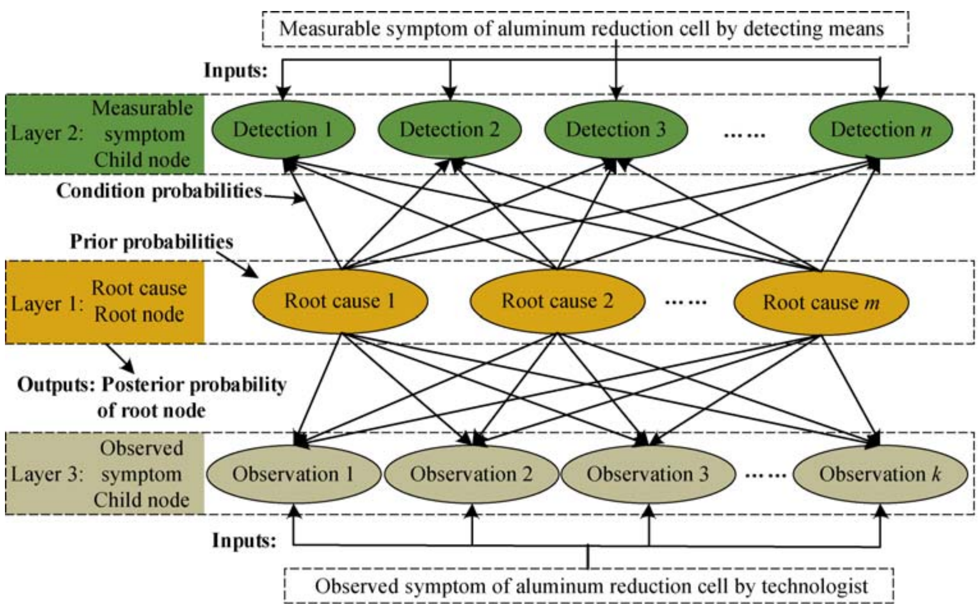

Fig. 2 RCA of ASARC based on FBN

aluminum electrolysis and statistic of historical data;

(5) based on the above steps, a multi-source knowledge solidification reasoning model is built. This model integrates the process knowledge and experiential knowledge of aluminum electrolysis, and the model is able to recognize the cause which leads to abnormal cell condition;

(6) testing and analyzing the actual abnormal cell condition, the reasoning result based on MSKR model is compared with the result of experienced technologist.

Based on the analysis of steps of constructing MSKR model, range of values of various states of the characteristic parameters are determined by the experiential knowledge of aluminum electrolysis experts. In the first step, the experiential knowledge of aluminum electrolysis process is integrated into the model. Next, the dependence among the nodes is determined by the process knowledge, and the process knowledge is integrated in the step. On the basis of the network architecture, the prior knowledge is given to each node by combining with the experience of aluminum electrolysis experts and the statistic of history data.

3.2.1 A framework of multi-source knowledge solidification reasoning model

The operating parameters *NB*, *AL*, *CV*, *MR*, *EL* are considered as root nodes. *VV*, *BT*, *LL*, *EC*, *BV*, *FC*, *FI*, *BC*, *BS*, *CS* are considered as child nodes of MSKR model based on the analysis in the Section 3.1. The causal relationship between the parameters is analyzed in Section 2. Combining with the results of the above analysis, a framework of causal relationship between the characteristic parameters is constructed based on FBN, as shown in Fig. 4.

The dominance of the causal relationship between the variables make the expression and reasoning of uncertainty more intuitive and clear in the causal diagram. For example, the state of *AL* has influence on *VV*. If state of *AL* was low, the phenomena will appear that the heat dissipation area approaches the bottom of cell, the bottom of the furnace temperature increasing, the furnace of cell being larger, expanding for aluminum liquid mirror. According to the above phenomena, the horizontal current density in the aluminum liquid will heighten. The fluidity of the aluminum liquid in the cell will accelerate under the strong electromagnetic force with current density increasing, and also may result in increasing of non-uniform electromagnetic force. Meanwhile, the liquid aluminum fluctuations increasing may result in increasing of *VV*. In addition, the state of *AL* has influence on *SD*. If the state of *AL* was high, the cell temperature caused by increasing of the heat dissipation from cell bottom, and it may result in being lower *SD*. Moreover, the state of *BT* and *SD* all have

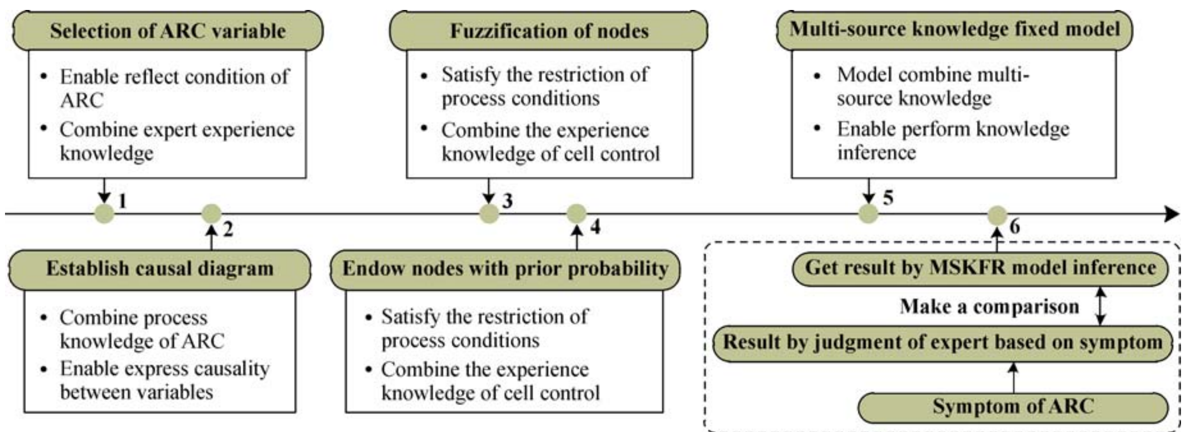

Fig. 3 Knowledge solidification modeling steps based on Fuzzy-Bayesian network

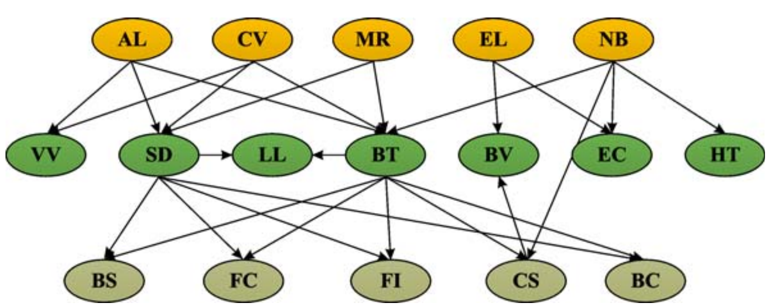

Fig. 4 Causal relationship among characteristic parameters

influenced on the states of BS, BC, FC, FI, CS and LL. The state of SD being high, BT being normal may result in boiling fiercely, being highlight, yellow, weak, normal, short of BS, BC, FC, FI, CS and LL, respectively.

Based on the causal relationship among the characteristic parameters and the constraints of aluminum electrolysis process condition, the nodes in the causal graph are fuzzification according to the knowledge of the aluminum electrolysis experts and the statistic of the measured data. Each node in the causal graph can be divided into several states, and these states cover the condition of cell during the aluminum electrolysis process. For the sake of convenience in writing, the state of a node is reduced to a letter instead. For example, AL_L and AL_H represent the state low and high of AL, respectively. Other nodes are the same to AL. First, the root nodes (operation parameters) are fuzzification, i.e., for example, the states of AL are always divided into three levels, i.e., low (18–20) cm, normal (19–21) cm, high (20–22) cm. The same to child nodes, the variable universe of SD is [5–12] °C. When it is varied from 5 to 8 °C, the state is low/SD_L, and varied from 7 to 10 °C, being normal/SD_N, moreover, varied from 9 to 12 °C, divided into high/SD_H. Combined with triangular fuzzy numbers, the three fuzzy states of aluminum level are shown in Fig. 5. The remaining fuzzy states of the feature parameters of Tables S1 and S2 are shown in the Supplementary Material A.

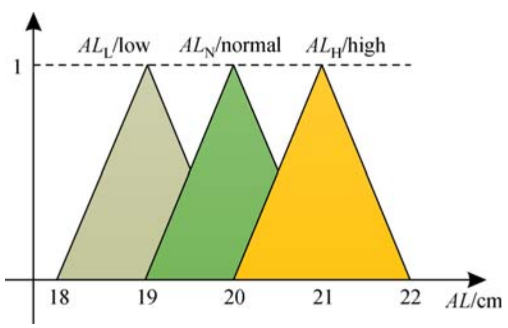

Fig. 5 The fuzzy state division of AL

The fuzzification of some characteristic parameters can't be measured by the detection device, which only depend on the observable information of the technicians. The fuzzification of the states of such nodes is usually based on technicians' experience. For example, the flame color is divided into three types by technicians. When the temperature of cell is cold, normal, high, the state of flame color are bluewhite, laverder and yellow, respectively, and other nodes that only depend on observing are the same to FC, and detailed analysis is as shown in Table 3.

3.2.2 Estimation of prior probability and conditional probability

The selection of the characteristic variables and the causal relationship among the variables are analyzed in Section 2, and MSKR model framework is constructed in Section 3.2.1. However, a complete MSKR model should obtain the probability of nodes according to the topology structure. The probability information contains marginal and conditional probability. They are stored in the probability table of FBN node, and the probability tables together with topology structure of the causal relationship graph are used to describe MSKR model. In order to maximize the accuracy of prior knowledge, the experienced aluminum electrolysis experts are consulted. Combing with fuzzy theory, the prior probability of root nodes and condition probability are determined according to the knowledge collection from experienced aluminum electrolysis experts and statistic of history data.

In recent years, the fuzzy mathematics method is widely used in the research of uncertainty for complex system, which is accurate and reliable with fuzzy and stochastic description for the system information [26,27]. Triangular fuzzy number is a method, which makes fuzzy language events to crisp value. The triangular fuzzy number is used to express the experiential knowledge, and the experiential knowledge is used to estimate probability of nodes under the abnormal cell condition in this paper, i.e., the fuzzy membership of triangular fuzzy numbers is introduced to represent the prior and condition probabilities of the nodes [28], and the triangular fuzzy numbers are defined as formula (5) [29].

$$\mu_X = \begin{cases} (x-l)/(m-l), & l \leqslant x \leqslant m, \\ (x-u)/(m-u), & m \leqslant x \leqslant u, \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

where l, m, u ($0 \leqslant l \leqslant m \leqslant u \leqslant 1$) represent the lower least likely value, the most likely valve, and the upper least

Table 3 State of observable nodes

Child node name	State of child node name		
FC	Blue white (FC_B)	Lavender (FC_L)	Yellow (FC_Y)
FI	Weak (FI_W)	Normal (FI_N)	
BC	Red (BC_R)	Red yellow (BC_{RY})	Highlight (BC_H)
BS	Hard (BS_H)	Equably (BS_E)	Fiercely (BS_F)
CS	Normal (CS_N)	Seriously (CS_S)	Very serious (CS_V)

likely value, respectively. So (l, m, n) represents triangular fuzzy numbers $X_i = (l, m, n)$ (i represents ith node). In order to get exact probability in the inference process, the triangular fuzzy numbers need to be defuzzified to get the crisp value. The center gravity of method is adopted, shown as formula (6) [30].

$$P_i = \frac{(u_i - l_i) + (m_i - l_i)}{3} + l_i, \tag{6}$$

where P_i represents the probability with defuzzification of triangular fuzzy number.

The probability of each state of AL is shown in the Table 4. When the state of AL is normal, (0.87, 0.88, 0.89) represents the probability of AL_N, and the most likely value of AL_N is 0.88, while 0.87 and 0.89 are the lower and upper least likely values, respectively. The probability values, such as (0.87,0.88,0.89) are mainly achieved from the experienced experts of aluminum electrolysis. The prior probability calculation procedure of the state of AL is as follows:

$$P(AL_N) = \frac{(0.89 - 0.87) + (0.88 - 0.87)}{3} + 0.87 = 0.88. \tag{7}$$

The prior probabilities of other root nodes are shown in Table 4 and Table 5.

With the help of the experienced aluminum electrolysis experts, the prior probabilities are given to the root nodes in Table 4 with triangular fuzzy numbers. With the same method, the conditional probabilities are given to the measurable child nodes, for example, the condition probabilities between VV and AL, CV are determined in Table 5. Table S1 shows that the CV and AL are all divided into three states: low, normal, and high. As can be seen from Table S2, there are three states of VV: normal/VV_N, serious/VV_S, and very serious/VV_V. The condition probabilities of VV are obtained according to the above method. The process of solving the conditional probabilities are omitted with triangular fuzzy numbers, and only crisp values are given. The condition probabilities are shown in Table 5.

The information of child nodes is obtained by the visual observation, and the judgment of the state of the node is based on the experiential knowledge of experts. Such as BC is influenced by BT and SD, which has three states: red/BC_R, red yellow/BC_{RY} and highlight/BC_H according to Figure 4 and Table S2. BT and SD both have three states: low, normal and high according to Table S2. The conditional probabilities of BC are also given in Table

Table 4 Prior probabilities of fuzzy states for root nodes

Root node name	Prior probability of each state of root node		
AL	AL_L	AL_N	AL_H
	(0.04,0.05,0.06)	(0.87,0.88,0.89)	(0.06,0.07,0.08)
	0.05	0.88	0.07
CV	CV_L	CV_N	CV_H
	(0.08,0.09,0.1)	(0.84,0.85,0.86)	(0.05,0.06,0.07)
	0.09	0.85	0.06
MR	MR_L	MR_N	MR_H
	(0.09,0.1,0.11)	(0.81,0.82,0.83)	(0.07,0.07,0.09)
	0.1	0.82	0.08
EL	EL_L	EL_N	EL_H
	(0.02,0.03,0.04)	(0.91,0.92,0.93)	(0.04,0.05,0.06)
	0.03	0.92	0.05
NB	NB_S	NB_N	NB_L
	(0.05,0.06,0.07)	(0.89,0.9,0.91)	(0.03,0.04,0.05)
	0.06	0.9	0.04

Table 5 Condition probability of variable 'voltage vibration (VV)'

Variable states	The prior probability of each state of child nodes								
	AL_L			AL_N			AL_H		
	CV_L	CV_N	CV_H	CV_L	CV_N	CV_H	CV_L	CV_N	CV_H
VV_N	0	0	0.05	0.15	0.96	0.85	0.11	0.16	0.7
VV_S	0.05	0.25	0.4	0.66	0.04	0.15	0.18	0.3	0.24
VV_V	0.95	0.75	0.55	0.19	0	0	0.71	0.54	0.06

S3 of the Supplementary Material B according to the method described above. The conditional probabilities of the remaining nodes, see the Supplementary Material B.

As we can see from the analysis of the above chapters, it should select characteristic variables first. Five characteristic variables are selected as root nodes which are able to adjust the equilibrium point of cell. In addition, twelve characteristic variables are selected as child nodes which are able to reflect the cell condition according to process knowledge and experiential knowledge of experts. The casual relationships among characteristic variables are determined according to process knowledge, and the MSKR framework is determined as shown in Fig. 4. The experiential knowledge of technicians is explicit with the introducing of fuzzy theory, result in dividing several states of variables. Besides, with the consulting of authoritative experts, the prior probabilities of root nodes and the conditional probabilities among the nodes are obtained according to experiential knowledge and statistic from history data, combing with triangular fuzzy number. According to the selection of characteristic variables, fuzzification of variables and the obtainment of prior knowledge, experiential knowledge and data knowledge are solidified in the model. What's more, process knowledge is assimilate into the model based on the modeling of casual relationship framework. At last, combining with Bayesian theory and by means of Netica software, the building of MSKR model is complete. Netica software was a comprehensive tool for working with Bayesian belief nets and decision nets (influence diagrams) which was designed to be simple, reliable, and high performing. For managing uncertainty in business, engineering, medicine, or ecology, it was able to build, learn, modify, transform and store nets, as well as answer queries or find optimal solutions using its powerful inference engine. The MSKR model based on FBN can be obtained with the help of software Netica that is shown as Fig. 6. The details of inference process in Fuzzy-Bayesian network with Netica shown as Algorithm 1.

Algorithm 1 The RCA of abnormal cell condition based on fuzzy-Bayesian network reasoning

```
// Initialization
// Input:
  BN: a fuzzy-Bayesian network
  E_i: the evidence of ith node of abnormal cell condition
  s_ij: the jth state of ith node
  N: the number of nodes
  M: the number of root nodes
  n: the group of abnormal cell conditions
  node_sizes: the set of state sizes of each node
  for groupth = 1: n // groupth is the number of conditions
    for i = 1: N
      s_i = TFN (E_i); // the states s_i of E_i divided by experienced experts
      with Triangular fuzzy number
      P(s_ij) = Defuzzification(s_ij); // according to experienced knowledge
      and data knowledge
    end
    for i = 1: N
      dag(E_i, E_{i+k}) = true; // the directed line of the the ith node to (i + k)
      th node is added based on the process knowledge
    end
    node_sizes = [ ]; the set of numbers of states of each node
    Bnet = mk_bnet (dag, node_sizes);
    bnet.CPD{E_i} = tabular_CPD (bnet, E_i, P(s_i)); // P(s_i) the set of prior
    probability of ith node's states
    Engine = jtree_inf_engine (bnet); // joint tree inference engine
    [engine,[]] = enter_evidence (engine, E); // adding evidence to model
    for Mth = 1: M
      Marg = marginal_nodes (engine, s_Mth); // Marginal probability
      calculation of Mthth root nodes
      Marg. T;
      P(Mth) = max(Marg. T(1), Marg. T(3)); // the maximum posterior
      probability of root node's
      p(s_Mthj | E) = max(P(Mth), P(Mth-1)); // the jth state of Mthth root
      node is the root cause for groupthth cell condition
    end
  end
// Output: Root cause
    The estimation of maximum posterior probability of each group
    condition.
```

4 Verification and validation

In this section, consulted experts are authoritative and experienced in the field of aluminum electrolysis. Make an assumption before validation and analysis of MSKR model that the accuracy rate of analysis and judgment for abnormal cell condition is able to be 100%.

The MSKR model is analyzed and constructed based on FBN in the above section, and the validity of MSKR model is validated by Netica software in the RCA process of abnormal cell condition. Verification process is knowledge reasoning process, which is realized by diagnosis inference. Diagnosis inference is bottom-up reasoning that inferred from conclusion to reason. Diagnosis inference is performed according to the Bayesian rule formulas (1), (2) and (3), and the posterior probability of the cause of the result is obtained. However, the cause of the abnormal condition for a variety of reasons, the root with the maximum posteriori probability can be considered as the root cause. In RCA process of abnormal cell condition, the reasoning computation is performed by taking abnormal cell condition as the known evidence into the MSKR model, and find the abnormal state of AL, CV, MR, NB, EL with maximum posterior probability. And RN_{ij}^{k} is the state

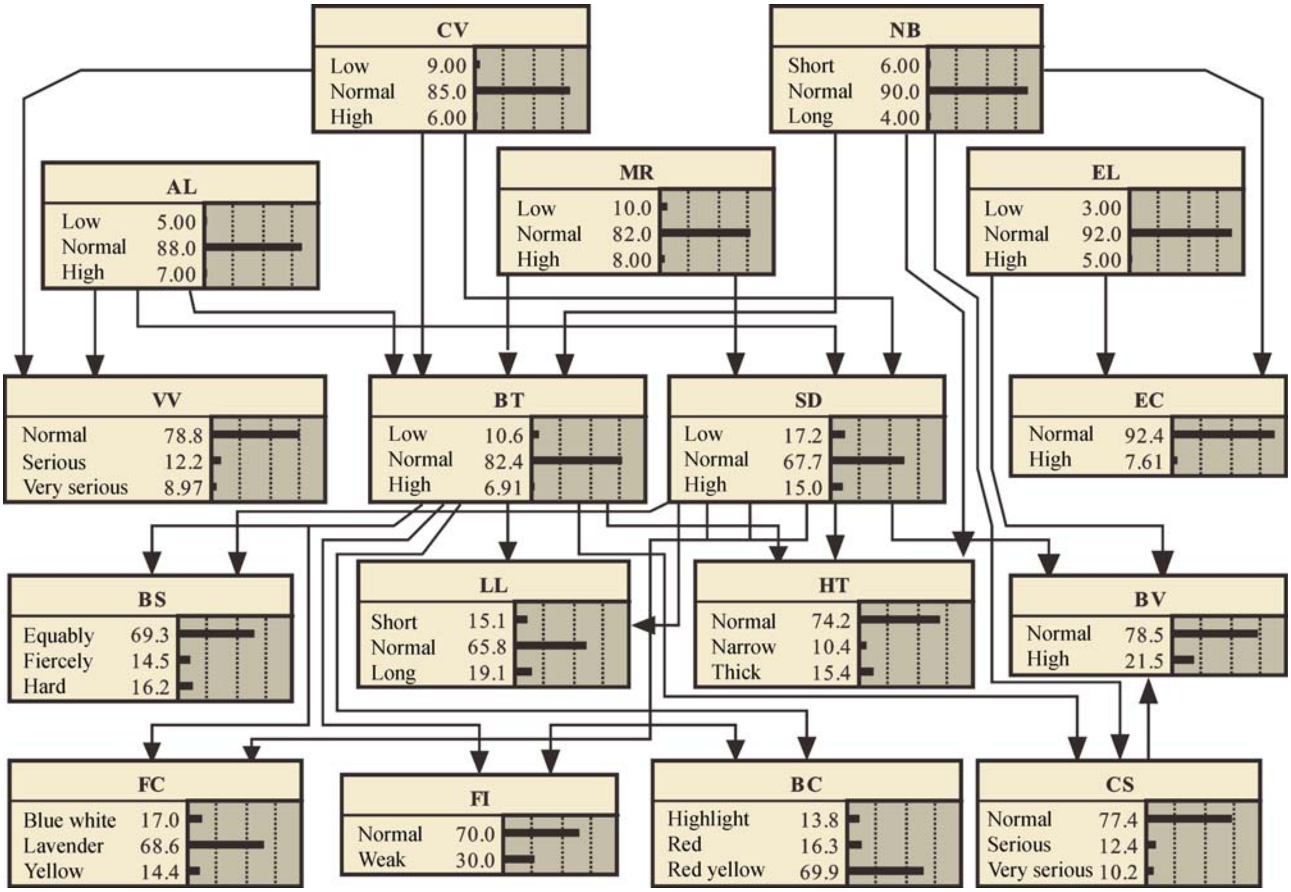

Fig. 6 MSKR model based on Fuzzy-Bayesian network

of root nodes corresponding to the status of the analysis in Table S1. The purpose of the reasoning calculation is to find the abnormal state of the root node with a maximum posteriori probability, which is the result of the reasoning calculation of the MSKR model based on FBN, i.e., root cause of abnormal cell condition:

$$RN^k_{ij} = \begin{pmatrix} AL^k_{11}/AL^K_{\mathrm{L}} & AL^k_{12}/AL^K_{\mathrm{N}} & AL^k_{13}/AL^K_{\mathrm{H}} \\ CV^k_{21}/CV^K_{\mathrm{L}} & CV^k_{22}/CV^K_{\mathrm{N}} & CV^k_{23}/CV^K_{\mathrm{H}} \\ MR^k_{31}/MR^k_{\mathrm{L}} & MR^k_{32}/MR^k_{\mathrm{N}} & MR^k_{33}/MR^k_{\mathrm{H}} \\ NB^k_{41}/NB^k_{\mathrm{S}} & NB^k_{42}/NB^k_{\mathrm{N}} & NB^k_{43}/NB^k_{\mathrm{L}} \\ EL^k_{51}/EL^k_{\mathrm{L}} & EL^k_{52}/EL^k_{\mathrm{N}} & EL^k_{53}/EL^k_{\mathrm{H}} \end{pmatrix}, \quad (8)$$

where i denotes ith node, j denotes jth state of the ith node, and k denotes the kth group abnormal cell condition.

In order to verify the validity of the proposed method, the method that comparing reasoning result based on the MSKR model with the results given by experts (*RGE*) of aluminum electrolysis is adopt, i.e., when $RN^k_{mn} = \max\ p(RN^k_{ij})$, the value of Δ is shown as Eq. (9).

$$\Delta = \begin{cases} 1 & RN^k_{mn} = RGE^k \\ 0 & \text{otherwise} \end{cases}, \quad (9)$$

where $m = 1,2,3,4,5$ and $n = 1,3$. And RN^k_{mn} denote the result based on MSKR model, and RGE^k is the result given by experts for the kth group abnormal cell condition.

The abnormal cell conditions for verification are all from the actual production site. The posterior probability of each abnormal state of the root node of MSKR model is obtained based on analyzing and reasoning of the 20 groups of abnormal cell conditions. The abnormal state with maximum posterior probability is compared with the expected result given by the experts, as shown in Table 7 validation results.

The 20 groups of abnormal cell conditions used for analysis and verification are shown in Table 6.

A series of abnormal cell conditions are reflected by the child nodes of the MSKR model which are listed in Table 6. For example, seventh group cell condition is that the state of *VV*, *BT*, *SD*, *EC*, *BS*, *LL*, *HT*, *BV*, *FC*, *FI*, *BC* and *CS is* VV_{N} /normal, BT_{H} /high, SD_{H} /high, EC_{N} /normal, BS_{F} /fierce, LL_{S} /short, HT_{Na} /narrow, BV_{N} /normal, FC_{Y} / yellow, FI_{W} /weak, BC_{H} /highlight, CS_{N} /normal, respectively. Verification is performed based on the proposed method with abnormal cell conditions, as shown in Table 6, validation results as shown in Table 7.

Remark: When the maximum posterior probability is selected in Table 7, only the abnormal state of each node is

Table 6 The abnormal aluminum electrolysis cell condition used to validate

Group	Corresponding state of the variables											
	VV	*BT*	*BT*	*EC*	*BS*	*LL*	*HT*	*BV*	*FC*	*FI*	*BC*	*CS*
1	VV_V	BT_N	SD_N	EC_N	BS_E	LL_N	HT_N	BV_N	FC_L	FI_N	BC_{RY}	CS_N
2	VV_N	BT_N	SD_N	EC_H	BS_E	LL_N	HT_{Na}	BV_N	FC_L	FI_N	BC_{RY}	CS_N
3	VV_N	BT_N	SD_N	EC_H	BS_E	LL_N	HT_N	BV_H	FC_L	FI_N	BC_{RY}	CS_V
4	VV_N	BT_L	SD_L	EC_N	BS_H	LL_N	HT_N	BV_H	FC_B	FI_W	BC_R	CS_S
5	VV_N	BT_N	SD_L	EC_N	BS_H	LL_N	HT_N	BV_H	FC_B	FI_W	BC_R	CS_S
6	VV_V	BT_L	SD_N	EC_N	BS_E	LL_N	HT_N	BV_N	FC_L	FI_N	BC_{RY}	CS_N
7	VV_N	BT_H	SD_H	EC_N	BS_F	LL_S	HT_{Na}	BV_N	FC_Y	FI_W	BC_H	CS_N
8	VV_N	BT_N	SD_H	EC_N	BS_F	LL_S	HT_{Na}	BV_S	FC_Y	FI_W	BC_H	CS_N
9	VV_V	BT_N	SD_L	EC_H	BS_E	LL_L	HT_T	BV_N	FC_L	FI_N	BC_{RY}	CS_N
10	VV_N	BT_N	SD_N	EC_H	BS_E	LL_L	HT_T	BV_H	FC_L	FI_N	BC_{RY}	CS_N
11	VV_N	BT_L	SD_H	EC_N	BS_F	LL_S	HT_{Na}	BV_N	FC_B	FI_W	BC_H	CS_N
12	VV_N	BT_N	SD_L	EC_H	BS_H	LL_L	HT_T	BV_H	FC_B	FI_W	BC_R	CS_S
13	VV_V	BT_L	SD_N	EC_H	BS_H	LL_N	HT_N	BV_N	FC_B	FI_W	BC_R	CS_N
14	VV_N	BT_N	SD_H	EC_N	BS_E	LL_N	HT_N	BV_N	FC_Y	FI_W	BC_{RY}	CS_S
15	VV_N	BT_N	SD_L	EC_N	BS_H	LL_N	HT_N	BV_H	FC_L	FI_N	BC_R	CS_V
16	VV_S	BT_L	SD_L	EC_N	BS_E	LL_N	HT_N	BV_H	FC_B	FI_W	BC_{RY}	CS_V
17	VV_N	BT_L	SD_N	EC_N	BS_E	LL_L	HT_T	BV_H	FC_L	FI_N	BC_{RY}	CS_V
18	VV_N	BT_H	SD_H	EC_N	BS_F	LL_N	HT_N	BV_N	FC_Y	FI_W	BC_H	CS_N
19	VV_S	BT_N	SD_L	EC_N	BS_E	LL_L	HT_T	BV_N	FC_B	FI_N	BC_R	CS_N
20	VV_V	BT_N	SD_H	EC_N	BS_F	LL_S	HT_{Na}	BV_N	FC_Y	FI_W	BC_H	CS_N

considered, for example, the states low, high of *CV*, the states short, long of *NB*, regardless of the node's state normal. The values with bold and italics in the table indicate the maximum posteriori probability of the root node's abnormal state, which corresponding to the group cell condition. That is, the abnormal state with the maximum posteriori probability is the root cause of the abnormal condition.

According to the proposed MSKR model based on FBN, 20 groups of abnormal cell conditions are verified and analyzed, and 19 groups of calculation results are the same to the results given by aluminum electrolysis experts, i.e., $\Delta = 1$ denotes that the decision results is correct based on MSKR model. The results of statistical analysis for Table 7 are shown in Fig. 7, easy to know the accuracy of test results to 95%.

While the test result is $\Delta = 0$ of 13th group cell condition, reasoning result is different to the given result by experts. Reasoning result shows that the state of NB is NB_L/Long, however, the result given by experts is CV_L/Low. Result of the 13th group shows that the states of *VV*, *BT*, *SD*, *EC*, *BS*, *LL*, *HT*, *BV*, *FC*, *FI*, *BC*, and *CS* are VV_V/very serious, BL_T/low, SD_N/normal, EC_H/high, BS_H/hard, LL_N/normal, HT_N/normal, BV_N/normal, FC_B/bluewhite, FI_W/weak, BC_R/red and CS_N/normal, respectively. The phenomena of the 13th group aluminum cell condition show that the condition of cell is the initial stages of cold trip. The reasoning result indicates that the root cause of the 13th group aluminum cell condition is NB_L/long.

Nevertheless, the root cause of the thirteen group aluminum cell condition is insufficient heat in the actual production site. Analysis to the computation result based on MSKR model, if NB was NB_L/long, the supply of alumina will be insufficient, and result in concentration reducing of $[AlOF_3]^{-2}$ in the bath. When reduced to a certain extent, interaction produce carbon fluoride between the precipitation of fluoride and anode, the anode surface will be passivated. The electrolyte will not wet the anode well, and form poor conductive gas film, resulting in increasing of *CV* and happening of anode effect. At last the electrolyte temperature increase significantly because of above phenomena, rather than the electrolyte temperature is low in the 13th group condition. Therefore, the root cause of the abnormal cell condition in the 13th group is CV_L/low rather than NB_L/long.

The reason why the result based on MSKR model is inconsistent with the result given by expert is analyzed as follows: enter the following cell condition in the Netica software.

VV_v BT_L SD_N BS_H LL_N HT_N BV_N FC_B FI_W BC_R CS_N

Table 7 Validation results

Variables	state	Group number 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
AL	Low	***0.5075***	0	0	0	0	0.252	0.0038	0.0001	0.0273	0	0	0	0.254	0.0002	0	0.0028	0	0.0038	0.0059	***0.986***
	High	0.48	0.003	0.004	0.112	0.14	0.407	0.0072	0.0021	***0.887***	0.003	0.003	0.053	0.395	0.0018	0.139	0.0662	0.054	0.0072	***0.591***	0.0136
CV	Low	0.0319	0.0003	0.0002	***0.685***	0.0334	***0.797***	0.0001	0.0005	0.1138	0.0007	0.0112	0	0.819	0.0001	0.0295	***0.949***	0.163	0	0.329	0.0018
	High	0.0031	0.0012	0.0018	0.001	0.0016	0.001	***0.998***	0.0785	0.0002	0.0013	0.0008	0.13	0.0002	0.0789	0.0015	0.0003	0.004	***1***	0.001	0.0062
MR	Low	0.198	0.0117	0.0124	0.0103	0.002	0.454	0.108	***0.935***	0.0116	0.0112	***1***	0.0066	0.4503	***0.935***	0.0019	0.013	0.258	0.1075	0.0086	0.1133
	High	0.129	0.0163	0.0176	0.3017	***0.847***	0.004	0.01	0.0008	0.1194	0.0168	0	***0.948***	0.0047	0.0001	0.851	0.087	0.093	0.0105	0.1914	0.0207
NB	Short	0.0003	0.0005	***1***	0	0.0222	0	0.0205	0.0022	0.006	0.017	0	0.624	0.0006	0.0255	***1***	0.0655	***0.632***	0.0453	0.0051	0.002
	Long	0.0007	***0.977***	0	0.006	0.0018	0.035	0.0055	0.0318	0.605	0.283	0.085	0	***0.901***	0.0015	0	0.0035	0.001	0.0007	0.0069	0.032
EL	Low	0.0226	0.0301	0.036	0.0272	0.0274	0.0218	0.0238	0.0231	0.0646	***1***	0.0219	0.0696	0.0358	0.0236	0.0294	0.0266	0.0283	0.024	0.0255	0.0231
	High	0.0464	0.0549	0.0441	0.0448	0.0456	0.0462	0.0462	0.0459	0.0914	0	0.0451	0.0814	0.0622	0.0414	0.0506	0.0464	0.0477	0.047	0.0475	0.0459
Reasoning result		*AL*	*NB*	*NB*	*CV*	*MR*	*CV*	*CV*	*MR*	*AL*	*EL*	*MR*	*MR*	*NB*	*MR*	*NB*	*CV*	*MR*	*CV*	*AL*	*AL*
		Low	Long	Short	Low	High	Low	High	Low	High	Low	Low	High	Long	Low	Short	Low	Low	High	High	Low
Result given by experts		*AL*	*NB*	*NB*	*CV*	*MR*	*CV*	*CV*	*MR*	*AL*	*EL*	*MR*	*MR*	*CV*	*MR*	*NB*	*CV*	*MR*	*CV*	*AL*	*AL*
		Low	Long	Short	Low	High	Low	High	Low	High	Low	Low	High	Low	Low	Short	Low	Low	High	High	Low
Δ		1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1

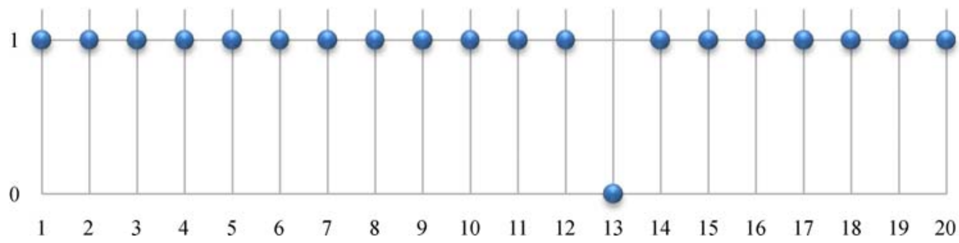

Fig. 7 The statistical results of verification

Now the status of the effect coefficient is unknown, and the sort of root cause for the above cell condition is as follows: $p(CV_{\mathrm{L}})>p(MR_{\mathrm{L}})>p(AL_{\mathrm{H}})>p(NB_{\mathrm{L}})>p(EL_{\mathrm{H}})$.

As a result, it reveals that $CV = CV_{\mathrm{L}}$ is the root cause of the above cell condition, at this point the results by reasoning is consistent with the result given by experts. When adding $EC = EC_{\mathrm{H}}$ to above cell condition, it will be the 13th group cell condition, and the sort of root cause for the 13th group cell condition is as follows: $p(NB_{\mathrm{L}}) > p(CV_{\mathrm{L}}) > p(MR_{\mathrm{L}}) > p(AL_{\mathrm{H}}) > p(EL_{\mathrm{H}})$.

The sort result is inconsistent to the result given by experts. So the main reason for inconsistency of the two results is that the state of *EC* has an influence on the inference result, and there is little deviation for the prior knowledge of *EC*.

The multi-source knowledge used for RCA of abnormal cell condition is solidified to build the MSKR model based on the proposed method. Aluminum electrolysis technicians are able to make judgment decision for the abnormal cell condition with the help of MSKR model. It's able to solve the problem that it is difficult to ensure the accuracy of finish the complex work because of dwindling and frequent flow of experienced technicians. Besides that, the problem that some knowledge is explicit with difficulty is resolved too.

5 Conclusions

The limitations of the existing detection device and the serious coupling characteristics of the electrolysis process, and with multi-source occurrence of abnormal cell condition result in that analysis for RCA of abnormal cell condition is complex. At the same time, the RCA of abnormal cell condition mainly relies on experienced technicians. However, it is difficult to ensure the accuracy of the RCA for abnormal cell condition because of dwindling and frequent flow of experienced technicians in the aluminum electrolysis industry. Process knowledge and experiential knowledge play an important role in RCA of abnormal cell condition, and the existing RCA of abnormal cell condition methods are very dependent on data analysis. These methods don't make full use of these knowledge, resulting in a great limitation for RCA of abnormal cell condition based on data analysis. The method proposed in this paper is able to solidify the various information, process knowledge and experiential knowledge effectively. Meanwhile, as the causes of abnormal cell conditions are multi-source, coupling, uncertainty and so on, the fuzzy Bayesian network proposed in this paper provides an effective and intuitionistic solution to solve the problem. Moreover, the introducing of FBN proposed in MSKR model realize solidification, fusion and inference calculation of multi-source knowledge, which is helpful for the technicians to analyze abnormal cell condition. The proposed method is validated in 20 groups of abnormal cell conditions, and the accuracy is up to 95%, which shows that the MSKR model based on FBN is feasible for RCA of abnormal cell condition.

Acknowledgements The paper is supported by the National Natural Science Foundation of China (Grant Nos. 61533020 and 61374156). The authors would like to thank Prof. Zou Zhong, Central South University, and the experts of Zhengzhou Faxiang Aluminum Industry Co., LTD for providing data and information of aluminum electrolysis cell.

Electronic Supplementary Material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s11705-017-1663-x and is accessible for authorized users.

References

1. Qiu Z X. Principles and Applications of Aluminum Electrolysis. Xuzhou: China University of Mining Press, 1998, 508–510
2. Gui W H, Wang C, Xie Y F, Song S, Meng Q F. Knowledge automation is a necessary method for universities to realize spanning development for process industry. China science founding, 2015(5): 337–342
3. Guo J, Gui W H, Wen X F. Multi-objective optimization for aluminum electrolysis production process. Journal of Central South University, 2012, 43(2): 548–553
4. Stam M A, Taylor M P, Chen J J J, Mulder A, Rodrigo R. Common behaviour and abnormalities in aluminum reduction cells. TMS Light Metals, 2008, 309–314
5. Majid N A A, Taylor M P, Chen J J J, Stam M A, Mulder A, Young B R. Aluminum process fault detection by multiway principal component analysis. Control Engineering Practice, 2011, 19(4): 367–379
6. Rooney J J, Heuvel L N V. Root cause analysis for beginners.

Quality progress, 2004, 37(7): 45–56

7. Doggett A M. Root cause analysis: A framework for tool selection. Quality Management Journal, 2005, 12(4): 34
8. Demirli K, Vijayakumar S. Fuzzy assignable cause diagnosis of control chart patterns. Fuzzy Information Processing Society, 2008, 1–6
9. Ruiz-Sarmiento J R, Galindo C, Gonzalez-Jimenez J. Scene object recognition for mobile robots through semantic knowledge and probabilistic graphical models. Expert Systems with Applications, 2015, 42(22): 8805–8816
10. Kordy B, Pouly M, Schweitzer P. Probabilistic reasoning with graphical security models. Information Sciences, 2016, 342: 111–131
11. Wee Y Y, Cheah W P, Tan S C, Wee K K. A method for root cause analysis with a Bayesian belief network and fuzzy cognitive map. Expert Systems with Applications, 2015, 42(1): 468–487
12. Alaeddini A, Dogan I. Using Bayesian networks for root cause analysis in statistical process control. Expert Systems with Applications, 2011, 38(9): 11230–11243
13. Weidl G, Madsen A L, Israelson S. Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes. Computers & Chemical Engineering, 2005, 29(9): 1996–2009
14. Ferreira L, Borenstein D. A fuzzy-Bayesian model for supplier selection. Expert Systems with Applications, 2012, 39(9): 7834–7844
15. Cai B, Liu Y, Fan Q, Zhang Y, Liu S Y, Ji R. Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network. Applied Energy, 2014, 114: 1–9
16. Zeng S P, Li J, Ding L. Fault diagnosis system for 350 KA pre-baked aluminum reduction cell based on BP neural network. TMS–Light Metals, 2007
17. Taylor M P, Zhang W D, Wills V, Schmid S. A dynamic model for the energy balance of an electrolysis cell. Chemical Engineering Research & Design, 1996, 74(8): 913–933
18. Zhou T. Study of alumina concentration control based on intelligent characteristic model. Advanced Materials Research. Trans Tech Publications, 2011, 317: 1314–1317
19. Tessier J, Tarcy G P, Batista E, Wang X. Towards on-line monitoring of alumina properties at a pot level. Light Metals, 2012: 633–638
20. Zeng S P, Wang S, Qu Y. Control of temperature and aluminum fluoride concentration based on model prediction in aluminum electrolysis. Advances in Materials Science and Engineering, 2014, 3:1–5
21. Kolås S, Støre T. Bath temperature and AlF_3 control of an aluminum electrolysis cell. Control Engineering Practice, 2009, 17(9): 1035–1043
22. Ayyub B M, Klir G J. Uncertainty modeling and analysis in engineering and the sciences. CRC Press, 2006, 3–7
23. Hancock J M. Bayesian network (belief network, causal network, knowledge map, probabilistic network). Dictionary of Bioinformatics and Computational Biology, 2004, 5–10
24. Pearl J. Fusion, propagation, and structuring in belief networks. Artificial Intelligence, 1986, 29(3): 241–288
25. Bishop C M. Pattern Recognition and Machine Learning. New York: Springer-Verlag, 2006, 373–375
26. Kabak M, Burmaoğlu S, Kazançoğlu Y. A fuzzy hybrid MCDM approach for professional selection. Expert Systems with Applications, 2012, 39(3): 3516–3525
27. Wan S. Power average operators of trapezoidal intuitionistic fuzzy numbers and application to multi-attribute group decision making. Applied Mathematical Modelling, 2013, 37(6): 4112–4126
28. Li P, Chen G, Dai L, Zhang L. A fuzzy Bayesian network approach to improve the quantification of organizational influences in HRA frameworks. Safety Science, 2012, 50(7): 1569–1583
29. Dubois D, Prade H. Operations on fuzzy numbers. International Journal of Systems Science, 1978, 9(6): 613–626
30. Hsu Y L, Lee C H, Kreng V B. The application of fuzzy Delphi method and fuzzy AHP in lubricant regenerative technology selection. Expert Systems with Applications, 2010, 37(1): 419–425