IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 15, 2020, accepted March 8, 2020, date of publication March 18, 2020, date of current version March 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2981763

Quantitative Analysis of Human Error Probability in High-Speed Railway Dispatching Tasks

YANHAO SUN[1,2], QI ZHANG[2,3], ZHIMING YUAN[2,3], YING GAO[2,3], AND SHUXIN DING[2,3]

[1]Postgraduate Department, China Academy of Railway Sciences, Beijing 100081, China
[2]Signal and Communication Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China
[3]National Research Center of Railway Intelligence Transportation System Engineering Technology, Beijing 100081, China

Corresponding author: Qi Zhang (zhangqi@rails.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61790575 and Grant U1834211, in part by the Foundation of China State Railway Group Company, Ltd., under Grant N2019G020 and Grant P2018G008, and in part by the Foundation of China Academy of Railway Sciences Corporation Ltd., under Grant 2019YJ071.

ABSTRACT Human error can be regarded as a significant factor contributing to high-speed railway accidents. Cognitive Reliability and Error Analysis Method (CREAM) is well-known approach applied to determine Human Error Probability (HEP). However, shortcomings are still disclosed and weaken the applicability of such approach. These include the lack of sufficient failure data, lack of valid description of Common Performance Condition (CPC) and does not consider the CPCs weights. In addition, Basic CREAM does not provide a method to calculate the concrete HEP. In this paper, a modified CREAM is proposed to assess HEP of high-speed railway dispatchers in dispatching tasks. The core of the modified method is to use 2-tuple linguistic term sets to describe CPCs evaluation, combine weighted CPCs by Evidential Reasoning (ER) approach, and adopt Multi-Attribute Group Decision-Making (MAGDM) method to calculate HEP. To make CPCs weights more accurate, dynamically adjusting weights is adopted in this paper. The rationality and validity of the modified CREAM approach are verified by two axioms and compared other models. Finally this modified CREAM approach is applied to human reliability analysis of high-speed railway dispatchers.

INDEX TERMS Cognitive reliability and error analysis method (CREAM), human error probability (HEP), common performance condition (CPC), high-speed railway dispatchers, 2-tuple linguistic term sets, evidential reasoning (ER), multi-attribute group decision-making (MAGDM).

I. INTRODUCTION

In recent years, high-speed railway has developed rapidly in China. By the end of 2019, the operating mileage of high-speed railway in China has exceeded 35000 kilometers, accounting for about 70% of the world's total. Due to the characteristics of high-speed railway, such as high speed, high density and large traffic volume, once an accident occurs, it will cause a large number of casualties, huge economic losses and bad social influence. There is a survey found that 75% of railway traffic accidents are related to human factors [1]. Therefore, it is necessary to study the human factor reliability of railway transportation. As the nerve center of high-speed railway transportation system, high-speed railway dispatching system plays an important role in ensuring the safety and punctuality of trains. High-speed railway dispatching system is a complex system composed of four elements: human, equipment, environment and information. As a system element with independent thinking, train dispatchers play a leading role in coordination and control system. It is necessary and significant to study the human error of train dispatchers to prevent the risk of railway transportation. However, due to the lack of relevant work, the quantitative analysis of human error of high-speed railway dispatchers faces great challenges.

Human Reliability refers to the ability of participants to complete the specified tasks without error within the specified time and under the specified conditions [2]. The human error has been recognized as a predominant causal factor in the occurrence of many accidents in numerous domains, and many experts and scholars have devoted to developing and facilitating methodology and theories related to Human Reliability Analysis (HRA) [3]–[6]. HRA method has been divided into three generations, The first generation

The associate editor coordinating the review of this manuscript and approving it for publication was Chong Leong Gan.

methodologies of HRA focus on the study of human behavior theory and error classification, and forms a statistical analysis and prediction method of human error probability (HEP) based on operator experience and expert judgment, among which Technique for Human Error Rate Prediction (THERP) is the representative method [7]. The second generation methodologies further study the internal course of human behavior, focusing on the mechanism and HEP in the whole process from human observation, diagnosis, decision-making and other cognitive activities to the execution of actions in a specific situation. Cognitive Reliability and Error Analysis Methods (CREAM) is one of the most recognized methods of the second generation for addressing such contextual influence [8]. The third generation is dynamic HRA methodology based on simulation.

CREAM contains two versions, namely the basic and the extended ones. The basic method is used for determination of control modes and corresponding error rate intervals at a screening stage, while the extended method is employed for error quantification of cognitive functions. In 2012, Professor Hollnagel, the founder of CREAM method, issued a disclaimer, which pointed out that only the A (for Analysis) and M (for Method) make sense, Cognitive Reliability (the CR) is defective. This caused the extended method to no longer be applicable. Therefore, in recent years, many scholars have not carried out in-depth research on the extension method, but have improved the basic method and achieved good results. For instance, Ung [5] proposed a rule-based fuzzy CREAM model considering the weight of Common Performance Conditions (CPC) for marine oil tanker leakage accidents. The introduction of fuzzy set theory to express uncertain information can make the model easily convert the qualitative information into quantitative probability result. Yang *et al.* [9] proposed modified IF-THEN rules to construct the relationship between the nine CPCs and the control modes, where the control modes are expressed by belief degree rather than 100% certainty. He *et al.* [10] and Sun *et al.* [11] used the Context Influence Index (CII) to represent the comprehensive level of CPC and use it to calculate HEP.

A method of estimating the level of CPC based on fuzzy sets, and then calculating HEP according to fuzzy knowledge reasoning and membership function of control mode was used by Konstandinidou *et al.* [12] and Nivolianitou and Konstantinidou [13]. In addition, the Bayesian Network (BN), which has been widely used for HRA was introduced in order to deal with the uncertainty in the reasoning process [14]–[16]. Although the above methods have made remarkable achievements in the field of HRA, if CREAM is used to determine the HEP of high-speed railway dispatchers, there are still some problems need to be solved in the existing research on the CREAM method:(1) Lack of reliable historical data. Most CREAM studies dependent on the domain knowledge and experiences of the HRA analyzers and experts. (2) The linguistic variables of each CPC cannot be precisely described. (3) Most of the literatures do not consider the CPCs weights, or use AHP to obtain the weight.

TABLE 1. The control modes and probability intervals.

COCOM	Probability interval
Strategic	(0. 00005， 0. 01)
Tactical	(0. 001， 0. 1)
Tactical	(0. 01， 0. 5)
Scrambled	(0. 1， 1. 0)

AHP is simple to calculate, but it is too subjective to make the weights accurate. (4) The method of obtaining CII is rough and not very accurate. (5) Although IF-THEN rules and BN can solve the uncertainty problem well, the number of inference rules that IF-THEN rule needs to set is too large, and the conditional probability table of BN needs a lot of prior data.

This article will provide a comprehensive method to solve all the above-mentioned problems based on prior research. The main contributions of this work are shown in the following: (1) Constructed the CPCs detailed evaluation rules for high-speed railway dispatchers, and used 2-tuple linguistic term sets to evaluate CPCs to characterize the fuzziness and uncertainty of information. (2) Proposed a method of dynamic adjustment of weights, which can not only obtain the accurate experts' weights and CPCs weights, but also reduce the conflicts between different experts' evaluations. (3) Proposed a simple method for converting binary semantics into confidence. Based on this, ER algorithm is used to combine the degree of belief of CPCs to calculate CII. (4) Adopted Multi-Attribute Group Decision-Making (MAGDM) to improve CREAM, MAGDM has the advantages of reasonable, brainstorming, and minimizing unreasonable factors in calculation.

The paper is arranged as follows. Section II gives a brief introduction about basic CREAM theory, 2-tuple linguistic term sets, and Evidential Reasoning approach. In Section III, the modified CREAM approach is proposed. Section IV gives a case of the high-speed railway dispatchers' performances to demonstrate the effectiveness of the proposed mode. Finally, Section V concludes the whole paper.

II. PRELIMINARIES

A. CREAM

The core of CREAM is that human error is not stochastic, but more shaped by the context of the task. In the basic CREAM, the Contextual Control Model (COCOM) is defined as the competence of operator to adapt to the environment. COCOM is defined by four characteristic control modes, namely, Scrambled, Tactical, Tactical and Strategic according to the human cognition and action context. Each control mode has its corresponding HEP probability interval, as shown in Table 1.

CREAM identifies nine CPCs, which are shown in Table 2. Each CPC has different effects on human performance, including reduced (negative), insignificant (neutral) or improved (positive). By calculating the total numbers of CPCs with improved, reduced effects are denoted as Σ_{improved} and Σ_{reduced} respectively. The control mode determined by coordinate mode (Σ_{improved}, Σ_{reduced}), which

TABLE 2. Nine CPCs name.

CPC number	CPC name
C_1	Adequacy of organization
C_2	Working conditions
C_3	Adequacy of MMI and operational support
C_4	Availability of procedures/plans
C_5	Number of simultaneous goals
C_6	Available time
C_7	Time of the day (circadian rhythm)
C_8	Adequacy of training and experience
C_9	Crew collaboration quality

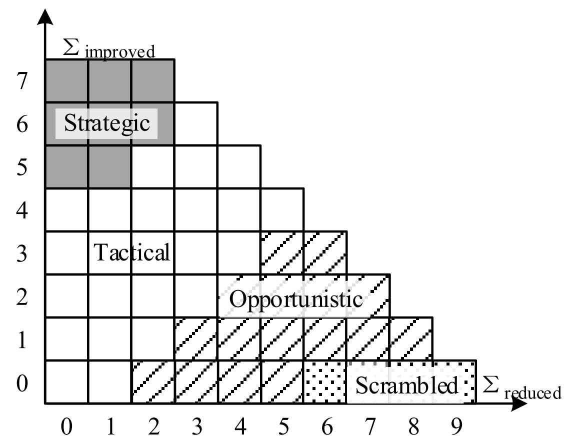

FIGURE 1. Determination of control modes.

is shown in Figure 1 below. Finally, the HEP probability interval is obtained by the control mode.

B. 2-TUPLE LINGUISTIC TERM SETS

Definition 1 [17]: Let $S = \{s_0, s_1, \ldots, s_g\}$ be a linguistic term set with odd cardinality. For any $s_i, s_j \in S$, where $i, j \in \{0, 1, \ldots, g\}$. The following attributes for S can be defined as:

(1) if $i < j$, then $s_i \prec s_j$;

(2) Negation operator: $Neg(s_i) = s_{g-i}$.

Definition 2 [17]: Let $S = \{s_0, s_1, \ldots, s_g\}$ be a linguistic term set. $\beta \in [0, g]$ be the result of an aggregation of the indexes of a set of labels assessed in a linguistic term set S, i.e., the result of a symbolic aggregation. Let $i = round(\beta)$ and $\alpha = \beta - i$ be two values such that $i \in [0, g]$ and $\alpha \in [-0.5, 0.5)$, then α is called a symbolic translation.

Definition 3 [17]: Let $S = \{s_0, s_1, \ldots, s_g\}$ be a linguistic term set. $\beta \in [0, g]$ be a value representing the result of a symbolic aggregation operation, then the 2-tuple that expresses the equivalent information to β is obtained by the function Δ:

$$\Delta : [0, g] \rightarrow S \times [-0.5, 0.5),$$
$$\Delta(\beta) = (r_i, \alpha_i) = \begin{cases} s_i & i = \text{round}(\beta) \\ \alpha = \beta - i, & \alpha \in [-0.5, 0.5). \end{cases}$$

where $i = \text{round}(\beta)$, $\alpha = \beta - i, \alpha \in [-0.5, 0.5)$, $s_i \in S$, round $(\cdot)$ is the usual round operation, s_i has the closest index label to β and α is the value of the symbolic translation. Contrarily, Δ is one to one function. There is always a function Δ^{-1}, from 2-tuple to the corresponding number:

$$\Delta^{-1} : S \times [-0.5, -0.5) \rightarrow [0, g],$$
$$\Delta^{-1}(s_i, \alpha_i) = i + \alpha_i = \beta.$$

Definition 4 [17]: Let (s_i, α_i) and (s_j, α_j) be two 2-tuples, then the distance between (s_i, α_i) and (s_j, α_j) is defined as follows:

$$d\left((s_i, \alpha_i), (s_j, \alpha_j)\right) = \sqrt{\left(\frac{\Delta^{-1}(s_i, \alpha_i)}{g}\right)^2 - \left(\frac{\Delta^{-1}(s_j, \alpha_j)}{g}\right)^2}. \quad (1)$$

C. EVIDENTIAL REASONING

The ER approach, which incorporates fuzzy set theory, decision theory, Bayesian probability theory, and the Dempster-Shafer (D-S) theory, is a way of dealing with different kinds of Uncertain Multi Attribute Decision Analysis (UMADA) problems [18]–[20]. The basic concepts and definitions of the ER algorithm relevant to this paper are briefly described as follows.

Suppose there is a simple two-level hierarchy of attributes with a general attribute Y at the top level and a number of basic attributes $E = \{e_i | i = 1, 2, \ldots, l\}$ at the bottom level. The weights of the attributes are given by $\omega = \{\omega_i | i = 1, 2, \ldots, l\}$, where $0 \leq \omega_i \leq 1$, $\sum_{i=1}^{l} \omega_i = 1$. Let Y be assessed at the l attributes on the basis of n distinctive evaluation grades $H = \{h_j | j = 1, 2 \ldots, n\}$.Without loss of generality, it is assumed that h_{j+1} is preferred to h_j. Then the evaluation of attribute e_i can be expressed as follows:

$$V(e_i) = \left\{\left(h_j, \beta_{i,j}\right), i = 1, 2, \ldots, l; j = 1, 2, \ldots, n\right\}. \quad (2)$$

where $\beta_{i,j} \geq 0$, $\sum_{j=1}^{n} \beta_{i,j} \leq 1$, and $\beta_{i,j}$ denotes a degree of belief of e_i. Assuming that the degree of belief of all the basic attributes is known, the reliability of the general attribute Y on H can be synthesized by ER algorithm. The process is briefly described as the following steps.

Let $m_{i,j}$ be a basic probability mass representing the degree to which the basic attribute e_i supports the hypothesis that the general attribute Y is assessed to the grade h_j. Let $m_{i,H}$ be a remaining probability mass unassigned to any individual grade after all the n grades have been considered for assessing the general attribute Y as far as e_i is concerned. $m_{i,j}$ and $m_{i,H}$ can be expressed as follows:

$$m_{i,j} = \omega_i \beta_{i,j}. \quad (3)$$

$$m_{i,H} = 1 - w_i \sum_{i=1}^{l} \beta_{i,j} = \bar{m}_{i,H} + \tilde{m}_{i,H}. \quad (4)$$

where $\bar{m}_{i,H}$ is caused by the relative importance of the attribute e_i and $\tilde{m}_{i,H}$ by the incompleteness of the assessment on e_i for the general attribute Y.

Define $E_{I(i)}$ as the subset of the first i basic attributes as follows:

$$E_{I(i)} = \{e_1, e_2, \ldots, e_i\}.$$

Suppose $m_{I(i),j}$ be a probability mass defined as the degree to which all the attributes in $E_{I(i)}$ support the hypothesis that Y is assessed to the grade h_j.$m_{I(i),H}$ is the remaining probability mass unassigned to individual grades after all the basic attributes in $E_{I(i)}$ have been assessed. $m_{I(i),j}$ and $m_{I(i),H}$ can be generated by the following recursive ER algorithm.

$$m_{I(i+1),j} = K_{I(i+1)}[m_{I(i),j} \times m_{i+1,j} + m_{i+1,j} \times m_{I(i),h} + m_{i+1,h} \times m_{I(i),j}]. \quad (5)$$

$$m_{I(i+1),H} = K_{I(i+1)} \times m_{I(i),H} \times m_{i+1,H}. \quad (6)$$

$$K_{I(i+1)} = \left[1 - \sum_{j=1}^{n} \sum_{\substack{p=1 \\ p \neq j}}^{n} m_{I(i),j} \times m_{p,i+1}\right]^{-1}, \quad i = 1, 2, \ldots, l-1. \quad (7)$$

where K is the conflict factor, indicating the extent to which different attributes support a certain evaluation grade. In the original ER approach, the combined degree of belief β_j is directly given by

$$\beta_j = \frac{m_{I(l),j}}{1 - \bar{m}_{I(l),H}}. \quad (8)$$

$$\beta_H = m_{I(l),H} = 1 - \sum_{j=1}^{n} \beta_j. \quad (9)$$

where β_H is the degree of belief unassigned to any individual evaluation grade after all the l basic attributes have been assessed. It denotes the degree of incompleteness in the assessment generated.

Therefore, the degree of belief of generalized attribute Y can be obtained, which is shown in:

$$V(Y) = \{(h_1, \beta_1), (h_2, \beta_2), \ldots, (h_n, \beta_n), (h_H, \beta_H)\}. \quad (10)$$

III. MODIFIED CREAM APPROACH

A. PRECONDITION DESCRIPTION

CREAM can be seen as a MAGDM problem, which has obvious characteristics of MAGDM. Experts evaluate multiple CPCs and then calculate HEP based on the evaluation results. So we can use the MAGDM method to analyze CREAM.

In this paper, l experts are invited to evaluate the n CPCs when the high-speed train dispatchers perform m different tasks, so as to evaluate the human reliability of high-speed railway dispatchers. Let $A = \{A^k | k = 1, 2, \ldots, l\}$ be a set of tasks performed by the high-speed railway dispatchers, $C = \{C_i | i = 1, 2, \ldots, m\}$ be a set of CPCs. And $\boldsymbol{w}_i(i = 1, 2, \ldots, m)$ be the CPCs weights vector, where $0 \leq w_i \leq 1$, $\sum_{i=1}^{m} w_i = 1$. Let $B = \{B_j | j = 1, 2, \ldots, n\}$ be a set of the evaluation experts, $\boldsymbol{\lambda}_j^s (j = 1, 2, \ldots, n)$ be the evaluation experts subjective weights vector, with $\lambda_j^s \geq 0$, $\sum_{k=1}^{l} \lambda_s^k = 1$. Suppose that $\boldsymbol{R}^k = (r_{ij}^k)_{m \times n}$ and $\boldsymbol{T} = (t_{ij})_{m \times n}$ are the two 2-tuple linguistic evaluation matrixes, r_{ij}^k takes the form of the 2-tuple linguistic, which is the evaluation of CPC C_i given by expert B_j for task A^k. t_{ij} also takes the 2-tuple linguistic form, which denotes that evaluation of the importance of CPC C_i by expert B_j.

B. CPCS ADJUSTMENT RULES

CREAM provides a simple description of the nine CPCs, and it did not give detailed evaluation rules. In order to apply CREAM to the HRA of the dispatchers, refer to the existing dispatching operation rules and handbooks, this article formulates detailed evaluation rules for each CPC as shown in Table 3.

Most of the literatures ignored the interaction of CPCs in the calculation of HEP, and could not accurately capture the joint CPC effects. The ER algorithm requires independent evidence when it is used, i.e., no correlation between each CPC, so the derivation of the combined CPC evaluation must consider how the dependencies can be concretely treated. The rules for considering dependencies and adjusting CPC primary effect were defined by Hollnagel [8] and are shown in Table 4.

According to the rules described in Table 4, for instance, the state of C_2 will be updated into the positive/negative status only when it has a neutral effect (other states don't work) on human reliability and 4 out of 5 CPCs it depends on have positive/negative ones simultaneously.

As this article uses 2-tuple linguistic to evaluate CPC, 2-tuple linguistic is a kind of fuzzy linguistic, so the adjusted evaluation of CPC needs to be amended. For instance, when the state of C_2 becomes positive/negative status, its evaluation should be changed to the min/max value in the rule CPC it depends on.

C. CALCULATING CPC WEIGHT

The most widely used method to obtain the CPCs weights is the Analytic Hierarchy Process (AHP) [9], [16]. Although the AHP is simple to calculate, it is too subjective. In order to make the calculation of weights more accurate, we provide a combined model to determine the CPCs weights in this paper.

First of all, we use experts' evaluations method to obtain the CPCs subjective weights as follows:

$$w_i^s = \frac{\sum_{j=1}^{n} \Delta^{-1}(t_{ij}) \lambda_j^s}{\sum_{i=1}^{m} \sum_{j=1}^{n} \Delta^{-1}(t_{ij}) \lambda_j^s} \quad i = 1, 2, \ldots, m; \; j = 1, 2, \ldots, n. \quad (11)$$

Next, we use the entropy method to determine the CPCs objective weights. Mon *et al.* [21] proposed that the

TABLE 3. CPCs detailed evaluation regulations.

CPC	Sub-index	The detailed evaluation regulations
Adequacy of organization C_1	Organization, responsibilities and management system	Whether the internal organization structure of the dispatching agency is complete, whether the division of responsibilities is clear, and whether a good job responsibility system is established.
	Communication and coordination	Whether the team meets regularly for communication and exchanges, and whether team members perform pre-shift handovers, site inspections and equipment inspections, and whether the communication is smooth.
	Safety performance assessment and rewards and punishments	Completeness of safety inspection, supervision, regular meetings and assessment systems, and whether the reward and punishment measures are qualified.
	Safety culture of team	Does the team carry out safety education activities, and does the dispatcher have a sense of safety in actively investigating hidden dangers and consciously operating according to standards? Are there sufficient funds for safety to improve safety technology measures and create a safety culture atmosphere?
Working conditions C_2	Physical environment	Whether the temperature and humidity of the dispatching hall are appropriate, whether the lighting is sufficient, and whether the noise around the work area is too loud.
	Soft environment	Whether the working area environment is tidy and orderly, and whether various tools and documents are put back in place in time after use.
Adequacy of MMI and operational support C_3	Equipment function and quality	Whether the functions of various hardware of the dispatching system are perfect and whether the communication equipment is intact.
	man-machine interface (MMI)	Whether the display of interface information is clear and comprehensive, and whether the operation of dispatching system software system is convenient and fast.
	Intelligent alarm	Whether the alarm content is accurate and comprehensive, and whether the alarm is clear.
Availability of procedures/plans C_4	Equipment operating procedures	Whether the operating instructions of the system's software and hardware are feasible and comprehensive.
	Emergency treatment specification	Whether there is a complete and feasible response plan.
Number of simultaneous goals C_5	Difficulty of task	Whether the technical complexity of the target task completion conforms to the objective law, the mental and physical exhaustion of the operator, and whether it is prone to fatigue, etc.
	Number of multiple targets	Whether overload occurs when multiple tasks are performed at the same time.
Available time C_6	Urgency of the task	The time and speed required to complete the task, the severity of the consequences of the emergency, and whether there are remedial measures.
	Time to complete the task	Whether the schedule for completing tasks is balanced and reasonable, and whether the tasks can be completed in a timely manner.
Time of the day (circadian rhythm) C_7	Physiological functions on duty	Influence degree of physical, mental and working state of dispatcher on operation quality.
	The law of work and rest	Whether the dispatcher has good work and rest rules.
Adequacy of training and experience C_8	Knowledge and skills training	Whether the training of dispatchers is comprehensive and has the necessary knowledge and culture level.
	Operational experience	Whether the dispatcher has rich operating experience and whether he frequently exchanges operating experience.
Crew collaboration quality C_9	Personality characteristics	Whether the dispatcher's age and physical conditions meet the requirements; whether the work mood, risk appetite and concentration are good.
	Division and Coordination	Whether the tasks of each jobber are clear and the allocation is reasonable; whether the communication and feedback are smooth, the process connections are tight, and coordination is consistent.
	Team atmosphere	Whether team members trust and cooperate with each other and have good team spirit.

TABLE 4. Rules for adjusting CPCs.

CPC	Depends on the following CPCs					Threshold
C_2	C_1	C_3	C_6	C_7	C_8	4
C_5	C_2	C_3	C_4	—	—	2
C_6	C_2	C_3	C_4	C_5	C_7	4
C_9	C_1	C_8	—	—	—	2

importance of an attribute can be measured by the entropy of the attribute. According to the entropy theory, if the evaluation value of a CPC on different tasks is closer, the entropy value of the CPC is larger. It's easy to see the larger the CPC entropy value, the smaller the degree of difference, and the smaller its weight [22]. So that the CPC C_i entropy value is defined as follows:

$$E_i = -\sum_{k=1}^{l}(\Delta^{-1}(r_i^k)\Big/\sum_{k=1}^{l}\Delta^{-1}(r_i^k)) \times \ln(\Delta^{-1}(r_i^k)\Big/\sum_{k=1}^{l}\Delta^{-1}(r_i^k)) \quad i = 1, 2, \ldots, m. \tag{12}$$

where $r_i^k = \sum_{j=1}^{n} r_{ij}^k \lambda_j^s$, if $r_i^k = 0$, then:

$$\Delta^{-1}(r_i^k)\Big/\sum_{k=1}^{l}\Delta^{-1}(r_i^k) \times \ln(\Delta^{-1}(r_i^k)\Big/\sum_{k=1}^{l}\Delta^{-1}(r_i^k)) = 0. \tag{13}$$

where $\Delta^{-1}(r_i^k)\Big/\sum_{k=1}^{l}\Delta^{-1}(r_i^k)(k = 1, 2, \ldots, l)$ are equal to each other, the entropy value is the largest at this time, and the maximum entropy is $(E_i)_{\max} = \ln k$. Normalize Eq. (12) by using $(E_i)_{\max}$, we get: $e_i = (1/\ln k)E_i(i = 1, 2, \ldots, m)$.

The CPCs objective weights can be calculated as follows:

$$w_i^o = \frac{1 - e_i}{m - \sum_{i=1}^{m} e_i} \quad i = 1, 2, \ldots, m. \tag{14}$$

Finally, the combined weights can be defined as follows:

$$w_i = \rho w_i^s + (1 - \rho) w_i^o \quad i = 1, 2, \ldots, m. \tag{15}$$

where ρ is the weight balance coefficient, Without loss of generality $\rho = 0.5$ is adopted in this paper.

D. OBTAINING EXPERTS WEIGHT

In the process of MAGDM, it is generally considered that there is a trend of consistency between individual evaluation and group evaluations [23]. Therefore, experts can be given different weights according to the degree of deviation between individual expert evaluations and group evaluation results. If the individual expert evaluation and group evaluation results are closer, the expert weight is greater. While the individual expert evaluation deviates from the group evaluation results, the expert weight is smaller.

For the expert B_j, we defined that the deviation D_j between the expert B_j evaluation and the group evaluation:

$$D_j = \sum_{k=1}^{l} d(r_j^k, r^k) \quad k = 1, 2, \ldots, l. \tag{16}$$

$$r_j^k = \sum_{i=1}^{m} r_{ij}^k w_i \quad j = 1, 2, \ldots, n;\ k = 1, 2, \ldots, l. \tag{17}$$

$$r^k = \sum_{j=1}^{n} r_j^k \lambda_j^s \quad j = 1, 2, \ldots, n. \tag{18}$$

where r_j^k, r^k represent individual and group evaluations of experts with task A^k respectively. Then the experts' weights can be defined as follows:

$$\lambda_j^o = \frac{D_j}{\sum_{j=1}^{n} D_j} \quad j = 1, 2, \ldots, n. \tag{19}$$

The experts combined weight can be calculated as follows:

$$\lambda_j = \eta\lambda_j^s + (1 - \eta)\lambda_j^o \quad j = 1, 2, \ldots, n. \tag{20}$$

where η is the weight balance factor, Without loss of generality $\eta = 0.5$ is adopted in this paper.

In order to reduce the conflict between experts' evaluations and make the group evaluation consistent, we dynamically adjust the experts' weights and CPCs weights.

Step 1: We replace the initial w_i and λ_j^s in Eq. (16) and Eq. (17) with new w_i and λ_j to obtain new r_j^k and r^k respectively, which are defined as $\left(r_j^k\right)'$ and $\left(r^k\right)'$.

Step 2: Compare the gap between the group evaluation results r^k and the last evaluation $\left(r^k\right)'$, and we define the gap as follows:

$$D(r) = \sqrt{\sum_{k=1}^{l}\left(d\left(\left(r^k\right)', r^k\right)\right)^2} \tag{21}$$

Then, we need to set a threshold δ for $D(r)$. If $D(r) \le \delta$, the gap of between r_j and r_j' is small. It believe that the group evaluation tends to be stable and consistent, then output the CPCs weights w_i and experts' weights λ_j at this time.

Step 3: If $D(r) > \delta$, the difference of between r_j and r_j' is large, so we need to dynamically adjust the CPCs and the experts' weights.

Bring $\left(r_j^k\right)'$ and $\left(r^k\right)'$ into Eq. (16) to get new D_j, and then bring D_j into Eq. (15) to get new experts' subjective weights λ_j^o, and use Eq. (19) to calculate the new experts' combined weights λ_j.

Then, replace λ_j^s in Eq. (11) with λ_j, and get the CPCs objective weights w_i^o by Eq. (14). Finally, the CPCs comprehensive weights w_j are calculated by Eq. (15).

Repeat the above steps until $D(r) \le \delta$, then output the CPCs weights w_i and experts' weights λ_j.

Step 4: Take the adjusted experts' weights and CPCs weights as their definitive weights.

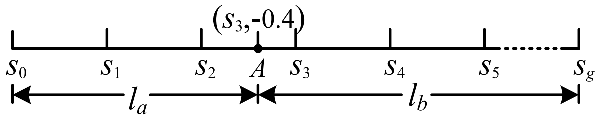

FIGURE 2. Geometric representation of 2-tuple linguistic.

E. CALCULATION OF HEP

In order to better illustrate the modified CREAM, two widely acknowledged assumptions must be presented be forehand.

(1) The control mode space is continuous [8], [24].

(2) HEP is also continuous, and it varies with the context exponentially [25]

The above assumptions have been used in many literatures and the results are validated to be acceptable in practice [10], [11], [26]. CII was defined by Sun *et al.* [11]:

$$CII = \frac{\Sigma_{\text{improved}}}{\max(\Sigma_{\text{improved}})} - \frac{\Sigma_{\text{reduced}}}{\min(\Sigma_{\text{reduced}})} = \frac{\Sigma_{\text{improved}}}{7} - \frac{\Sigma_{\text{improved}}}{9}. \tag{22}$$

The functional relation between the CII variable and HEP can be constructed as follows [11]:

$$\text{HEP} = \text{HEP}_0 \times \exp(\mu \times \text{CII}). \tag{23}$$

Obviously, the maximum and minimum values of CII are 1 and -1 respectively. By substituting them into Eq. (23), we have:

$$\begin{cases} \text{HEP}_{\min} = \text{HEP}_0 \times \exp(\mu) \\ \text{HEP}_{\max} = \text{HEP}_0 \times \exp(-\mu). \end{cases} \tag{24}$$

From Table 1, it can be found that $\text{HEP}_{\min} = 0.00005$ and $\text{HEP}_{\max} = 1$. Then, the values of HEP_0 and μ could be calculated by Eq. (24).

$$\text{HEP}_0 = 7.07 \times 10^{-3}; \quad \mu = -4.9517. \tag{25}$$

The model key issue for calculating HEP is to obtain CII. Similarly with the CII proposed by Sun *et al.*, we also use integrated positive reliability performance (improved) minus the integrated negative reliability performance (reduced) as the CII. However, since this article uses 2-tuple linguistic to evaluate CPC. Only when the 2-tuple linguistic evaluation is converted into a degree of belief can the CII be calculated using the ER algorithm. Take Figure 2 as an example.

The 2-tuple linguistic coordinate of point A in Figure 2 is $(s_3, -0.4)$, then the degree of belief of "reduced" at point A is l_b/g, and the degree of belief of "improved" is l_a/g, where $l_a = \Delta^{-1}(s_3, -0.4), l_b = g - l_a$. If $l_a = l_b$, then the state of CPC is neutral effect, i.e., neither "improved" nor "reduced". Let β_1, β_2 and β_3 be the degree of belief of "reduced", "not significant" and "improved", respectively. Then β_1, β_2 and β_3 are shown as follows:

$$\beta_1 = l_b/g, \quad \beta_2 = \begin{cases} 1 & \text{when } l_b = l_a \\ 0 & \text{when } l_b \neq l_a, \end{cases} \beta_3 = l_a/g. \tag{26}$$

The HEP can be calculated by Eq. (24)

$$\text{HEP} = 7.07 \times 10^{-3} \times \exp[-4.9517 \times \text{CII}] = 7.07 \times 10^{-3} \times \exp[-4.9517 \times (\beta_3 - \beta_1)]. \tag{27}$$

The neutral effects are not considered in Eq. (24), because the neutral effects are not significant for calculation of the HEP which has been validated by many literatures [5], [9], [11].

The steps of the modified CREAM approach are given as follows:

Step1: Obtain experts' evaluation matrix $\boldsymbol{R}^k = (r_{ij}^k)_{m\times n}$ with CPCs and their importance assessment matrix $\boldsymbol{T} = (t_{ij})_{m\times n}$.

Step2: Check whether the CPCs evaluation matrix $\boldsymbol{R}^k = (r_{ij}^k)_{m\times n}$ needs to be adjusted. If $\boldsymbol{R}^k = (r_{ij}^k)_{m\times n}$ needs to be adjusted, adjust $\boldsymbol{R}^k = (r_{ij}^k)_{m\times n}$ to $\hat{\boldsymbol{R}}^k = (\hat{r}_{ij}^k)_{m\times n}$ according to Table 4.

Step3: Calculate the CPCs subjective weights based on importance matrix $\boldsymbol{T} = (t_{ij})_{m\times n}$, and calculate the CPCs objective weights and the experts' objective weights based on matrix $\boldsymbol{R}^k = (r_{ij}^k)_{m\times n}$ or $\hat{\boldsymbol{R}}^k = (\hat{r}_{ij}^k)_{m\times n}$. Then get the definitive CPCs weights $w_i(i = 1, 2, \ldots, m)$ and experts' weights $\lambda_j(j = 1, 2, \ldots, n)$ after dynamic adjustment.

Step4: The matrix $\boldsymbol{R}^k = (r_{ij}^k)_{m\times n}$ or $\hat{\boldsymbol{R}}^k = (\hat{r}_{ij}^k)_{m\times n}$ is transformed into the degree of belief of CPC corresponding effects, and then the degree of belief of nine CPCs effects are synthesized with ER algorithm to get the degree of belief of each expert about integral CPC effects. Then the degree of belief of each expert about integral CPC effects are combined with ER algorithm to obtain the degree of belief of all experts about integral CPC effects (β_1, β_2 and β_3).

Step5: Finally, the HEP is obtained by Eq. (27).

The flowchart in Figure 3 shows the modified CREAM approach for process.

IV. CASE STUDY

A. TASK SELECTION AND ANALYSIS

High speed railway dispatching system is a safety-critical systems. Therefore, the human reliability requirements of the dispatcher is relatively high.

To calculate the HEP of the dispatcher, we choose two common tasks, which are "Temporary speed restriction of train control system" as task 1 and "Centralized traffic control (CTC) system control mode conversion" as task 2. Two tasks disposal workflow can be provided in Table 5 and Table 6.

B. CALCULATION OF HEP

Task1: A railway line of a railway group company suffered a strong wind at 15:32 on 12 July 2019. For the safety of the trains, speed restriction is required. Therefore, the dispatchers needs to perform temporary speed restriction of train control system. The range of the speed restriction range is the downward line 1286km+868m to 1293km +300m. At this time, the ambient temperature was 26.5°C with 52% humidity in the dispatch hall. The train dispatcher, train assistant dispatcher, and deputy director on duty are all employees with many years of work experience. The

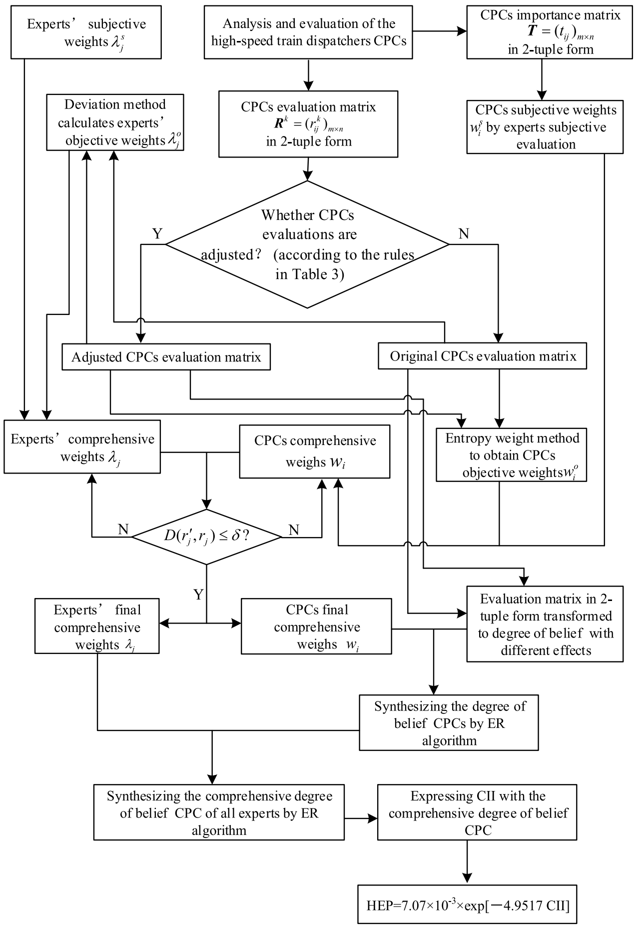

FIGURE 3. The flowchart of the modified CREAM approach.

team members have a good cooperative relationship and they communicate smoothly with other types of workers. The human-machine interface interaction of the CTC dispatch terminal is good, and the communication equipment is fault-free.

Task2: At 11 am on November 16, 2019, due to equipment failure, the CTC control mode needs to be changed from DCCM to ASCM. After rush repair by the staff, the CTC equipment was repaired and the CTC control mode was restored from ASCM to DCCM. The dispatchers of the two tasks are the same. At this time, the ambient temperature was 26°C with 50% humidity in the dispatch hall. The dispatching hall is shown in Figure 4.

Four experts B_1, B_2, B_3 and B_4 with different experience and knowledge (According to their knowledge and experience, the subjective weights of four experts were 0.4, 0.3, 0.2 and 0.1) were invited to analyze the operation video and log of the high-speed train dispatchers for two tasks, Table 7 to Table 9 shows the evaluations of each CPC given by the four experts. The CPCs evaluation with 2-tuple linguistic is defined as

$$S=\{s_0: \text{very poor}, s_1: \text{poor}, s_2: \text{slightly poor}, s_3: \text{medium}, s_4: \text{slightly good}, s_5: \text{good}, s_6: \text{very good}\}.$$

TABLE 5. Disposal workflow of "Temporary speed restriction of train control system".

No.		Disposal workflow of task 1.
1		Confirm speed restriction mileage and speed restriction value.
2		Handling of related trains.
	1	Trains that have entered the speed restriction section will immediately notify the driver of speed limit operation.
	2	Trains that have left the departure signal but have not started (toward the speed limit section) will immediately inform the driver to stop in the station.
	3	Set the departure route of the first train that is about to enter the speed limit zone to be triggered manually ("The departure route of XX stops of XX trains at XX stations is manually triggered).
3		Issue speed restriction command.
	1	Enter the command at the dispatch command terminal: "XX line XX Km XX m to XX Km XX m speed limit XX kilometers, set train control speed limit".
	2	Enter speed restriction parameters.
	3	Arrange train assistant dispatcher to write command.
	4	Check the contents of dispatch command and release it
	5	Confirm whether the speed restriction command is implemented one by one.
4		Let relevant trains pass.
	1	Notify the parking train in the station to continue to run and then resume the departure route one by one for automatic triggering.
	2	Set the departure route of subsequent trains requiring speed limit one by one as manual triggering.
5		Cancel train control speed restriction.
	1	Enter the command at the dispatch command terminal: "XX line XX Km XX m to XX Km XX m speed limit XX kilometers, cancel train control speed limit".
	2	Input the speed limit parameter and issue the command.

And CPCs importance evaluation with 2-tuple linguistic is defined as

$$H=\{h_0: \text{unimportant}, h_1: \text{slightly unimportant}, h_2: \text{medium}, h_3: \text{slightly important}, h_4: \text{important}\}.$$

The concrete calculation steps are listed as follows:

Step1: For the CPC that needs to be adjusted, it can be known from the adjustment rules in Table 4 that the evaluation matrix $\boldsymbol{R}^1 = (r_{ij}^1)_{9\times 4}$ needs to be adjusted. The adjusted evaluation matrix $\hat{\boldsymbol{R}}^1 = (\hat{r}_{ij}^1)_{9\times 4}$ shown in Table 10.

TABLE 6. Disposal workflow of "CTC control mode conversion".

No.		Disposal workflow of task 2.
1		CTC equipment failure and needs to temporarily switch dispatching control mode.
2		Switching the CTC control mode. Dispatching Center Control Mode (DCCM) changed to Abnormal Station Control Mode (ASCM).
	1	Notify the deputy director on duty and the staff on duty at the station.
	2	Registered in the "CTC control mode conversion register" and signed and confirmed by the duty officer.
	3	Notify the station to change to ASCM.
	4	Confirm whether the conversion is successful by controlling the mode indicator light.
	5	After the control mode conversion is successful, confirm that the CTC terminal can display normally and the route is correct.
3		CTC equipment failure repaired, control mode switched back again. ASCM changed to DCCM.
	1	Registered in the "CTC control mode conversion register" and signed and confirmed by the duty officer again.
	2	Notify the station to switch back to the DCCM.
	3	After confirming that the dispatch terminal received the conversion request, it then dictated the order: "Agree XX station to be converted to central control."
	4	Execute a command that allows conversion.
	5	Confirm whether the conversion is successful by controlling the mode indicator light.
	6	Adjust the train operation plan in time, and monitor the follow-up status of the route sequence.

Step2: Calculate the definitive combined weights of experts and CPCs.

Adopt the data in Table 9 to obtain the CPCs subjective weights with Eq. (11).

$$w_1^o = 0.073, w_2^o = 0.092, w_3^o = 0.115, w_4^o = 0.085, w_5^o = 0.140, w_6^o = 0.132, w_7^o = 0.100, w_8^o = 0.148, w_9^o = 0.115.$$

Compute the CPCs objective weights by Eq. (9) and Eq. (10).

$$w_1^s = 0.080, w_2^s = 0.110, w_3^s = 0.140, w_4^s = 0.056, w_5^s = 0.074, w_6^s = 0.141, w_7^s = 0.091, w_8^s = 0.184, w_9^s = 0.124.$$

FIGURE 4. The dispatching hall.

TABLE 7. CPCs Evaluation matrix $R^1 = (r_{ij}^1)_{9\times 4}$ for task "Temporary speed restriction of train control system".

CPC	B_1	B_2	B_3	B_4
C_1	$(s_5,-0.4)$	$(s_6,0)$	$(s_5,0.2)$	$(s_6,-0.3)$
C_2	$(s_6,0)$	$(s_5,0.3)$	$(s_5,0.3)$	$(s_5,0)$
C_3	$(s_5,-0.2)$	$(s_4,0.4)$	$(s_6,0)$	$(s_5,0.1)$
C_4	$(s_3,0.2)$	$(s_3,0)$	$(s_3,0.1)$	$(s_4,-0.3)$
C_5	$(s_4,0.3)$	$(s_4,0.2)$	$(s_3,0)$	$(s_3,0)$
C_6	$(s_3,0)$	$(s_4,0)$	$(s_4,-0.2)$	$(s_4,0.3)$
C_7	$(s_4,0.2)$	$(s_5,-0.3)$	$(s_5,0)$	$(s_5,-0.1)$
C_8	$(s_4,0.4)$	$(s_4,0)$	$(s_4,0)$	$(s_4,-0.2)$
C_9	$(s_5,0)$	$(s_6,-0.3)$	$(s_6,0)$	$(s_6,0)$

TABLE 8. CPCs Evaluation of matrix $R^2 = (r_{ij}^2)_{9\times 4}$ for task "CTC control mode conversion".

CPC	B_1	B_2	B_3	B_4
C_1	$(s_5,-0.4)$	$(s_6,0)$	$(s_5,0.3)$	$(s_6,-0.1)$
C_2	$(s_5,0.1)$	$(s_5,0)$	$(s_5,-0.2)$	$(s_5,0)$
C_3	$(s_5,-0.2)$	$(s_5,0.4)$	$(s_6,0)$	$(s_5,0)$
C_4	$(s_3,0.4)$	$(s_4,0)$	$(s_4,0.1)$	$(s_4,-0.2)$
C_5	$(s_5,0.2)$	$(s_5,0)$	$(s_4,-0.5)$	$(s_4,0)$
C_6	$(s_5,0)$	$(s_5,0.2)$	$(s_5,-0.3)$	$(s_5,0.1)$
C_7	$(s_4,0.4)$	$(s_5,0)$	$(s_5,0)$	$(s_5,-0.3)$
C_8	$(s_4,0.4)$	$(s_4,0)$	$(s_4,0)$	$(s_4,-0.2)$
C_9	$(s_5,0)$	$(s_6,-0.3)$	$(s_6,0)$	$(s_6,0)$

Determine the CPCs combined weights by Eq. (15).

$$w_1 = 0.076, w_2 = 0.101, w_3 = 0.128, w_4 = 0.071, w_5 = 0.107, w_6 = 0.136, w_7 = 0.096, w_8 = 0.166, w_9 = 0.119.$$

TABLE 9. CPCs importance evaluation matrix $T = (t_{ij})_{9\times4}$.

CPC	B_1	B_2	B_3	B_4
C_1	$(h_2,-0.2)$	$(h_2,0)$	$(h_2,0.1)$	$(h_2,0)$
C_2	$(h_2,0)$	$(h_3,0)$	$(h_2,0.3)$	$(h_3,0)$
C_3	$(h_3,-0.1)$	$(h_3,0)$	$(h_3,0.3)$	$(h_4,-0.4)$
C_4	$(h_2,0.2)$	$(h_2,0.3)$	$(h_2,0)$	$(h_3,0)$
C_5	$(h_4,-0.3)$	$(h_4,-0.5)$	$(h_4,0)$	$(h_4,0)$
C_6	$(h_4,0)$	$(h_3,0.3)$	$(h_3,0.1)$	$(h_3,0)$
C_7	$(h_3,-0.2)$	$(h_3,0)$	$(h_2,0)$	$(h_2,0.4)$
C_8	$(h_4,0)$	$(h_4,0)$	$(h_4,-0.2)$	$(h_4,0)$
C_9	$(h_3,0)$	$(h_3,0.2)$	$(h_3,0)$	$(h_3,0)$

TABLE 10. CPCs adjusted evaluation matrix $\hat{R}^1 = (\hat{r}^1_{ij})_{9\times4}$ for task 1.

CPC	B_1	B_2	B_3	B_4
C_1	$(s_5,-0.4)$	$(s_6,0)$	$(s_5,0.2)$	$(s_6,-0.3)$
C_2	$(s_6,0)$	$(s_5,0.3)$	$(s_5,0.3)$	$(s_5,0)$
C_3	$(s_5,-0.2)$	$(s_4,0.4)$	$(s_6,0)$	$(s_5,0.1)$
C_4	$(s_3,0.2)$	$(s_3,0)$	$(s_3,0.1)$	$(s_4,-0.3)$
C_5	$(s_4,0.3)$	$(s_4,0.2)$	$(s_3,0.1)$	$(s_4,-0.3)$
C_6	$(s_3,0.2)$	$(s_4,0)$	$(s_4,-0.2)$	$(s_4,0.3)$
C_7	$(s_4,0.2)$	$(s_5,-0.3)$	$(s_5,0)$	$(s_5,-0.1)$
C_8	$(s_4,0.4)$	$(s_4,0)$	$(s_4,0)$	$(s_4,-0.2)$
C_9	$(s_5,0)$	$(s_6,-0.3)$	$(s_6,0)$	$(s_6,0)$

The experts' subjective weights were known, which are 0.4, 0.3, 0.2 and 0.1 respectively, and get the experts' objective weights by Eq. (19) as 0.2745, 0.3197, 0.1464 and 0.2594. Then determine the experts' combined weights by Eq. (20) as follows:

$$\lambda_1 = 0.337, \lambda_2 = 0.310, \lambda_3 = 0.173, \lambda_4 = 0.180.$$

The gap between the group evaluation result r^k and the last evaluation $\left(r^k\right)'$ is $D(r)$. We define the threshold of $D(r)$ as $\delta = 0.0001$ in this paper. So we need to dynamically adjust the experts' weights and CPCs weights.

After six adjustments $D(r) = 0.000072 < \delta$,stop adjusting, and the definitive CPCs weights and experts' weights are as follows:

$$w_1 = 0.065, w_2 = 0.089, w_3 = 0.075, w_4 = 0.069,$$
$$w_5 = 0.058, w_6 = 0.203, w_7 = 0.161, w_8 = 0.132,$$
$$w_9 = 0.148.\lambda_1 = 0.274, \lambda_2 = 0.321, \lambda_3 = 0.153,$$
$$\lambda_4 = 0.252.$$

Step3: The CPCs evaluation matrix transformed into the degree of belief of corresponding effects.

The adjusted matrix $\hat{\boldsymbol{R}}^1 = (\hat{r}^1_{ij})_{9\times4}$ of task 1 and matrix $\boldsymbol{R}^2 = (r^2_{ij})_{9\times4}$ of task 2 transformed into the degree of belief of corresponding effects are shown in Table 11 and Table 12.

Step4: Determining HEP by CII.

Each experts' degree of belief of nine CPCs effects are combined by ER algorithm. Table 13 shows degree of belief of 4 experts' CPC effects for two tasks.

Combine the data of 4 experts in Table 13 by ER algorithm again, and the results are shown in Table 14.

According to Table 14, we can calculate CII_1 and CII_2 as follows:

$$\text{CII}_1 = 0.8384 - 0.1497 = 0.6887.$$
$$\text{CII}_2 = 0.8885 - 0.1115 = 0.7770.$$

Then HEP is calculated by Eq. (23)

$$\begin{aligned}\text{HEP}_1 &= 7.07 \times 10^{-3} \times \exp[-4.9517 \times 0.6887] \\ &= 2.3355 \times 10^{-4} \\ \text{HEP}_2 &= 7.07 \times 10^{-3} \times \exp[-4.9517 \times 0.7770] \\ &= 1.5083 \times 10^{-4}\end{aligned}$$

So the HEP for Temporary speed restriction of train control system and CTC system control mode conversion is 2.3355×10^{-4} and 1.5083×10^{-4} respectively.

TABLE 11. The degree of belief of corresponding effects for task 1.

CPC	B_1			B_2			B_3			B_4		
	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3	β_1	β_2	β_3
C_1	0.2333	0	0.7667	0.0000	0	1.0000	0.1333	0	0.8667	0.0500	0	0.9500
C_2	0.0000	0	1.0000	0.1167	0	0.8833	0.1167	0	0.8833	0.1667	0	0.8333
C_3	0.2000	0	0.8000	0.2667	0	0.7333	0.0000	0	1.0000	0.1500	0	0.8500
C_4	0.4667	0	0.5333	0.0000	1.000	0.0000	0.4833	0	0.5167	0.3833	0	0.6167
C_5	0.2833	0	0.7167	0.3000	0	0.7000	0.4833	0	0.5167	0.3833	0	0.6167
C_6	0.4667	0	0.5333	0.3333	0	0.6667	0.3667	0	0.6333	0.2833	0	0.7167
C_7	0.3000	0	0.7000	0.2167	0	0.7833	0.1667	0	0.8333	0.1833	0	0.8167
C_8	0.2667	0	0.7333	0.3333	0	0.6667	0.3333	0	0.6667	0.3667	0	0.6333
C_9	0.1667	0	0.8333	0.0500	0	0.9500	0.0000	0	1.0000	0.0000	0	1.0000

TABLE 12. The degree of belief of corresponding effects for task 2.

CPC		B_1			B_2			B_3			B_4	
	β_1	β_2	β_3	β_1	β_1	β_2	β_3	β_1	β_1	β_2	β_3	β_1
C_1	0.2333	0	0.7667	0.0000	0	1.0000	0.1167	0	0.8833	0.0167	0	0.9833
C_2	0.1500	0	0.8500	0.1667	0	0.8333	0.2000	0	0.8000	0.1667	0	0.8333
C_3	0.2000	0	0.8000	0.1000	0	0.9000	0.0000	0	1.0000	0.1667	0	0.8333
C_4	0.4333	0	0.5667	0.3333	0	0.6667	0.3167	0	0.6833	0.3667	0	0.6333
C_5	0.1333	0	0.8667	0.1667	0	0.8333	0.4167	0	0.5833	0.3333	0	0.6667
C_6	0.1667	0	0.8333	0.1333	0	0.8667	0.2167	0	0.7833	0.1500	0	0.8500
C_7	0.2667	0	0.7333	0.1667	0	0.8333	0.1667	0	0.8333	0.2167	0	0.7833
C_8	0.2667	0	0.7333	0.3333	0	0.6667	0.3333	0	0.6667	0.3667	0	0.6333
C_9	0.1667	0	0.8333	0.0500	0	0.9500	0.0000	0	1.0000	0.0000	0	1.0000

TABLE 13. Degree of belief of 4 experts' CPC effects for two tasks.

Task		B_1			B_2			B_3			B_4	
type	β_1	β_2	β_3	β_1	β_1	β_2	β_3	β_1	β_1	β_2	β_3	β_1
Task 1	0.2406	0	0.7594	0.1678	0.0478	0.7844	0.1802	0	0.8198	0.1679	0	0.8321
Task 2	0.1718	0	0.8282	0.1182	0	0.8818	0.1416	0	0.8584	0.1415	0	0.8585

TABLE 14. The definitive degree of belief for two tasks.

Task	Degree of belief		
type	β_1	β_2	β_3
Task 1	0.1497	0.0119	0.8384
Task 2	0.1115	0	0.8885

C. COMPARED WITH THE QUANTIFICATION RESULTS OF THE OTHER CREAM METHODS

In order to verify the model calculation results validation, the results of this model are compared with those of literature [5] and literature [9].

Since neither of the two literatures uses MAGDM method, it is necessary to combine the data of four experts in Table 11 and table 12 respectively, which is different from step 4 in part B of this section. Each CPC Effects on human reliability are shown in Table 15.

For comparison purposes, given the same CPCs weights as this model, the data in Table 16 are used to calculate the HEP by the model in literature [5] and literature [9] respectively. The calculation results are shown in Table 16.

Obviously, due to different calculation methods, the difference of the results by three approaches seems to be inevitable, the model in this paper is a bit conservative. Since the literature [9] is also calculated based on the ER algorithm, compared with the literature [5], the results are closer to the results calculated in this paper. However, a conservative HEP is not unacceptable for reliability assessment of safety-critical systems.

It shows the validation of the model in this paper. However, literature [5] and literature [9] need to establish the membership function and a large number of fuzzy rules

TABLE 15. The degree of belief of nine CPCs for two tasks.

CPC	Task 1	Task2
	Effects on human reliability	
C_1	5.72% Negative	6.55% Negative
	94.28% Positive	93.45% Positive
C_2	12.71% Negative	6.77% Negative
	87.29% Positive	93.23% Positive
C_3	9.50% Negative	14.07% Negative
	90.50% Positive	85.93% Positive
C_4	33.77% Negative	29.25% Negative
	66.23% Positive	31.45% Neutral
		39.30% Positive
C_5	19.10% Negative	30.92% Negative
	80.90% Positive	69.08% Positive
C_6	11.94% Negative	33.23% Negative
	88.06% Positive	66.77% Positive
C_7	16.64% Negative	18.17% Negative
	83.36% Positive	81.83% Positive
C_8	28.75% Negative	28.75% Negative
	71.25% Positive	71.25% Positive
C_9	4.02% Negative	4.02% Negative
	95.98% Positive	95.98% Positive

(The number of rules in literature [5] and literature [9] are 23328 and 46656 respectively). The establishment of membership function requires a lot of reliable data. And the

TABLE 16. HEP for two tasks by different approaches.

Comparative literature	HEP	
	Task 1	Task2
literature[5]	2.0286×10^{-4}	1.3293×10^{-4}
literature[9]	2.3045×10^{-4}	1.5007×10^{-4}
This paper	2.3355×10^{-4}	1.5083×10^{-4}

methods for calculating CPCs weights in both literatures are too subjective. These weakens the applicability of these two approaches.

D. MODEL RATIONALITY VERIFICATION

1) THE MODEL RATIONALITY ANALYSIS

In this section, to verify the model rationality, we introduce the following two axioms [5].

Axiom 1: A slight positive/negative change in the status of the CPCs would definitely result in the decrement/increment HEP, but it should not cause the mutation of HEP.

Axiom 2: Given the same observations of the CPCs, the input variable with higher degrees of importance would certainly give rise to a significant influence on the magnitude of HEP.

Take the evaluation of task 1 by expert 1 as an example, firstly, keep the evaluation of C_2 to C_9 unchanged, and the evaluation of C_1 increases from $(s_0, 0)$ to $(s_6, 0)$ at a step of 0.1, then keep the evaluation of C_1 to C_8 unchanged, and the evaluation of C_9 also increases from $(s_0, 0)$ to $(s_6, 0)$ at a step of 0.1.

It can be seen from Table 17 and Table 18 that a slight positive/negative change in the status of the CPCs would definitely result in the decrement/increment HEP. From Figure 5,

TABLE 17. HEP with increasing C_1 evaluation.

C_1	$(s_0,0)$	$(s_0,0.1)$	$(s_0,0.2)$	$(s_0,0.3)$	$(s_0,0.4)$	$(s_1,-0.5)$	$(s_1,-0.4)$	$(s_1,-0.3)$	$(s_1,-0.2)$	$(s_1,-0.1)$	—
HEP ($\times10^{-4}$)	4.4215	4.3736	4.3262	4.2794	4.2372	4.1913	4.1459	4.1009	4.0605	4.0165	—
C_1	$(s_1,0)$	$(s_1,0.1)$	$(s_1,0.2)$	$(s_1,0.3)$	$(s_1,0.4)$	$(s_2,-0.5)$	$(s_2,-0.4)$	$(s_2,-0.3)$	$(s_2,-0.2)$	$(s_2,-0.1)$	—
HEP ($\times10^{-4}$)	3.9730	3.9339	3.8912	3.8491	3.8112	3.7699	3.729	3.6923	3.6523	3.6163	—
C_1	$(s_2,0)$	$(s_2,0.1)$	$(s_2,0.2)$	$(s_2,0.3)$	$(s_2,0.4)$	$(s_3,-0.5)$	$(s_3,-0.4)$	$(s_3,-0.3)$	$(s_3,-0.2)$	$(s_3,-0.1)$	—
HEP ($\times10^{-4}$)	3.5771	3.5418	3.5035	3.4689	3.4314	3.3975	3.3641	3.3276	3.2916	3.2591	—
C_1	$(s_3,0)$	$(s_3,0.1)$	$(s_3,0.2)$	$(s_3,0.3)$	$(s_3,0.4)$	$(s_4,-0.5)$	$(s_4,-0.4)$	$(s_4,-0.3)$	$(s_4,-0.2)$	$(s_4,-0.1)$	—
HEP ($\times10^{-4}$)	3.2414	3.1920	3.1606	3.1263	3.0955	3.065	3.0348	3.0019	2.9724	2.9431	—
C_1	$(s_4,0)$	$(s_4,0.1)$	$(s_4,0.2)$	$(s_4,0.3)$	$(s_4,0.4)$	$(s_5,-0.5)$	$(s_5,-0.4)$	$(s_5,-0.3)$	$(s_5,-0.2)$	$(s_5,-0.1)$	—
HEP ($\times10^{-4}$)	2.9112	2.8825	2.8541	2.826	2.7981	2.7678	2.7405	2.7135	2.6868	2.6603	—
C_1	$(s_5,0)$	$(s_5,0.1)$	$(s_5,0.2)$	$(s_5,0.3)$	$(s_5,0.4)$	$(s_6,-0.5)$	$(s_6,-0.4)$	$(s_6,-0.3)$	$(s_6,-0.2)$	$(s_6,-0.1)$	$(s_6,0)$
HEP ($\times10^{-4}$)	2.6341	2.6081	2.5799	2.5544	2.5293	2.5043	2.4797	2.4552	2.4310	2.4071	2.3834

TABLE 18. HEP with increasing C_9 evaluation.

C_9	$(s_0,0)$	$(s_0,0.1)$	$(s_0,0.2)$	$(s_0,0.3)$	$(s_0,0.4)$	$(s_1,-0.5)$	$(s_1,-0.4)$	$(s_1,-0.3)$	$(s_1,-0.2)$	$(s_1,-0.1)$	—
HEP ($\times10^{-4}$)	10.1893	9.9107	9.6301	9.3667	9.1015	8.8526	8.6105	8.3750	8.1460	7.9232	—
C_9	$(s_1,0)$	$(s_1,0.1)$	$(s_1,0.2)$	$(s_1,0.3)$	$(s_1,0.4)$	$(s_2,-0.5)$	$(s_2,-0.4)$	$(s_2,-0.3)$	$(s_2,-0.2)$	$(s_2,-0.1)$	—
HEP ($\times10^{-4}$)	7.7141	7.5032	7.2980	7.1054	6.9179	6.7287	6.5512	6.3784	6.2101	6.0462	—
C_9	$(s_2,0)$	$(s_2,0.1)$	$(s_2,0.2)$	$(s_2,0.3)$	$(s_2,0.4)$	$(s_3,-0.5)$	$(s_3,-0.4)$	$(s_3,-0.3)$	$(s_3,-0.2)$	$(s_3,-0.1)$	—
HEP ($\times10^{-4}$)	5.8867	5.7314	5.5801	5.4383	5.2948	5.1551	5.0241	4.8963	4.7672	4.6483	—
C_9	$(s_3,0)$	$(s_3,0.1)$	$(s_3,0.2)$	$(s_3,0.3)$	$(s_3,0.4)$	$(s_4,-0.5)$	$(s_4,-0.4)$	$(s_4,-0.3)$	$(s_4,-0.2)$	$(s_4,-0.1)$	—
HEP ($\times10^{-4}$)	4.6460	4.4128	4.3006	4.1913	4.0847	3.9848	3.8835	3.7848	3.6923	3.5984	—
C_9	$(s_4,0)$	$(s_4,0.1)$	$(s_4,0.2)$	$(s_4,0.3)$	$(s_4,0.4)$	$(s_5,-0.5)$	$(s_5,-0.4)$	$(s_5,-0.3)$	$(s_5,-0.2)$	$(s_5,-0.1)$	—
HEP ($\times10^{-4}$)	3.5104	3.4246	3.3375	3.2559	3.1763	3.0986	3.0228	2.9489	2.8796	2.8092	—
C_9	$(s_5,0)$	$(s_5,0.1)$	$(s_5,0.2)$	$(s_5,0.3)$	$(s_5,0.4)$	$(s_6,-0.5)$	$(s_6,-0.4)$	$(s_6,-0.3)$	$(s_6,-0.2)$	$(s_6,-0.1)$	$(s_6,0)$
HEP ($\times10^{-4}$)	2.7405	2.6762	2.6107	2.5494	2.4870	2.4286	2.3716	2.3159	2.3022	2.2084	2.1565

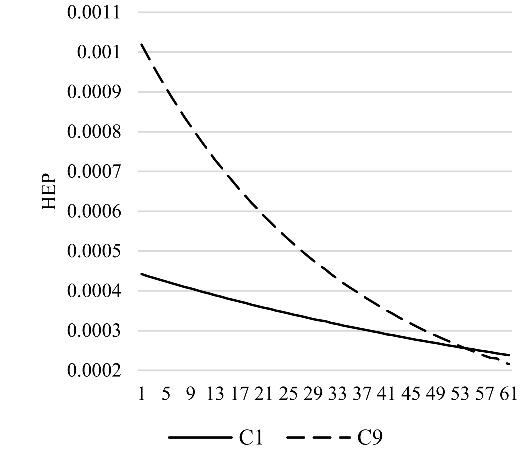

FIGURE 5. The HEP trend lines of two scenarios.

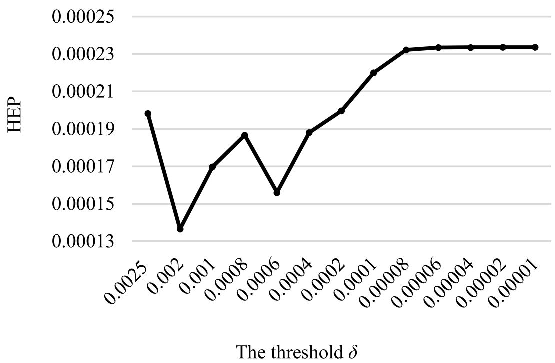

FIGURE 6. The HEP with different thresholds.

we can see that the change of the HEP trend line is smooth without any mutation. Therefore, it is reasonable to judge that the logicality of the proposed modified CREAM is validated. It can also be seen that C_9 has a significant influence on HEP than C_1(the weight of C_9 is higher than that of C_1), this is consistent with the principle of Axiom 2.. Therefore, the rationality and logicality of the model are verified.

2) THE MODEL SENSITIVITY ANALYSIS

It can be seen from part C and D of section III that the gap between group evaluation results $D(r)$ has an important impact on the weight of CPC and experts.

To test the influence of threshold δ for $D(r)$ setting on HEP, a sensitivity analysis is conducted by changing the threshold δ values. The sensitivity analysis results according to different threshold δ is shown as Fig. 6. (Take task 1 as an example). Since the initial $D(r)$ is 0.002614, the abscissa δ starts at 0.0025.

As it can be seen from Figure.6, the HEP indeed influenced by changing the different δ values, but when $\delta < 0.0001$, the impact on HEP becomes less significant. This is because when experts do not reach a certain consensus, it has a great impact on the attributes (i.e. CPCs) and experts' weights, which in turn affects the calculation of HEP. Therefore, when calculating HEP, the impact of weight must be considered. The method of dynamically adjusting weights proposed in this paper takes into account the role of experts' opinions and CPCs information in determining the weights, making full use of the advantages of MAGDM, making HEP evaluation closer to the reality.

V. CONCLUSION

This paper proposed a modified CREAM for human reliability in high-speed railway dispatching tasks. Due to Lack of reliable historical data, most CREAM studies dependent on the domain knowledge and experiences of the HRA analyzers and experts. Consequently, the HEP calculation result is a bit subjective inevitably. Different from other previous CREAM based HEP quantification studies, the new method has some unique and significant characteristics as follows:

(1) 2-tuple linguistic term sets can well characterize the fuzziness and uncertainty of CPCs information, then 2-tuple linguistic is transformed into degree of belief in a simple way. The degree of belief approach, instead of a deterministic one-or-zero way in specification of CPCs, can be used to well model the uncertainty and be more practical when adopting experts' judgments.

(2) Considering both subjective and objective weights of CPCs and experts by the dynamic adjustment method, which provides a prerequisite for the application of the ER algorithm.

(3) The ER algorithm with the degree of belief can deal with the uncertainties caused by insufficient information and data in the evaluation of CPCs

(4) Since the application of the MAGDM method, the evaluation of the expert group is crucial to the result of analysis. The multiple experts avoid bias that may be presented considering the subjective judgments of a single expert. Accordingly, the MAGDM CREAM model is able to produce reliable HEP results

Based on the existing research, the proposed model overcomes the shortcomings caused by the lack of data. Therefore, the model can not only use the HRA field of high-speed railways, but also applicable to other fields. Furthermore, the extended method of CREAM is more valuable but complex than the basic method. In the future, we will try to do some modifications on the extended method.

REFERENCES

[1] M. Kyriakidis, K. T. Pak, and A. Majumdar, "Railway accidents caused by human error: Historic analysis of UK railways, 1945 to 2012," *Transp. Res. record.*, vol. 2476, no. 1, pp. 126–136, Jan.2015.

[2] M. A. Sujan, D. Embrey, and H. Huang, "On the application of human reliability analysis in healthcare: Opportunities and challenges," *Rel. Eng. Syst. Saf.*, vol. 194, Feb. 2020, Art. no. 106189.

[3] X. Deng and W. Jiang, "Dependence assessment in human reliability analysis using an evidential network approach extended by belief rules and uncertainty measures," *Ann. Nucl. Energy*, vol. 117, pp. 183–193, Jul. 2018.

[4] N. J. Ekanem, A. Mosleh, and S.-H. Shen, "Phoenix—A model-based human reliability analysis methodology: Qualitative analysis procedure," *Rel. Eng. Syst. Saf.*, vol. 145, pp. 301–315, Jan. 2016.

[5] S.-T. Ung, "A weighted CREAM model for maritime human reliability analysis," *Saf. Sci.*, vol. 72, pp. 144–152, Feb. 2015.

[6] J.-L. Zhou, Y. Lei, and Y. Chen, "A hybrid HEART method to estimate human error probabilities in locomotive driving process," *Rel. Eng. Syst. Saf.*, vol. 188, pp. 80–89, Aug. 2019.

[7] A. D. Swain and H. E. Guttmann, "Handbook of human-reliability analysis with emphasis on nuclear power plant applications," Sandia Nat. Labs., Albuquerque, NM, USA, Tech. Rep. NUREG/CR-1278, Aug. 1983.
[8] E. Hollnagel, "Cream a second generation HRA method," in *Cognitive Reliability and Error Analysis Method*, 1st ed. Oxford, U.K.: Elsevier, 1988, pp. 151–184.
[9] Z. L. Yang, S. Bonsall, A. Wall, J. Wang, and M. Usman, "A modified CREAM to human reliability quantification in marine engineering," *Ocean Eng.*, vol. 58, pp. 293–303, Jan. 2013.
[10] X. He, Y. Wang, Z. Shen, and X. Huang, "A simplified CREAM prospective quantification process and its application," *Rel. Eng. Syst. Saf.*, vol. 93, no. 2, pp. 298–306, Feb. 2008.
[11] Z. Sun, Z. Li, E. Gong, and H. Xie, "Estimating human error probability using a modified CREAM," *Rel. Eng. Syst. Saf.*, vol. 100, pp. 28–32, Apr. 2012.
[12] M. Konstandinidou, Z. Nivolianitou, C. Kiranoudis, and N. Markatos, "A fuzzy modeling application of CREAM methodology for human reliability analysis," *Rel. Eng. Syst. Saf.*, vol. 91, no. 6, pp. 706–716, Jun. 2006.
[13] Z. Nivolianitou and M. Konstantinidou, "A fuzzy modeling application for human reliability analysis in the process industry," in *Human Factors and Reliability Engineering for Safety and Security in Critical Infrastructures*. Cham, Switzerland: Springer, 2018, pp. 109–154.
[14] G. A. Shirali, T. Hosseinzadeh, K. A. Angali, and S. R. N. Kalhori, "Modifying a method for human reliability assessment based on CREAM-BN: A case study in control room of a petrochemical plant," *MethodsX*, vol. 6, pp. 300–315, Jan. 2019.
[15] Q. Zhou, Y. D. Wong, H. S. Loh, and K. F. Yuen, "A fuzzy and Bayesian network CREAM model for human reliability analysis—The case of tanker shipping," *Saf. Sci.*, vol. 105, pp. 149–157, Jun. 2018.
[16] S.-T. Ung, "Evaluation of human error contribution to oil tanker collision using fault tree analysis and modified fuzzy Bayesian network based CREAM," *Ocean Eng.*, vol. 179, pp. 159–172, May 2019.
[17] L. Martinez and F. Herrera, "A 2-tuple fuzzy linguistic representation model for computing with words," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 6, pp. 746–752, Dec. 2000.
[18] J.-B. Yang and M. G. Singh, "An evidential reasoning approach for multiple-attribute decision making with uncertainty," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 1, pp. 1–18, Jan. 1994.
[19] J.-B. Yang and D.-L. Xu, "On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 32, no. 3, pp. 289–304, May 2002.
[20] J. B. Yang, Y. M. Wang, D. L. Xu, and K. S. Chin, "The evidential reasoning approach for MADA under both probabilistic and fuzzy uncertainties," *Eur. J. Oper. Res.*, vol. 171, no. 1, pp. 309–343, May 2006.
[21] D.-L. Mon, C.-H. Cheng, and J.-C. Lin, "Evaluating weapon system using fuzzy analytic hierarchy process based on entropy weight," *Fuzzy Sets Syst.*, vol. 62, no. 2, pp. 127–134, Mar. 1994.
[22] F. Jin, L. Pei, H. Chen, and L. Zhou, "Interval-valued intuitionistic fuzzy continuous weighted entropy and its application to multi-criteria fuzzy group decision making," *Knowl.-Based Syst.*, vol. 59, pp. 132–141, Mar. 2014.
[23] W. Zhang, Y. Xu, and H. Wang, "A consensus reaching model for 2-tuple linguistic multiple attribute group decision making with incomplete weight information," *Int. J. Syst. Sci.*, vol. 47, no. 2, pp. 389–405, Jan. 2016.
[24] Y. Fujita and E. Hollnagel, "Failures without errors: Quantification of context in HRA," *Rel. Eng. Syst. Saf.*, vol. 83, no. 2, pp. 145–151, Feb. 2004.
[25] G. Apostolakis, V. Bier, and A. Mosleh, "A critique of recent models for human error rate assessment," *Rel. Eng. Syst. Saf.*, vol. 22, nos. 1–4, pp. 201–217, Dec. 1988.
[26] Y. T. Xi, Z. L. Yang, Q. G. Fang, W. J. Chen, and J. Wang, "A new hybrid approach to human error probability quantification–applications in maritime operations," *Ocean Eng.*, vol. 138, pp. 45–54, Jul. 2017.

YANHAO SUN received the B.S. degree in traffic engineering from the Shandong University of Technology, Zibo, China, in 2012, and the M.S degree in traffic and transportation engineering from the China Academy of Railway Sciences, Beijing, China, in 2017, where he is currently pursuing the Ph.D. degree in traffic information engineering and control. His research interests include railway signal and communication, railway traffic modeling and simulation, railway intelligent control and dispatching, and railway transportation safety.

QI ZHANG received the B.S. degree in electronic engineering from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1991, and the M.S degree in computer application and the Ph.D. degree in traffic information engineering and control from the China Academy of Railway Sciences, Beijing, in 1993 and 1998, respectively. He is currently a Chief Researcher with the China Academy of Railway Sciences Corporation Limited and the Leader of railway technical expertise. His research interests include railway signal and communication, automatic train operation, train operation control, intelligent dispatching, and cooperative control of multiple trains. He has directed many national scientific projects, in China and has a lot of achievements. One of his research achievements, the Centralized Traffic Control (CTC) System, has become one of the most important technical equipment in China railway.

ZHIMING YUAN received the Ph.D. degree from the China Academy of Railway Sciences, Beijing, China, in 2016. From March 2017 to March 2018, he worked as a Researcher with North Carolina State University. He is currently a Professor with the Communication and Signal Research Institute, China Academy of Railway Sciences Corporation Limited. His research interests include railway traffic modeling and simulation, railway operations research, and railway traffic control.

YING GAO received the B.S. degree in electrical engineering and automation from Qingdao Agriculture University, Qingdao, China, in 2008, the M.S degree in detection technology and automation from Beijing Jiaotong University, Beijing, in 2010, and the Ph.D. degrees in traffic information engineering and control from the China Academy of Railway Sciences, in 2019. She is currently a Safety Assessment Engineer with the China Academy of Railway Sciences Corporation Limited and a Product Certification Engineer in railway signal equipment. Her research interests include railway signal and train operation control. She has participated in many national scientific projects, in China. Her achievements, in railway signal equipment safety assessment and safety audit, have provided important support for the safety and security of China's high-speed railway.

SHUXIN DING received the B.E. degree in automation and the Ph.D. degree in control science and engineering from the Beijing Institute of Technology, Beijing, China, in 2012 and 2019, respectively. From 2016 to 2017, he was a Visiting Scholar of industrial and systems engineering with the University of Florida, Gainesville, FL, USA. He is currently an Assistant Researcher with the Communication and Signal Research Institute, China Academy of Railway Sciences Corporation Limited. His current research interests include evolutionary computation, multiobjective optimization, optimization under uncertainty, and railway scheduling.