Reliability Engineering and System Safety 234 (2023) 109136

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

A railway accident prevention method based on reinforcement learning – Active preventive strategy by multi-modal data

Dongyang Yan [a], Keping Li [a,*], Qiaozhen Zhu [b], Yanyan Liu [a]

[a] State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China
[b] Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing 100044, China

ARTICLE INFO

Keywords:
Railway
Accident prediction
Artificial intelligence
Railway safety
Text data

ABSTRACT

Railway systems are entering an era of highly intelligent automation where stability and safety are becoming increasingly important. However, there is still a lack of intelligent and effective ways for railway accident prevention, especially active accident prevention strategies. This paper presents a railway accident prevention method based on the reinforcement learning model and multi-modal data to achieve active railway accident prevention strategies. Three metrics are designed to show the performance of active prevention methods. Based on the three metrics and the data from Federal Railroad Administration, the effectiveness of the proposed method is verified in the case study by introducing two methods as baselines. The results also show that nearly 30% of accidents can be effectively prevented through active preventive measures with the proposed method. Finally, this paper analyzes the influence of personal skills on the proposed model and makes relevant suggestions for improving railway safety based on the analysis of the results.

1. Introduction

Safety is a long-lasting topic for railway systems. The occurrence of railway accidents may cause substantial property losses and casualties. According to reports from the Federal Railroad Administration (FRA), the total number of accidents/incidents has reached 92,000 in the past ten years (2012–2021), resulting in 60,948 injuries and 6486 fatalities.[1] Meanwhile, the total number of accidents/incidents was 9215 in 2012 and 8035 in 2021(9745 in 2019), while the total fatalities were 558 in 2012 and 766 in 2021. Although efforts [1–3] over the past decade have achieved certain results, the number of accidents is still at a high level. Traditional accident prevention strategies have reached a bottleneck, and a more active strategy is needed to prevent the happening of railway accidents.

Active accident prevention is to recognize and eliminate the factors that will trigger accidents in advance through automatic risk detection technologies and effective responses. For example, through vision sensors, some possible conflicts can be detected and used to warn passengers to prevent collision risks [4]. Conversely, passive prevention refers to the methods that examine as many accident factors as possible in the hope of ruling out future factors which may have caused accidents [5]. Recognition of possible hidden risks of accidents is the main characteristic of active prevention [6]. Moreover, active accident prevention is a combination of accident prediction, causal analysis, and accident data analysis.

In railway systems, the implementation of active accident prevention suffers from the challenges of accident detection and the limitations of insufficient pre-accident response. First of all, the response to prevent accidents needs a risk predictor with high reliability to ensure that the response is valid. So, risk prediction is required to be as accurate and specific as possible. However, many factors, including weather, geography, equipment, train condition, passenger behavior, and even unobservable factors, could be the trigger for a railway accident, while many prediction models can only consider part of them because it is hard to get the information of all the factors. Secondly, the pre-accident response, which means the response to prevent an accident after accident risk is detected, requires specific information about the future accident. This means that only knowing whether an accident will happen is not enough, and the cause of this accident is also needed. Therefore, many accident prediction methods cannot satisfy the need for active accident prevention because they only predict the occurrence or frequency of accidents [7–9]. Moreover, response time should also be

* Corresponding author.
E-mail address: kpli@bjtu.edu.cn (K. Li).

[1] https://railroads.dot.gov/safety-data/accident-and-incident-reporting/accidentincident-dashboards-data-downloads

https://doi.org/10.1016/j.ress.2023.109136
Received 23 September 2022; Received in revised form 17 January 2023; Accepted 1 February 2023
Available online 4 February 2023
0951-8320/

considered an important factor. Some responses for railway accidents may need a long time, and thus the prediction model should ensure the data can be accessed in advance to reserve enough time for active accident prevention. To sum up, successful implementation of active accident prevention for railway systems needs to address the following issues: an accurate prediction of accident occurrence, a specific prediction of accident cause, and full use of multi-source data to reserve enough time for pre-accident response.

This paper presents a reinforcement learning-based method to predict the causes of railway operational accidents along with time flows to achieve active accident prevention for railway systems by addressing the above issues. The main idea of this method is to achieve a dynamic scenario-response mode to anticipate the cause of a possible accident and take the corresponding preventive measures in advance to take the initiative to prevent accidents. The proposed model is mainly based on the historical information of accidents and related characteristics, so data for prediction are within broader time constraints, ensuring that possible accident information can be learned in advance and corresponding prevention and control can be carried out. Moreover, the multi-modal data, including structured data and unstructured data (text), are combined through embedding vectors and natural language processing (NLP) technologies to achieve full use of the information contained in the accident records.

The main contributions of this paper are as follows:

(1) A reinforcement learning (RL) method is introduced to achieve railway active accident prevention. The RL method achieves an effective way to learn useful information from accident records and predict the cause of the coming accident ahead of a period, allowing the staff to make a good preparation for accident prevention. Three metrics are designed to verify the performance of the proposed method.
(2) Multi-modal data, including text and numerical data, is organized for accident prediction, representing categorical data by embedding vectors.
(3) Randomness is introduced to be an indicator denoting the influence of personal skills to make the proposed method aware of the impact of human factors. Considering that accident prevention highly depends on the experiences of employees, the randomness makes the proposed method able to reflect the influence of personal skill level. In the experimental section, this paper presents the results from the different levels of randomness and advises some measures for improving personal skills.

The rest of this paper is organized as follows. The implementation of active accident prevention requires the combination of accident prediction, accident cause analysis, and full use of railway accident data. So, Section 2 reviews the related works, including railway accident prediction, cause analysis, and data analysis in railway safety. Moreover, RL-based methods that are applied to railway systems are also reviewed to shed light on the advance of RL-based methods in railways. In Section 3, the data used in this paper is described from point of the motivation of this research. Section 4 presents the proposed reinforcement learning-based method. Section 5 shows the experimental results based on the operational accident data. Finally, some discussions and conclusions are drawn in Section 6 and Section 7.

2. Literature review

2.1. Railway accident prediction

Railway accident prediction is a technology to predict whether or not a railway accident will happen or how many railway accidents may occur in a particular future period at a certain place. Specifically, if the model does not estimate the number of accidents but predicts whether accidents will happen, we can call it risk prediction [1].

The highway-rail grade crossing (HRGC) is a concern in railway accident prediction. Accident prediction for HRGC hopes to explore the place with the most risk, prioritize prevention and control, and reasonably allocate the limited budget expenditure. A review and assessment of related prediction models can be accessed in Abioye's study [1] and Pasha's study [8]. The commonly used models are binary logit models, multinomial models, negative binomial models, and other generalized linear modeling models [2]. Some other models, like the zero-inflated negative binomial model [7], are proposed to cope with the problem that crash data contain a preponderance of zeros. Machine learning and deep learning methods are introduced to develop the technologies of railway accident prediction in a model-free approach. Based on these methods, like random forest algorithm, decision tree, K-nearest neighbors, naïve Bayes classifier, and convolution neural network (CNN), researchers can borrow improvements from machine learning in other fields to further improve the prediction accuracy in railway accident prediction. Generally, these machine learning methods show fluctuations in performance for specific tasks, and it is necessary to test different methods before deciding which one to use [2,3].

Data availability and data accuracy are also essential for prediction accuracy. Most prediction models are based on certain characteristics of HRGC, e.g., traffic volume, train volume, train speed, vehicle speed, and the number of workers on duty. Human factors [10], and related factors like pavement markings in the form of strop lines [9], play an important role in accidents happening. Singh et al. [11] also analyze the influence of spatial, temporal, and environmental characteristics. It can be observed that in many research, time constraints are not considered for their studies because they aim at long-term accident prediction. If active accident prevention is required, some factors, like vehicle speed, number of workers, and traffic volume, may not be available because active accident response may be implemented before the time when these factors are not detected. In this paper, the accessibility of data will also be our consideration during data selection.

In recent years, the focus on using text data for railway accident prediction is raising. Because the text data is unstructured, the method to use this kind of data shows many differences from the numerical data. Before analyzing, the text data should be first represented in vectors and clustered to get its valid information [12]. Furthermore, the technologies used in text data mining, e.g., the long-short-term memory (LSTM) model, inspire researchers to study accident prevention in a text data mining view based on a context ontology of railway accidents [13]. The use of text data is also highlighted in this paper, and the way to use both numerical data and text data for railway accident prediction is explored.

2.2. Railway accident cause analysis

There is an overlap between the method of causation analysis and accident prediction because some prediction models, e.g., the binary logit model, can predict the accident hazard as well as estimate the risk degree of each contributing factor. In the past decades, lots of studies based on GLMs have been implemented to discover different contributing factors to railway accidents. It is hard to use a unified framework to conclude the main factors that may cause accidents because of the complex aspects of railway systems. Specifically, the factors related to pedestrian behavior play key roles in accidents around the highway-rail crossing and riding time, the number of stops a passenger passes, transgressions, and abrupt violations are the main contributors to accidents [14]. Even the non-crossing places should also be cautious of the railway accidents caused by pedestrian trespassing [15], which shows a spatial variation. As for the railroad tracks, the load of the train also affects the occurrence and severity of an accident [16]. Kyriakidis et al. [17] select 27 precursory factors of the metro and divide them into six categories. The relationship between each precursory factor and accident injury is analyzed statistically by questionnaire. Some suggestions are put forward to improve the safety level. In Xing's study [18], the authors analyze the cases of escalator-related pedestrian injuries in the

Table 1
The possibility of obtaining the features related to a certain accident at a certain time.

Factors	Time until the accident happens				
	Days before	Hours before	Just happens	Hours after	Days after
Time	possible	possible	possible	possible	possible
Place	possible	possible	possible	possible	possible
CIT	impossible	impossible	possible	possible	possible
CN	impossible	impossible	possible	possible	possible
CS	impossible	impossible	possible	possible	possible
DA	impossible	impossible	impossible	impossible	possible

Note: CIT-condition of influenced trains; CN-condition of nature; CS-condition of staff; DA-details of accidents.

Guangzhou metro from 2013 to 2015 and find out the differences in contributing factors and crowd characteristics. Some works analyze the cause of railway accidents from a holistic view, which explores the evolution process of accident causes at different levels in a qualitative way based on STAMP [19] and AcciMap [20].

The complex network is commonly used in the cause analysis of railway accidents. Complex network represents the elements related to accidents as nodes and connects the nodes with edges if they contain certain relationships [21]. It is a way to understand the systems and interactions of components from a graph-based view. The network attribute is an important metric to determine the influence of causes. Degree, shortest paths, and betweenness are the key attributes used to rank the importance of nodes [22,23]. Discovering the key causes of accidents can effectively monitor railway safety. The above work is based on an undirected complex network, and the nodes in the network have the same properties. Liu et al. [5] introduce the directed heterogeneous network to consider a variety of relationships between different entities related to railway operational accidents.

From the above works, we can conclude that the railway accident cause analysis aims to find the main causes and give suggestions for accident prevention by eliminating the occurrence of the main causes. However, the main causes cannot cover all the accidents, and also, the strategies for eliminating the main causes may gradually lose influence because these causes will gradually change with the intervention of human factors and the development of the railway system. This paper attempts to use all the different causes in a new dynamic manner for accident prevention. The causes that should be eliminated are evaluated based on the accident history and the most likely causes to be prevented currently are extracted. Thus, the prevention strategy is a dynamic-scenario-coping strategy that can evolve with human intervention and railway system development.

2.3. Data analysis on railway safety

Commonly used data in railway accident prediction and prevention is structured numerical data because most of the accident data is recorded as row data. In recent years, researchers have started to explore the contribution of unstructured data like text, pictures, and videos. As discussed in Section 2.1, text data is now more and more considered in railway safety. Using text data as a part of railway safety analysis has its own advantages because many accident records contain a narrative part to explain and expand the details of the accidents, and so text data is easier to access than other unstructured data [24]. Text data mining technologies are used first before text data is analyzed further. Generally, the information contained in text data shows a sparsity characteristic, which is highly different from the numerical data. The effectiveness of text data mining will influence the results of safety analysis. Pre-processing operations for texts, like lemmatization, removing punctuations and stopwords, and theme clustering are needed to facilitate the machine to understand texts. After that, texts are further categorized [25] and vectorized [26] to measure the degree of similarity between the semantic meanings of texts. Some studies go further about the use of texts by transferring the incident records into a stream of text formats and then analyzing them with text information extraction methods [27]. Besides, the short text-based streaming media platform will also more and more play an important role in accident warnings with the help of text analysis technology [28]. A review of the data-driven models and methods for high-speed train fault diagnosis is presented in Chen's study [29].

With the development of text mining, the representation of text also becomes more efficient. The above research is all based on the representation of text with BoW models [30]. In this paper, we introduce the word embedding method for the effective representation of text, as well as categorical data.

2.4. Application of reinforcement learning to railway system

In 2015, the AlphaGo designed by DeepMind beat human players and won the championship, pushing RL into a stage of prosperity and development in the next decades. For now, RL has been applied in various engineering fields, including gas supply optimization [31], power allocation [32], production scheduling [33], system maintenance [34–36], train scheduling [37], rescheduling [38], and passenger flow control [39].

The railway system consists of multiple sub-systems, among which the subsystems for railway power and wireless network are put much attention regarding the use of RL. For railway power, RL can be applied to improve the reliability of pantograph-catenary contact force (PCCF) [40], to get a better charge and discharge policy [41,42], as well as to get a better energy management strategy to save energy consumption and operational cost [43,44]. The application in wireless networks also highlights the need for energy-saving control, and based on this demand, researchers use RL-based methods to achieve balanced power allocation and reduce the total computational cost [45,46]. Facing the high unpredictability of high-speed railway (HSR) networks, RL shows its tremendous potential in improving HSR network quality by reliving the impact of receiving buffer blocking problems and handover delay [47–49]. In railway planning management, researchers take efforts to improve the strategies of site selection [50], train scheduling [51], and timetable rescheduling [38].

Though lots of successful applications of RL in railways are reported, the number of research on using RL for railway accident prevention is still limited. So, it is significant to explore the way to improve the safety level by introducing RL to railway accident prevention, which is one of the motivations of this paper.

3. Data description

This paper uses the accident report from FRA to develop the RL-based active accident prevention method. FRA investigates and records the accidents/incidents meeting specific criteria.[2] The records of accidents/incidents can be accessed on its *FRA Safety Data Site*. These data are categorized into four classes based on the accident-related attributes, including rail equipment accident/incident (REAI) data, injury/illness summary–operational (OP) data, injury/illness summary–casualty (CA) data, and highway-rail grade crossing accident/incident (HRGC) data. The case study is based on the REAI data from 2015 to 2019.

3.1. Data selection

REAI data contains a series of detailed factors related to different types of railway equipment accidents, including derailment, head-on

[2] https://railroads.dot.gov/railroad-safety/accident-data-reporting-and-investigations

Table 2
Factors used in this paper.

Format	Factor	Description
Numerical data	TEMP	Temperature (in degrees Fahrenheit)
	TONS	Gross tonnage (excluding power units)
	TRNSPD	Train speed (in miles per hour)
	HIGHSPD	Maximum train speed (in miles per hour)
	HRGCFREQ	Highway-rail crossing accident frequency during a period at a certain place (COOK, JEFFERSON, HARRIS)
	CANUM	Number of casualties during a period at a certain place (COOK, JEFFERSON, HARRIS)
Categorical data	VISIBLTY	Daylight period (4 classes)
	WEATHER	Weather condition (6 classes)
	TYPEQ	Type of consist (14 classes)
	TYPTRK	Type of track (4 classes)
	CAUSE	Primary cause of accident/incident (389 classes/5 titles/28 categories)
Narrative data	NARRS	Data with text format to describe the details of accidents. NARRS is the concatenation of NARR1 – NARR15.

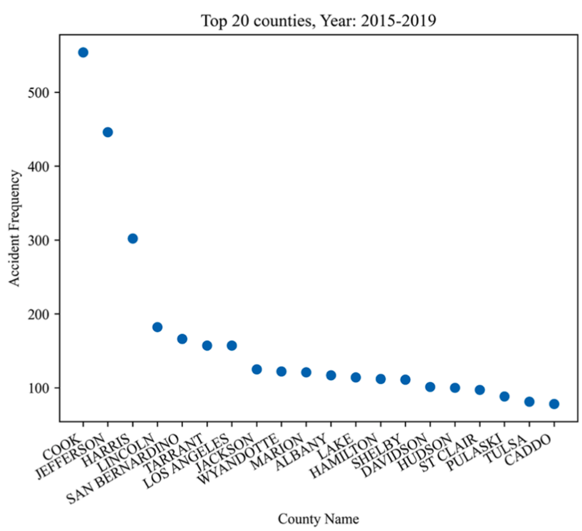

Fig. 1. Top 20 counties with high accident frequency from 2015 to 2019.

collision, rear-end collision, etc. For each record, FRA requires the reporter to submit more than 100 related factors when an accident happens, including the time (year, month, day, hour, minute), place (state, county, station), condition of influenced trains (weight, load, speed, type, affiliation), condition of nature (weather, temperature, visibility), condition of staff (#workers, working hours), details of accidents (cause, damage, type) and other features.

In previous studies, researchers did not consider the possibility of specifying the time cost of data collecting labor to get these factors. Specifically, Table 1 shows the possibility of obtaining the factors related to a certain accident at a certain time. For example, the details of accidents are impossible to obtain until the accident happens days after and the accident investigation work is done. From the above point of view, most of the factors that are essential in the previous accident prediction studies cannot be obtained in advance (days or hours before an accident happens), and thus difficult to be the indicators to predict accidents on time. Furthermore, although some important features, like condition of influenced trains, condition of nature, and condition of staff, can be accessed when the accident is about to happen, the prediction results will become useless once obtained because no time is left for implementing the corresponding accident prevention actions based on these prediction results.

With the motivation of this paper, i.e., to achieve active accident prevention, which needs effective prediction of accident causes and sufficient time for accident prevention, some factors, which contribute to accident prediction, like condition of nature, cannot be reached. In this paper, we only choose the factors that can be obtained before accident prevention actions are implemented. In other words, if an action is implemented on the *i*th day, this action must be predicted before the *i*th day, e.g., on the (*i*-1)th day, and the factors to predict this action must be obtained no later than the (*i*-1)th day.

3.2. Details of chosen data

Table 2 shows the factors used in the proposed model. Three data formats, Numerical, categorical, and narrative data, are selected to make full use of the collected data. Numerical and categorical data is structured data, which is commonly used in accident prediction research, while narrative data is unstructured text data, which needs to be processed with natural language processing (NLP) technologies. The processing method of all these data will be described in Section 4.

The chosen factors, when formed into the input data for active accident prevention, are all the information from the already happened accidents, which we call historical information (HI) in this paper, instead of the information for the accident about to predict. The reasons for this treatment have already been explained in the previous subsection. We design an example to explain the details of this treatment. Assuming there are four successive accidents $\{acc_1, acc_2, acc_3, acc_4\}$ and they happened on the 1st, 3rd, 5th, and 6th day, then if we want to predict the accident cause of acc_4 for implementing an active prevention action for acc_4 (note that in this scenario, acc_4, either its occurrence or cause, is unknown while acc_1, acc_2, and acc_3 happen already), the HI of factor "TEMP" can be represented as [TEMP_{acc1}, PAD, TEMP_{acc2}, PAD, TEMP_{acc3}], where "TEMP_{acc1}" is the temperature when acc_1 happens, and "PAD" denotes that no accident happens so no information collected on this day. In the above example, the HI is composed of the data 5 days prior to the predicted time point. In this paper, the length of referenced days is a hyper-parameter, which is denoted as α.

The treatment for other factors is the same as the above example shows, and the whole historical information is the concatenation of all the HI from different factors. The HI can be accessed days before the accident happens, and thus can be used for an active accident prevention method. Note that the values of factors "HRGCFREQ" and "CANUM" are obtained from HRGC data and CA data.

Moreover, it is important to delineate the scope of the study area. We choose two counties, "COOK" and "JEFFERSON", as the research objects because these counties are with high accident occurrences (as Fig. 1 shows) and most accidents are concentrated in a few stations (as Fig. 2 shows), making it possible to organize these stations to take active accident prevention actions with our proposed methods.

3.3. Accident cause

One important part of this paper is to predict the accident cause to specify how to prevent a possible accident. For example, if the proposed model correctly predicts the cause of the possible accident (e.g., wrongly making a line switch) which will happen the next day, and the staff takes effective actions according to the prediction results (e.g., emphasizing the process and guidelines for line switching), the accident will be avoided with a high possibility.

FRA has concluded more than 389 different causes into five titles (E: Mechanical and Electrical Failures; M: Miscellaneous Causes Not Otherwise Listed; T: Rack, Roadbed and Structures; S: Signal and Communication; H: Train operation - Human Factors). More specifically, they are extended to 28 categories (see Supplementary Material). In this paper, we use granularity to denote the level of detail of causes, i.e., title or category. Generally, the more accurately we know about the cause of the coming accident, the better action we can take to avoid it from happening by specific strategy. So, to predict more than 389 different

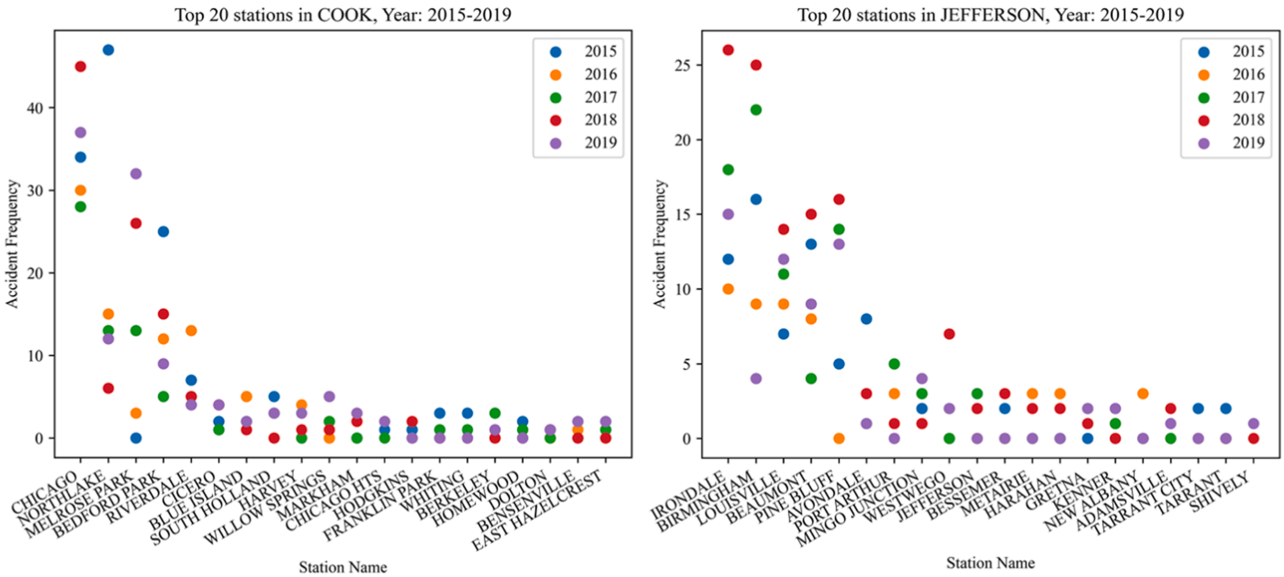

Fig. 2. Top 20 stations with high accident frequency from 2015 to 2019 in COOK and JEFFERSON.

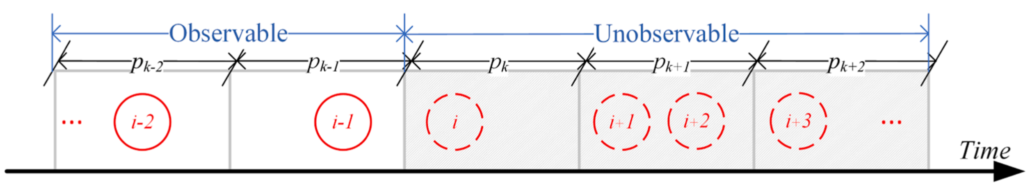

Fig. 3. Problem definition.

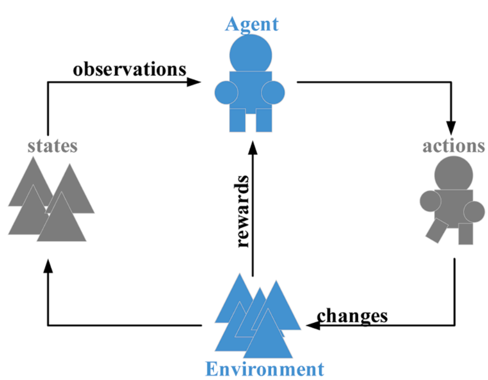

Fig. 4. The framework of reinforcement learning.

causes directly is surely more efficient in avoiding accidents than predicting the five titles of the causes. However, it is a fact that predicting 389 different causes needs much more detail than predicting five titles of the causes, which may reduce the performance of our proposed method. This paper aims to predict the cause in 5 titles and 28 categories. In the proposed model, a cause that corresponds to a null action, which means that nothing is needed to do, is added, and so the total number of predicted objects is 6 for Titles and 29 for Category.

The difficulty of predicting causes varies in different granularities, while successful accident prevention is influenced by how specific a cause is predicted and how effective an action the related staff will take. As a result, the effectiveness of active accident prevention is influenced by a complex factor, which comes mainly from human operations. Since it is difficult to define this factor precisely, a parameter, called valid prevention probability (*vpr*), is introduced in this paper to account for the effect of this factor. The *vpr* is the possibility that the preventive measures are successfully achieved, and we will describe it in detail in Sections 4 and 5.

4. Methodology

4.1. Problem definition

As shown in Fig. 3, the railway accident happening in a certain place can be simplified as a series of nodes appearing with time, where the red circles denote the railway accidents. We define the *i*th occurrence of the railway accident on the timeline as acc_i, and separate the time flows into a series of equal-length time slices (periods), e.g., the *k*th time slice (period) is p_k, thus cutting all the accidents into separate time slices. Now assuming that acc_{i-2} and acc_{i-1} have already happened and the time is at the end of p_{k-1}, the aim of this paper is to predict the occurrence and the causes of acc_i, acc_{i+1}, acc_{i+2}, acc_{i+3}, and the following accidents in turns in advance at least a period, and then take timely preventive actions to avoid the happening of them. In this scenario, the factors affecting acc_i cannot be collected to predict acc_i because they remain unobservable in time, so we have to collect only the information that is available for the observable part before p_k.

In this paper, the information from the observable part is called HI. The dynamics that exist in HI make the prediction of railway accident causes different from the classical accident prediction methods, which are based on a supervised learning approach. For example, from the accident records, we may train an accident prediction model that uses HI of acc_{i-2} to predict acc_{i-1}, uses that of acc_{i-2} and acc_{i-1} to predict acc_i, …, uses that of acc_{i-2}, …, acc_{i+2} to predict acc_{i+3}. However, when this well-trained model is used to predict accident causes and there are the same accident series as the accidents trained before, we will see that if acc_{i-1} is correctly predicted and prevented, HI becomes only from acc_{i-2}. So, the prediction of acc_i may be failed because the well-trained model uses HI

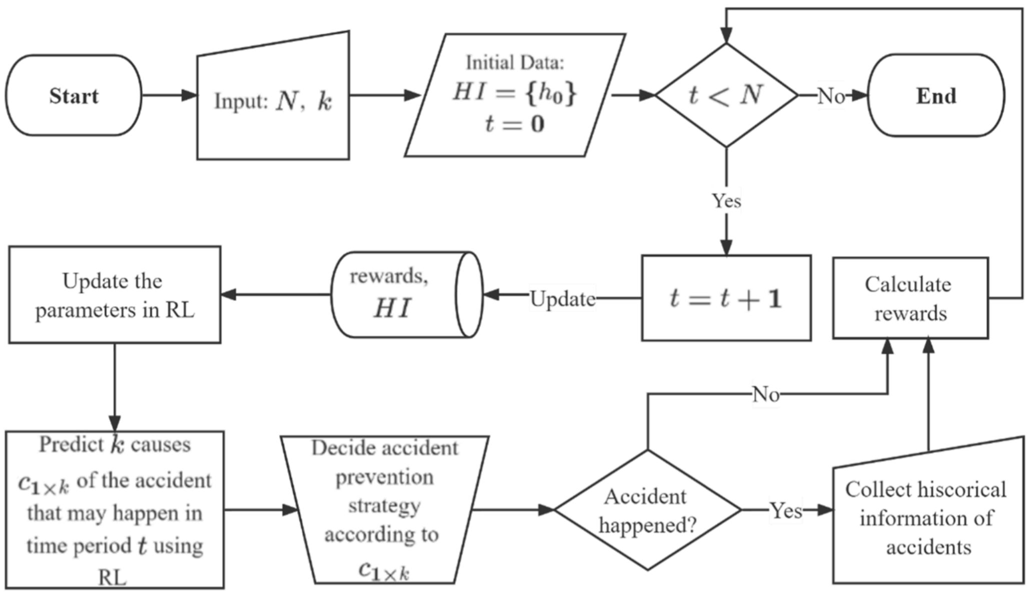

Fig. 5. Flow chart of the proposed method.

of $acc_{i\text{-}2}$ and $acc_{i\text{-}1}$ to predict acc_i but now it becomes only of $acc_{i\text{-}2}$.

From the above content, we can observe that our question can be more suitable to be described as a decision-making process where actions (predicting cause and preventing accidents) are decided by the observations (observed HI) and the observations are in turn influenced by previously adopted actions [43]. So, we proposed a reinforcement learning method, which is commonly used to solve decision-making problems, to solve this problem.

4.2. Accident prevention model based on reinforcement learning

As Fig. 4 shows, a classical RL framework contains agent(s) and an environment. The agent(s) react to the environment changes, and the environment evolves with agent(s)' actions. By interacting with the environment, the agent(s) gradually learns how to adapt to changes in the environment and make the best response. Once the agent(s) decide on an action based on the current state, a reward is returned, denoting the contribution of this action to the final target. In the initial stages, the agent(s) don't know the best reaction to the environment. As the training deepens, the appropriate actions are gradually learned according to the rewards from every step.

Let E denotes the environment of the studied object, i.e., hazard-inducing and hazard-preventing environment related to a certain part of the railway systems. Considering that the implementation of the proposed model is generally in a specific area, the environment is also confined to a certain place, e.g., a railway station, a city, or a county. In the RL model, E is defined by a finite set of states, S, actions, A, a mechanism, P, for predicting actions according to states, and a reward function, r. As described in the previous sub-section, the studied time is split into several periods, cutting all the accidents into separate time slices. In every period, called a step t, the decision-maker, i.e., the agent in RL, takes actions a_t according to the current state, s_t, then gets a reward, $r(s_t,a_t)$, denoting how well the accidents are prevented and how much labor is costed, and proceeds to the next step $t+1$. In this paper, the action that the agent will take at every step is the corresponding precautionary measures to probable causes of accidents predicted to happen in the future. For simplicity, this paper ignores the details of precautionary measures, and only considers the accurate prediction of causes and whether or not accidents will be prevented. The randomness from the correctly predicted causes to successful accident prevention is considered as a parameter, called valid prevention ratio (*vpr*), indicating the possibility to achieve successful accident prevention when the accident causes are correctly predicted. The agent directly interacts with E, and learns to act optimally according to the obtained reward step by step. After finite steps (N), the agent reaches the terminal state, and one epoch of RL is finished, producing a sequence of actions. The action sequences, which are state-dependent, define the policy of the agent, π. The policy can either be deterministic or stochastic, and this paper adopts a stochastic form: $\pi(a_t|s_t) : S \rightarrow P(A)$ [34], where π is represented as a vector denoting the probability over all actions, which corresponds to all the possible causes, C. To predict the occurrence and cause of an accident, it should also be considered that no accident will happen. So, the actions should include a null action ($|A| = |C| + 1$), which means that no action is needed to be taken. Without the null action, the model will always produce actions even if no accident will happen in the next step. So, the null action is a consideration of labor cost, which indicates when preventive actions are necessary and when saving labor costs is necessary by taking no action.

By evolving with the environment, the policy, π, is updated according to the corresponding total return, R_t^π:

$$R_t^\pi = \sum_{i=t}^{N} \gamma^{i-t} r(s_i, a_i) \tag{1}$$

where γ is a discount factor in the range of 0 to 1 indicating the importance of future actions to the current action; the step reward, $r(s_t, a_t)$, is calculated according to Eq. (2).

$$r(s_t, a_t) = -\frac{n_{acc}|_{s_t,a_t}}{N} - \varepsilon(s_t, a_t) \tag{2}$$

where $n_{acc}|_{s_t,a_t}$ denotes the number of accidents in step t and $\varepsilon(s_t, a_t)$ is the indicator of how much invalid labors are paid in step t. The value of $\varepsilon(s_t, a_t)$ can be calculated according to Eq. (3), where κ is a constant that indicates the value of the work put into each action, ϕ denotes whether a null action exists (if True, $\phi = 1$, else, $\phi = 0$.), and $|a_t|$ denotes the number of actions produced by the model in step t.

$$\varepsilon(s_t, a_t) = \kappa\left(1 - \phi|_{s_t,a_t}\right)|a_t| \tag{3}$$

The evolving of policy, π, falls on the updating of the parameters of the policy function P. Numerous studies have been implemented to achieve a stable and fast-converging policy function like DQN [43], which is out of the scope of this paper. This paper adopts the deep policy

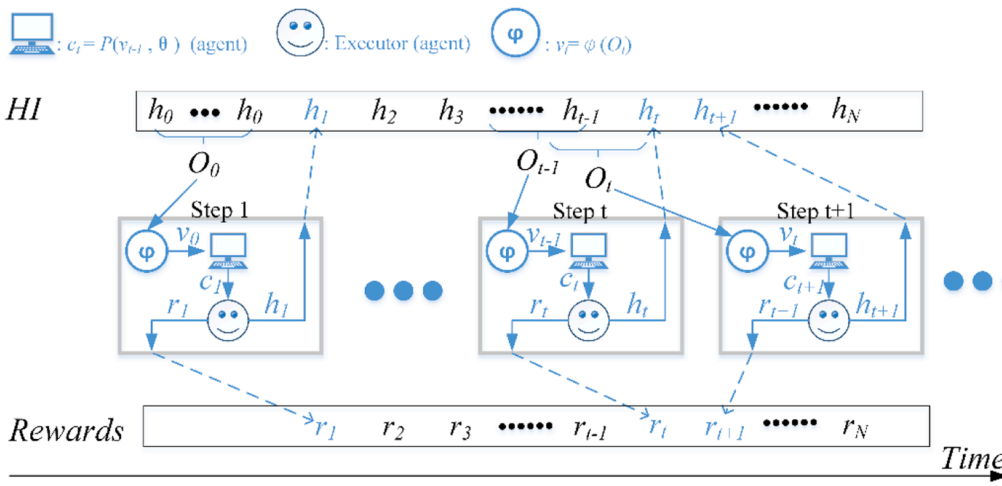

Fig. 6. Data flow of RL.

Algorithm 1
Deep Policy Gradient Algorithm.

Input: Total #episodes $\boldsymbol{M}$, Total #steps $\boldsymbol{N}$, #Actions $\boldsymbol{k}$
Initialize: Neural network (***net***) weights $\boldsymbol{\theta}$, Replay buffer ***bf***
for episode = 1 to $\boldsymbol{M}$ **do**
  Initialize the environment, $\boldsymbol{E}$, and obtain the observation $\boldsymbol{O}_t(t=1)$
  for $t = 1$ to $\boldsymbol{N}$ **do**
    Select $\boldsymbol{k}$ actions, $\boldsymbol{a}_t$, according to $\pi(\boldsymbol{O}_t, \boldsymbol{\theta})$ sampled from ***net*** if for training
    Select top $\boldsymbol{k}$ actions, $\boldsymbol{a}_t$, according to $\pi(\boldsymbol{O}_t, \boldsymbol{\theta})$ predicted from ***net*** otherwise
    Update $\boldsymbol{E}$ with $\boldsymbol{a}_t$ and get the step reward, $\boldsymbol{r}(\boldsymbol{O}_t, \boldsymbol{a}_t)$, by Eq. (3)
    Store experience $(\boldsymbol{O}_t,\ \boldsymbol{a}_t,\ \boldsymbol{r}(\boldsymbol{O}_t, \boldsymbol{a}_t))$ in ***bf***
    Obtain the observation $\boldsymbol{O}_{t+1}$ from $\boldsymbol{E}$
  end for
  Calculate $\boldsymbol{R}_t^{\pi}$ by Eqs. (1) and (6) based on the ***bf***
  Update $\boldsymbol{\theta}$ according to the gradient obtained by Eq. (5)
  Clean ***bf***
end for

gradient to learn the policy, π, with a two-layer fully-connected neural network (FCN). The optimization process of the neural network is to iteratively update the parameters, θ, within it by constructing an objective function that is minimized during training. The loss of the objective function can be defined as:

$$L(\theta) = \sum_t \log\pi(a_t|s_t, \theta) R_t^{\pi} \tag{4}$$

and the policy gradient will be estimated by:

$$g_\theta = \mathbb{E}\left[\sum_t \Psi_t \nabla_\theta \log\pi(a_t|s_t, \theta)\right] \tag{5}$$

where Ψ_t is calculated by Eq. (1) while using the normalized $r_n(s_t, a_t)$, calculated by Eq. (6), to replace $r(s_t, a_t)$.

$$r_n(s_t, a_t) = r(s_t, a_t) - \bar{r}(s_t, a_t) \tag{6}$$

where $\bar{r}(s_t, a_t)$ denotes the average of the sequence of $r(s_t, a_t)$.

In many occasions of RL, the states are partially observable in every step, thus the policy is generated with observation values in step t, denoted as O_t. In this paper, it consists of a series of the latest data in HI, i.e., $O_t = \{h_j\}_{t-\alpha+1<j\leq t}$, $HI = \{h_j\}_{j\geq 0}$. The h_0 denotes that no accident happens, and thus HI about accidents is empty. Further, let o_{tk} denote the kth factor related to the occurrence of accidents, e.g., weather, temperature, and accident damage value, observed in step t, and we get $h_j = \{o_{jk}\}_{k\leq K}$, where K is the number of factors. For the non-numerical factors, let φ be a function to convert them to vectors, i.e., $v_t = \varphi(O_t)$, where v_t represents a vector synthesized by vectorizing all the factors contained in O_t.

Fig. 5 shows the flow chart of the method presented in this paper. N denotes the number of short periods (steps) in an epoch, with each end of the period yielding k probable accident causes in the next period through RL based on HI. The start step contains no HI, and we set an h_0 to denote that the history is empty. Then, related staff on duty can take some accident prevention strategies in advance to avoid potential risks. At the end of the next period, the information about the happened accident is collected to update a database, including HI and rewards. Further, with the updated rewards, the parameters in RL are updated. To clearly describe the details of the proposed method, we show the data flow of RL in Fig. 6 and present the algorithm to implement the deep policy gradient of the proposed method in Algorithm 1.

4.3. Vectorization for non-numerical factors

The HI contains different data formats of factors, e.g., the temperature value is numerical, visibility is categorical data, and this paper also uses narrative data as part of HI.

The numerical data are all analyzed to get their mean and standard deviation and then normalized through Eq. (7), where x_{si} is the ith real value of factor s, u_s, σ_s are mean and standard deviation of factor s, respectively.

$$x_{si}^* = \frac{x_{si} - \mu_s}{\sigma_s} \tag{7}$$

The categorical data are vectorized through the embedding method. Embedding vectors, the standard paradigm of neural network-based natural language processing, are used to vectorize the constituent units of text data, such as words [52–54]. Let noc denote the total number of categories for all the categorical data, and $E_{noc\times d}$ is a matrix with noc rows and d columns. An embedding method maps any category value to a unique index to the row of $E_{noc\times d}$. And then every category value corresponds to a vector in a certain row of $E_{noc\times d}$. Let the above map be M, then Eq. (8) shows the embedding from categorical data y_{si} to its corresponding vector v_{si} (with dimensionality of d). "[]" represents the index of a value at a particular position in the matrix, e.g., $E_{noc}[i,j]$ is the value at ith row and jth column of E_{noc}. ":" denotes indexing all the value of corresponding row or column.

$$v_{si} = E_{noc}[\mathrm{M}(y_{si}), :] \tag{8}$$

The narrative data is unstructured data. The vectorization of narrative data (t_{si}) is based on the AEBoW model proposed in our previously published paper [30]. Then the dimensionality of vectors is further reduced by LSA [55], as shown in Eq. (9). Some records contain a short narrative text (the number of total words is less than 20), which is a

Table 3
vpr for category and Title of the causes.

Granularity	#Action	vpr
Category	28 + 1 (29)	0.8
Title	5 + 1 (6)	0.4

problem of imbalanced text data. This paper adopts a strategy that we proposed in the literature [52], which expands the content of text data by using the network-based random walking method, to solve this problem.

$$v_{si} = LSA(AEBoW(t_{si})) \tag{9}$$

5. Case study

5.1. Setup

As discussed before, we design a *vpr* to denote how likely a correct prediction can lead to a successful prevention strategy. Unfortunately, it is hard to know the value of *vpr* precisely, so in this paper, we define the value of *vpr* with empirical knowledge. Table 3 shows the value of *vpr* for Category and Title of the causes, where the “+1″ in #Action denotes that each granularity of actions includes a null action, which means do nothing.

The length of the epoch and the length of time in one step are two factors influencing the complexity of the environment in the proposed model. The length of one epoch is fixed in 365 steps, and the length of time in one step is one day during the case study.

In Section 3.2, we use α to denote how many elements in HI are used in prediction. To set large α is beneficial to the performance, but will highly increase the scale of learnable parameters of the proposed method. This paper sets $\alpha = 20$ to make a balance between the performance and the scale of parameters. For the prediction of action, the number of actions, i.e., the number of varied causes, to be implemented is not necessarily set as one. In this paper, we use k to denote this hyperparameter. Intuitively, higher value of k can obtain more accurate accident prevention because more actions can be adopted for accident prevention, but higher k needs more financial and labor costs. The default value of k is 2.

Section 4.3 has discussed the vectorization of chosen factors. For categorical data, set the embedding dimensionality $d = 32$, and total number of categories $noc = 4 + 6 + 14 + 4 + 6(29) = 34(57)$. For the narrative data, the reduced dimensionality for each text is 16.

The implementation of the proposed method is based on a Windows system and Python(version 3.7) interpreter. The RL is achieved by using PaddlePaddle(version 2.3)[3] and PARL[4] modules. The maximum training episode for RL is set as 4000; the learning rate is set as 0.001; $\kappa = 0.14$ and $\gamma = 0.9$.

To show the effectiveness of the proposed method, we introduce two baselines to make a comparison. It should be noted that there is little other research aiming at railway active accident prevention, and thus it is difficult to make a fair comparison between the proposed method and the other methods. So we only show the comparison of the performance on how effectively to predict the accident cause based on HI, which is an essential part of active accident prevention. The two baselines are further designed to fit the above aim, as shown in the following.

(1) **Random Forest (RF):** RF is proven to have good performance in railway accident prediction [2]. Here, RF is introduced as the baseline to demonstrate the advantages of the proposed method compared to machine learning methods.

(2) **Fully-connect Neural Network (FCN):** FCN is a basic structure of the neural network models, which is also an important part of our proposed RL-based method to learn the non-linear relationship between variables. The structure of this FCN is the same as the structure in the proposed method. By removing the part of RL, we would like to show the improvement from the RL model in our proposed RL-based method.

The training of the RF and FCN model is different from our proposed method: RF and FCN are trained in a supervision learning way, while our proposed method is trained by the way of RL.

Furthermore, we design three metrics to show the performance of each method on active accident prevention, as shown in the following.

(1) **Correctly Predicted Rate (CPR):** CPR is the indicator of how accurate is a model in predicting the accident causes. Predicting the cause is the first step in the active accident prevention scenario proposed in this paper, high CPR is the guarantee of effective accident prevention. The calculation of CPR is presented in Eq. (10), where $n_{predict}$ is the number of correctly predicted causes, and n_{actual} is the number of all happened accidents in an epoch.

$$CPR = \frac{n_{predict}}{n_{actual}} \tag{10}$$

(2) **Invalid Labor Rate (ILR):** ILR denotes the degree of invalid labor during accident prevention. The accident occurrence is of sparsity characteristic, i.e., the chance of having an accident is much smaller than the chance of not having an accident. So a long-lasting preventive measure will unavoidably fall into the problem of low returns, and most of the accident prevention costs become must-paying but unnecessary. ILR can be used to reflect how much ineffective payoffs are avoided by active accident prevention models. The smaller the value of ILR, the more ineffective payoffs a model avoids. ILR is calculated as Eq. (11) shows, where $d_{act\text{-}noacc}$ is the days that no accident happens but the model advises that no less than one cause should be avoided, and d_{noacc} is the days that no accident happens in a period.

$$ILR = \frac{d_{act-noacc}}{d_{noacc}} \tag{11}$$

(3) **Prevention Missing Rate (PMR):** PMR can represent the model’s ability to correctly judge the chance an accident will happen. The definition of PMR is the proportion of the model advising that no action is required during the days occurring accidents, as Eq. (12) shows. In this equation, $d_{noact\text{-}acc}$ is the days that no less than one accident happens but the model advises that no action is needed, while d_{acc} is the days that no less than one accident happens in a period. The smaller the value of PMR, the less the model misses a preventive measure.

$$PMR = \frac{d_{noact-acc}}{d_{acc}} \tag{12}$$

5.2. Results and comparison for accident prevention

This part presents the prevention performance based on the proposed method and the setup in Section 5.1. First of all, we show the comparison among results from the proposed method and the other two baselines. The results neglect the influence of *vpr*, i.e., let *vpr* be 1, to eliminate the randomness because the baselines cannot consider the value of *vpr*, and are all based on the granularity of “Category”.

[3] https://github.com/PaddlePaddle
[4] https://github.com/PaddlePaddle/PARL

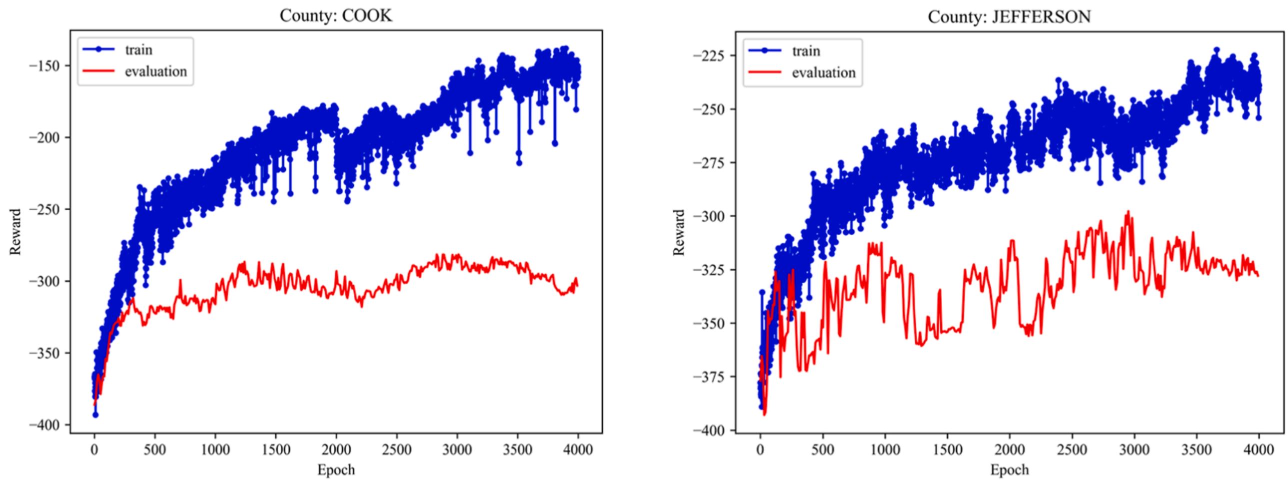

Fig. 7. The trajectory of rewards during training.

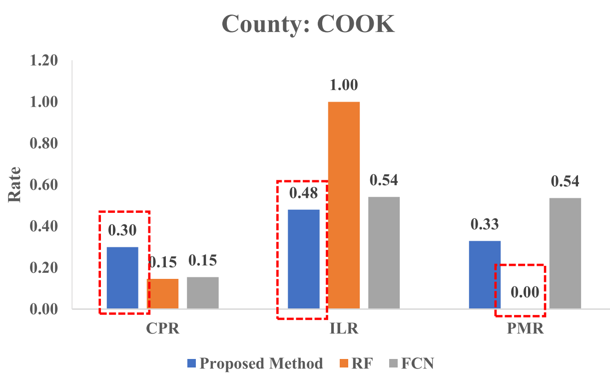

Fig. 8. Comparison of results based on the data from COOK.

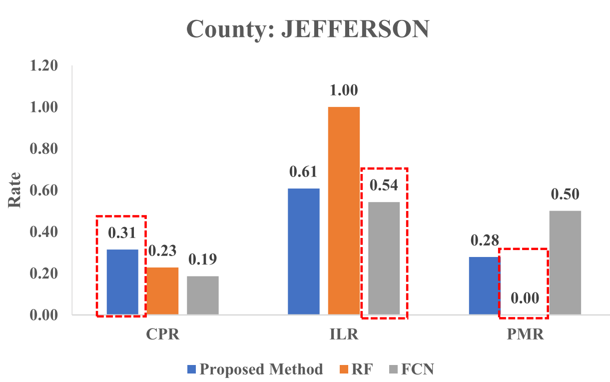

Fig. 9. Comparison of results based on the data from JEFFERSON.

Fig. 7 shows the trajectory of rewards during training, where the training data of a certain epoch is chosen from records during the years 2015–2018. The records in 2019 are selected as the test data. The two subfigures all show the training processes finally reach their convergency states after 4000 epochs. We can also observe that the curves of evaluation rewards quickly rise to a state of near convergence (at nearly 1500 epochs), and rewards are getting better slowly as the epochs increase after that. Nevertheless, it is still necessary to train up to 4000 epochs, because we have observed that the results for three metrics after 4000 steps are better than those of 1500 steps. Another obvious feature of the reward trajectory is that curves of train rewards and curves of evaluation rewards show significant gaps with the epochs increasing. These gaps come from the inherent randomness of accidents occurring over different periods, i.e., a portion of the accidents' patterns cannot be summarized from HI. The fact that both trajectories are rising, in turn, indicates that some accidents can be predicted from HI, and we will also show with experimental results that the proportion of this fraction is encouragingly large enough for our efforts to study them.

Before analyzing the results, it is necessary to demonstrate how to correctly understand the results in the figures. Firstly, CPR is the most important index among these three metrics because CPR is a direct indicator of the ability to predict the accident causes, which may achieve successful accident prevention. ILR and PMR are the indicators that show a model's ability to give appropriate advice for accident prevention implementation. For example, if the model correctly predicts the cause of an accident but gives the wrong advice that today is not necessary to take action to prevent the accident, the accident will still happen. Conversely, if the model always advises that active actions should be applied, the staff will be caught in an endless accident prevention effort that will cost a great deal of manpower and resources. These two parameters are mutually constraining, and we certainly want them both to be small, but it is not desirable to make one value too large in order to reduce the other.

The results based on the data from COOK and JEFFERSON are listed in Figs. 8 and 9, where the red dashed boxes highlight the optimal values relative to the individual metric. From the results, it can be observed that the three methods, including the proposed methods and two baselines, behave with a big difference. RF model, as an effective model in railway accident prediction [2], has very poor performance when applied to active accident prevention based on a complex combination of data. From the definition of ILR and PMR, we can know that when ILR=1 and PMR=0, the model neglects the constraints of labor cost and advises that preventive actions should be taken every day. After analyzing the details of the prediction results, we also observed that RF gives the same prediction results of accident causes all the time. In other words, RF learns little useful information from the data. The main reason may be that RF cannot adapt to complex data with high-dimensional and multiple sources. FCN model can effectively learn information from HI, and get a fairly balanced result among CPR, ILR, and PMR. However, the value of CPR in the results of FCN lags significantly behind that of the proposed method. We attribute this to the dynamic nature that HI exhibits in training, i.e., the action adopted in this step will influence the value of HI which is used to predict the accident cause in the next step. The supervision learning of FCN can only learn the already happened patterns and failed to adapt the dynamics caused by the actions for accident prevention. The proposed method, in an RL-based way, can not only learn valid information from HI, but also adapt to the dynamic nature. Thus, the proposed method can effectively predict the accident

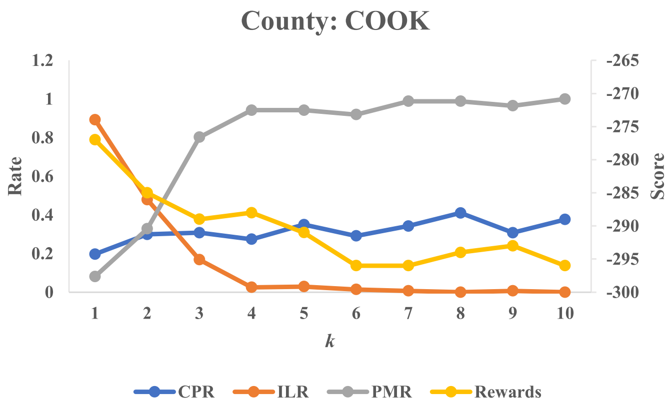

Fig. 10. The trend of k influencing the CPR, ILR, PMR, and evaluation reward ($k = 1,2,\ldots,10$) based on the accident records from COOK.

causes. The CPR from the proposed method is 30% for COOK and 31% for JEFFERSON, exceeding FCN by 15% and 12%, respectively. At the same time, the proposed method achieves a balance between ILR and PMR. For COOK, the ILR and PMR of the proposed method are lower than that of FCN by 6% and 21%, respectively. For JEFFERSON, the value of ILR is 7% higher than that of FCN, but the value of PMR is 22% lower than that of FCN. The results show the advantage of our proposed method of active railway accident prevention, as well as the potential to save labor costs as much as possible.

As we have verified the performance of the proposed method in more than one area, the following analysis will concentrate on COOK to save the length of this paper.The results from JEFFERSON are available in Figs. A1-A6 of the Appendix.

Besides, we also explore the influence of k for action prediction in the proposed method (the default k in this paper is 2). Though a higher value of k can produce more actions simultaneously for accident prevention, which may get better results, higher k needs more financial and labor costs. The proposed method takes the influence of labor cost into account, so the results with different k will somehow reflect the model's preference between accuracy and cost savings. Fig. 10 shows the trend of CPR, ILR, PMR, and evaluation reward with varied values of k in COOK, where the value of CPR, ILR, and PMR refers to the left y-axis and that of rewards refers to the right y-axis.

We have mentioned above that the results with larger k might obtain more accurate predictions, the value of CPR in Fig. 10 validates this trend. But it is too early to get the conclusion that higher k makes better. From the trend of ILR and PMR, it can be observed that the model tries to reduce the value of ILR and increase the value of PMR with k increasing. To explain this, we need to know the average number of accidents that occur each day in COOK, which is easy to obtain from Fig. 1, i.e., less than 600 in 5 years, that is, less than 0.3 per day. More specifically, there are at most 3 accidents one day in COOK in 2019. As a result, the model has to find ways to reduce the extra labor costs due to the excessive number of actions. When k is large, the labor cost of the excessive number of actions will major in the rewards, and so the model chooses to take no action, which is one option in the setup of the model, thus yielding the increase of PMR and decrease of ILR. This is not a desirable option in practice, as it means that most of the time nothing is done for accident prevention. The intersection point in the figure appears in the vicinity of 2, which is one of the reasons for choosing $k = 2$ in this paper. It must be noted that the timing of the appearance of this intersection point varies for different accident frequencies, and therefore the ideal values are slightly different for different counties.

5.3. Challenge in practical using: randomness

As far as we have verified the advantage of our proposed method in railway active accident prevention, in this section, we will put our attention on what will happen once the proposed model is applied in practical use.

It is difficult to guarantee that every accident will be successfully prevented in practice even if the cause of this accident is correctly predicted. The influencing factors on failed prevention are very complex, including the main part that comes from human operation. We reduce this effect to a parameter, called valid prevention ratio (*vpr*), which is included in the model. In the previous sub-section, we compare the results with other baselines based on the setup that $vpr = 1$ to verify the advantage of our proposed model in theory. In this section, we will present the results considering the influence of randomness, i.e., set *vpr* to a certain value lower than 1 according to how difficult an accident can

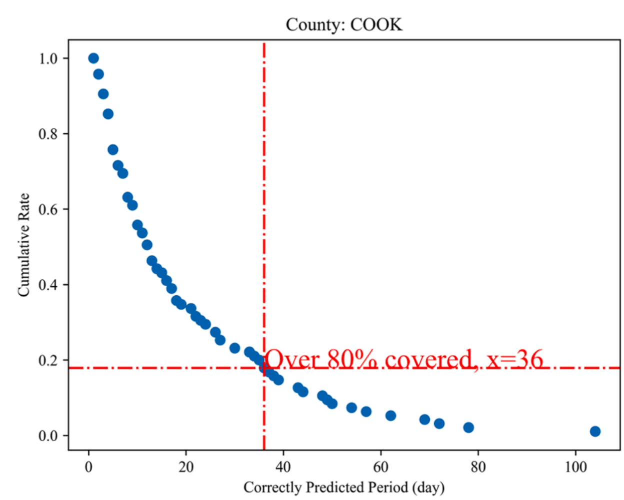

Fig. 12. The cumulative rate of the correctly predicted period in COOK (2015–2019) (Granularity: Category).

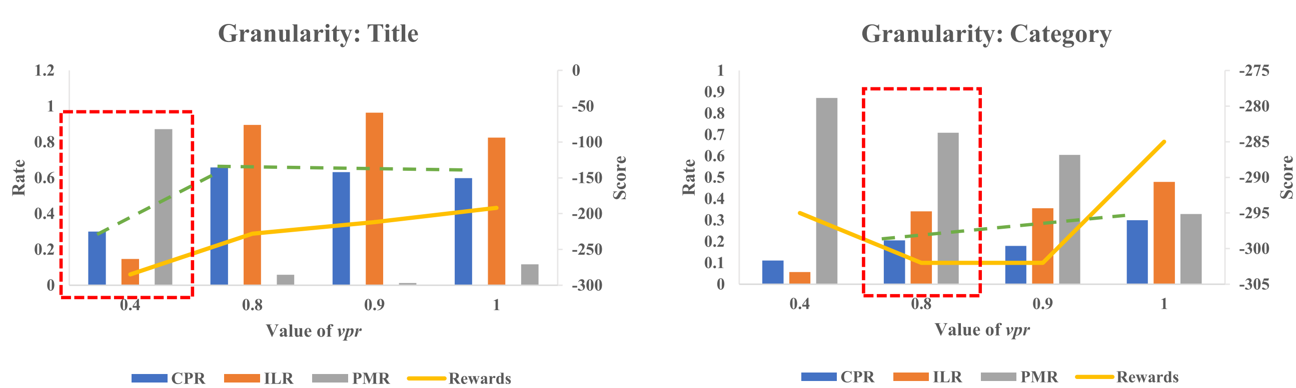

Fig. 11. The influence of *vpr* in different granularities based on the accident records from COOK.

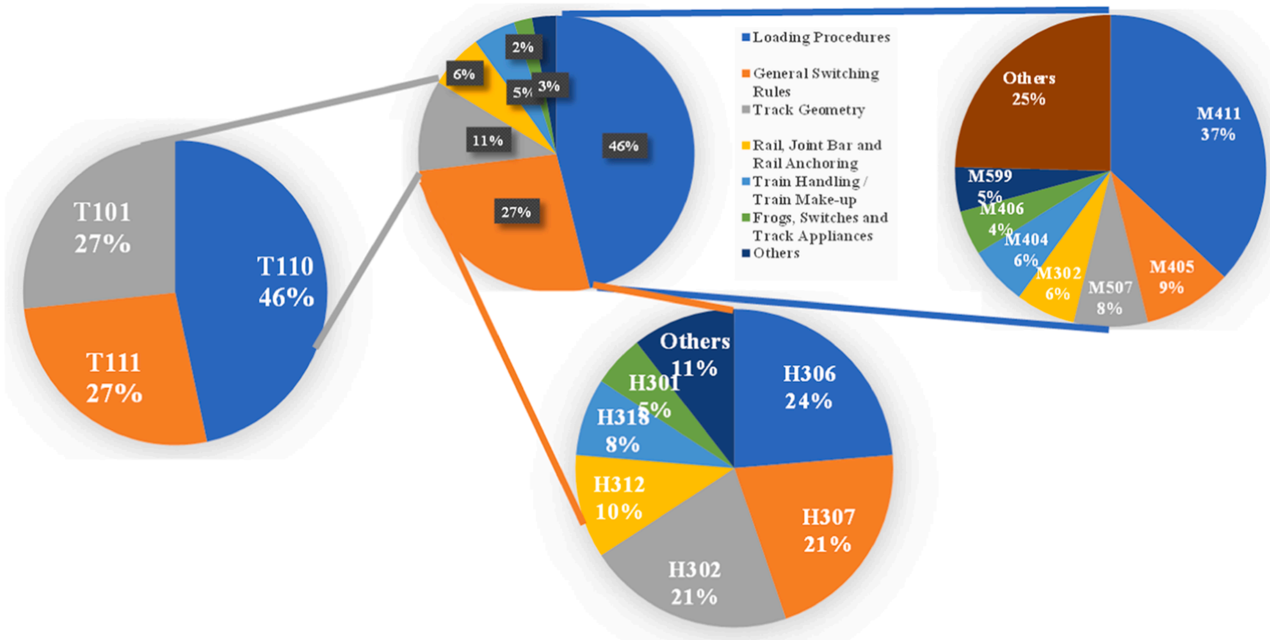

Fig. 13. Overview chart of the correctly predicted causes in COOK from 2015 to 2019 (Granularity: Category).

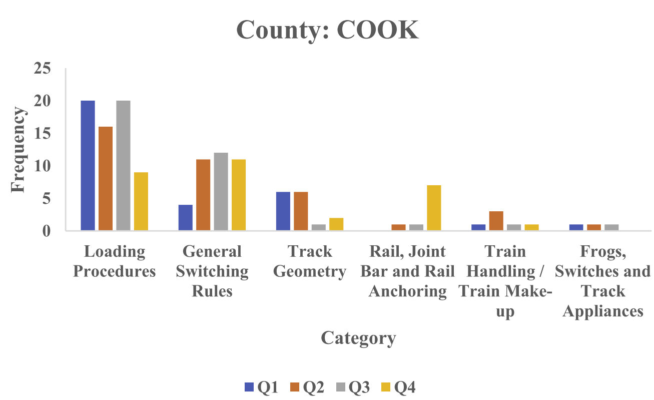

Fig. 14. Frequency of correctly predicted causes by quarters in COOK from 2015 to 2019 (Granularity: Category).

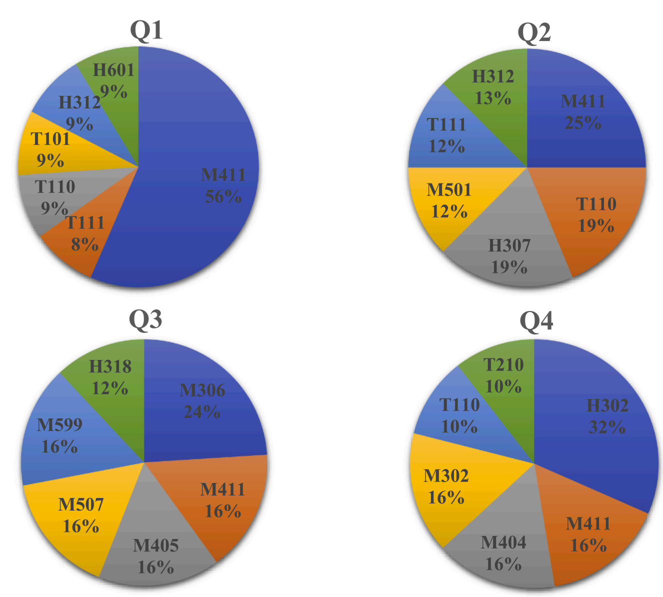

Fig. 15. Top 6 correctly predicted causes by quarters in COOK from 2015 to 2019 (Granularity: Category).

be prevented in practice.

Fig. 11 shows the influence of *vpr* in different granularities based on the accident records from COOK. Granularity is the consideration of the precision of accident cause prediction, as is shown in Table 3. A more precise cause will more easily lead to successful accident prevention, so correspond to a larger *vpr*. Table 3 shows our default setup of *vpr* for different granularities, while in experimental results we also explore other values of *vpr* influencing the results for a certain granularity. The results from the default setup are marked with red dashed boxes. From the results, we can observe that once the value of *vpr* is considered, the performance of the model will unavoidably degrade. The degradation level of CPR can be indicated by the green lines in the figure. Also, with low *vpr*, high judgment error, corresponding to high PMR, occurs when judging an accident occurring. Although the ILR is low, it does not indicate the labor cost is saved because the model always advises to not take action. We can use other words to conclude these results: with low *vpr*, the staff is free of labor for accident prevention, but the railway sectors will not be free of handling the impact of accidents.

Furthermore, some significant conclusions can be drawn from Fig. 11. Firstly, the model with larger granularity has a higher potential for enhancement. The model with the granularity of "Title" will witness a great improvement, including a significant rise in CPR and a significant drop in PMR, as long as its *vpr* increase from 0.4 to 0.8. However, the decisions based on this model have to afford a big increase in labor cost, i.e., a significant rise of ILR, to achieve the improvement. Secondly, as *vpr* increases, the enhancement from unit *vpr* increase will be reduced. This may be helpful in practice because the quantitative-based analysis can find the boundaries of optimal efficiency in practice and improve railway safety effectiveness at a minimal cost. For example, the model with the granularity of "Title" is effective enough when $vpr = 0.8$, while for that of "Category", a more suitable *vpr* is between 0.9 and 1.

5.4. Safety enhancement of railways – start from the staff

After analyzing the results from the proposed methods and exploring

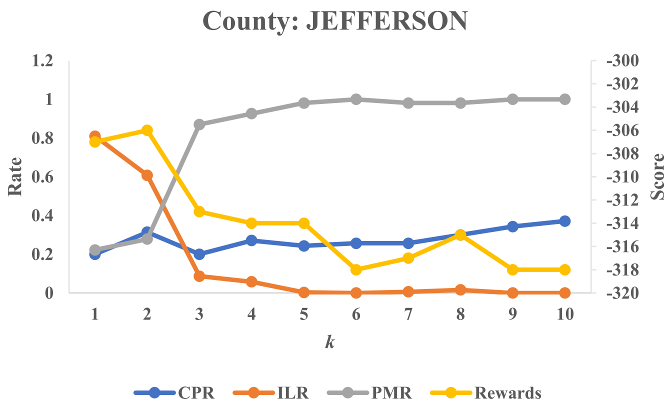

Fig. A1. The trend of k influencing the CPR, ILR, PMR, and evaluation reward ($k = 1,2,\ldots,10$) based on the accident records from JEFFERSON.

the influence of the value of *vpr*, we have learned the importance of reducing the randomness that occurs in practice. In this section, we will show some possible strategies for the safety enhancement of railways from the staff's point of view based on the results from the proposed model.

Note that our focus on safety enhancement is to make the proposed model obtain performance close to theoretical analysis. To achieve this goal, we need to raise the value of *vpr* to its optimal state in the model. As is shown in Table 3, the default values of *vpr*, which are obtained from our estimation, all have a certain gap with the optimal. Though we do not know what the specific gap is, it is certain that the direction of optimization is necessarily the direction of improving *vpr* value. Active accident prevention actions are taken mainly by the staff. So, our conclusions focus on how staff will do.

There are two aspects to getting a better *vpr* value: timely and regular safety skills training, as well as precise training for staff. For timely and regular training, we need to know how often training is appropriate; for precise training, we need to know which skill is the most necessary to be trained for certain training. For the first aspect, we count the causes of accidents that the model could correctly predict and determine the length of the training cycle according to the correctly predicted period; for the second aspect, we further count the finer causal factors corresponding to the accident causes that were correctly predicted. Because the analyzing object is the correctly predicted causes, training employees on these contents will be highly relevant to the accident about to happen, which is targeted and precise.

In Fig. 12, we show the cumulative rate of the correctly predicted period in COOK from 2015 to 2019. The red dashed lines mark the point whose value of the x-axis ($x = 36$) is over 80% of correctly predicted periods. The reason that we count the correctly predicted period is that we hope once an employee has been trained, the skills he/she trained are to be used as much as possible before the next training. Considering that 80% of the accidents with a high possibility to be correctly predicted will happen within 36 days, we believe that setting the training cycle as 36 days will meet the requirements. So, for timely and regular training, the staff in COOK are supposed to be trained with active accident prevention skills within 36 days.

Next, we will explore which skill is the most necessary to be trained for the staff. Fig. 13 shows the overview of the correctly predicted cause in COOK from 2015–2019. The causes related to "Loading Procedures", "General Switching Rules", and "Track Geometry" consist nearly 84% of the total correctly predicted causes. This means that the related staff members, if proficient in accident prevention skills that target these three types of accidents, will greatly increase the likelihood of successful accident prevention based on the proposed method in this paper. The details of the causes related to these three kinds of corrected predicted causes, where M411, M405, M507, H306, H307, H302, T101, T110, and T110 should be put more effort, should be invested in training to improve the individual's ability to prevent these accidents with precision. We also take into account the four seasons of the year and obtained Figs. 14 and 15, where Q1 denotes the first quarter of a year. From Fig. 14, we can observe that different kinds of causes should be treated with different approaches. For causes related to "Loading Procedures" and "General Switching Rules", we should always treat them with active and adequate measures. For causes related to "Track Geometry" and "Rail, Joint Bar and Rail Anchoring", active measures should be taken in certain seasons, e.g., Q1, Q2 for "Track Geometry" and Q4 for "Rail, Joint Bar and Rail Anchoring". The other kind of causes may be targeted trained in longer cycles because the correct prediction frequency of these causes is low all the time. In Fig. 15, the top 6 frequently correctly

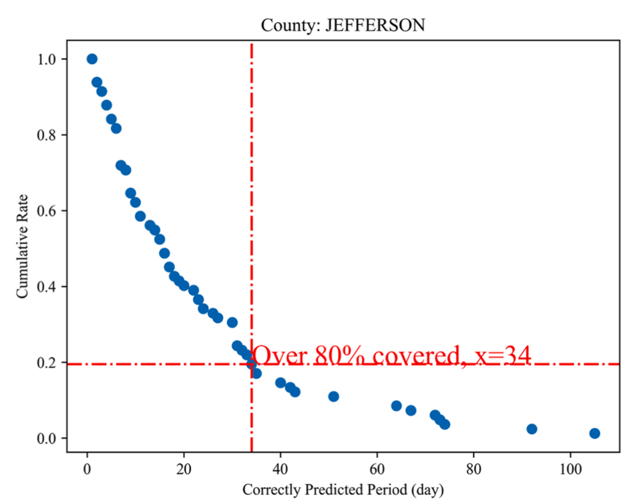

Fig. A3. The cumulative rate of the correctly predicted period in JEFFERSON (2015–2019) (Granularity: Category).

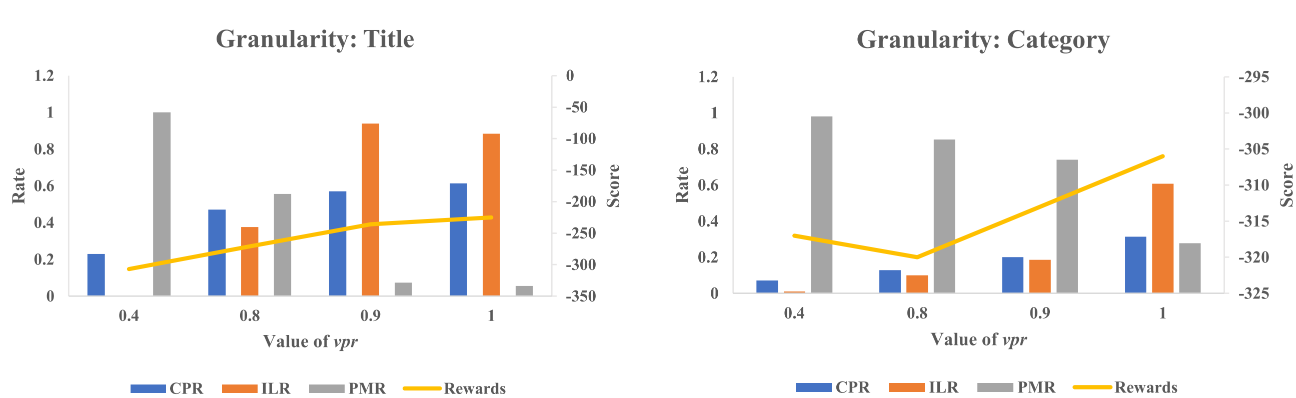

Fig. A2. The influence of *vpr* in different granularities based on the accident records from JEFFERSON.

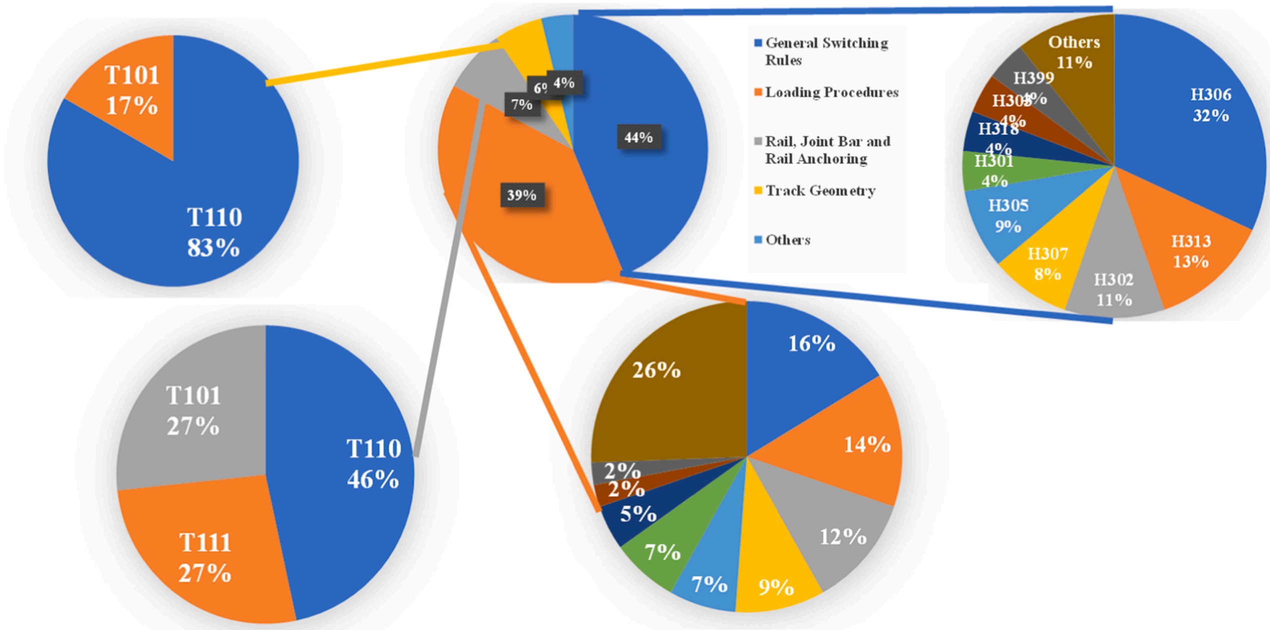

Fig. A4. Overview chart of the correctly predicted causes in JEFFERSON from 2015 to 2019 (Granularity: Category).

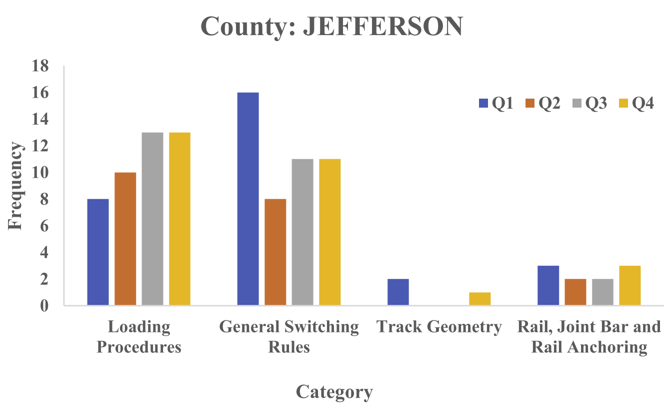

Fig. A5. Frequency of correctly predicted causes by quarters in JEFFERSON from 2015 to 2019 (Granularity: Category).

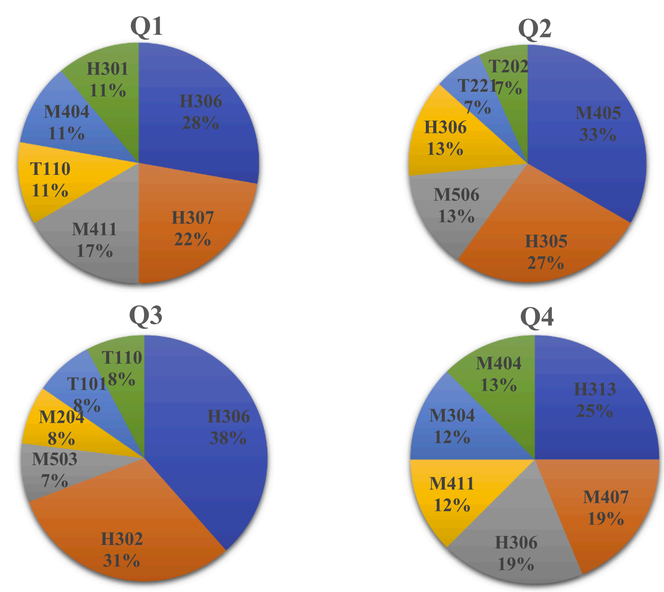

Fig. A6. Top 6 correctly predicted causes by quarters in JEFFERSON from 2015 to 2019 (Granularity: Category).

predicted causes are listed. It can be observed that except M411, each season has its own type of accident that requires a special response. For example, T111 in Q1, T110 and H307 in Q2, M306 in Q3, and H302 in Q4.

6. Discussion

Accident prediction and prevention is an urgent topic that needs a long-lasting endeavor in railway systems. Followed by system design and technology development, railway accident prevention has become more complex and uncertain. The railway accident analysis needs to focus more on how to eliminate the factors that may cause accidents effectively and timely, i.e., active accident prevention. Moreover, the characteristics of complex systems determine that the maneuver time left for accident prevention is very limited, and thus there urges for a method to predict accidents in advance, which need to make use of historical data. This paper makes a trial on active accident prevention for railway systems by using the historical information of accident records. At the beginning of the paper, we set our goal to present an effective model for conducting active accident prevention, to explore what our model can achieve in practice, as well as to explore some insights into our approach for practical work. In this section, we will discuss the above issues.

Our first step is to design an RL-based model to achieve active accident prevention by using HI. One of the important steps of this part is to verify that our proposed model is effective. In the previous section, we have verified that the proposed method has the advantage in railway active accident prevention compared with other machine learning methods like RF and FCN. We conclude two difficulties in the scenario of active accident prevention in this paper. Firstly, HI becomes the major in accident-cause prediction. This motivates us to propose a new way to accomplish the prediction task, which can adapt to the complexity of HI, e.g., non-linearity and multi-modal, and learn valid information from the data. We have also tried to simplify the composition of data, but this will cause performance reduction. For example, if we only use part of the numerical data, neglecting the categorical data and text data, the proposed model will fail to learn valid information and always predict the same output, just like how the RF behaves. The second difficulty is the dynamics of the HI, which is explained in Section 4.1. Supervised learning methods like FCN cannot take into account these dynamics during training, thus failing to learn part of the information. The proposed RL-based method can learn the dynamics, as well as find the optimal state of the dynamics, which means that the model can not only predict accident causes effectively but also learn the chance to take preventive actions. This is one of the advantages of using reinforcement learning.

Once the effectiveness of the proposed method is verified, we begin to think about the challenges of the method if used in practice, and we reduce the practical impact to a parameter, denoted as *vpr*, and performed a detailed analysis. On the one hand, the results show a direction to better use the proposed method, i.e., to promote the personal skill of staff to make every correctly predicted accident be avoided as much as possible. On the other hand, the parameter *vpr* quantitatively points out the degree of continued effort needed. For example, if our assessment is reasonable, i.e., the *vpr* for Category is 0.8, then an effort is needed to raise it to higher than 0.9 to make the proposed model in the best state. It is important to note that though we have given the value directly instead of defining precisely how *vpr* is calculated, it is not too difficult to calculate *vpr*. A straightforward way is to let it equal one minus the rate of the happened accidents whose causes are correctly predicted before. This can only be calculated when the proposed model is implemented for a while and the related data is collected.

Finally, some advice is provided based on the analysis of the correctly predicted causes to raise the value of *vpr* in practice. In the case study, we have concluded that some key causes like M411 should be paid much attention on the training of personal skills, and other causes like T111 should be highlighted in some specific periods. Here, we further give some advice on improving the personal skills (for COOK) of staff based on the conclusion mentioned in Section 5.4, as listed in the following.

(1) Every month, the related staff members must: learn the skills in inspecting and fixing coupler defects; learn the related factors affecting the lateral/vertical forces, and skills to eliminate these factors; learn to timely set warnings for highway users when extreme weather occurs, and make sure every highway user gets these warnings; learn the skills in inspecting the defects of train activated warning devices, and learn and be skilled in the standardized operation of train activated warning devices; learn to know all the irregular condition of tracks, especially in the crossing. (According to the results in Fig. 13)
(2) At the beginning of the first quarter, the related staff members must be trained with the skills in: accurately determining the status switches of tracks and replacing damaged switches in a timely manner; operating couplers according to the standardized operation rules. (According to the results in Fig. 15-Q1)
(3) At the beginning of the second quarter, the related staff members must be trained with the skills in: accurately determining the status switches of tracks and replacing damaged switches in a timely manner; knowing the ways to effectively protect the on-track equipment to avoid the on-track equipment be damaged by external forces. (According to the results in Fig. 15-Q2)
(4) At the beginning of the third quarter, the related staff members must be trained with the skills in inspecting the defects of turnout frogs and knowing the condition to replace or fix the turnout frogs. (According to the results in Fig. 15-Q3)
(5) At the beginning of the fourth quarter, the related staff members must be trained with the skills in: quickly and accurately detecting and eliminating foreign objects on or around the rail; quickly inspecting the damage in the rail, and fixing the rail. (According to the results in Fig. 15-Q4)

Note that predicting accidents and causes may be better to consider the occurrence and the information of near-misses because these data are of magnitude greater than the already happened accidents, and can be an important precursor for accident occurrence [24]. However, near-misses are not investigated in many accident reports, including the adopted data in this paper. The use of near-misses to enhance the performance of the proposed method will be one of our attention in future work.

7. Conclusion

This paper proposes an accident prediction and prevention method based on the reinforcement learning model to achieve active accident prevention. This method uses historical information on railway accidents to predict the causes of the accidents that may happen in the future and take active strategies to prevent them in advance. In the case study section, three metrics are designed to compare the proposed method with two baselines (RF and FCN) to verify the effectiveness of the proposed method. The randomness of operation and environment denoting the difficulty of successfully avoiding an accident when knowing its cause is also considered, and *vpr* is proposed and analyzed in detail as the indicator of randomness, which can be a quantitative index to show the direction of the efforts to improve the performance of active accident prevention in practice. Moreover, based on the analysis of the results obtained from the proposed method, we give some advice on improving the personal skills of staff for railway accident prevention.

There are some limitations of the proposed model. First of all, the historical information only takes the numerical data and text data into account without analyzing the contribution of different data to the performance of the proposed model. In future work, we will increase the richness of the historical data by introducing the data from more diverse unstructured data like vision data. Gains of each modal of data to the model will also be made in a detailed comparison. Furthermore, the proposed model does not take the damage of accidents into account. Accidents with larger damage should be prioritized to be prevented. In future work, the damage will be considered as a weighted parameter in the model to enable the model to determine what kind of accidents should put on more time and manpower into prevention.

CRediT authorship contribution statement

Dongyang Yan: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Software, Formal analysis, Investigation, Resources. **Keping Li:** Writing – original draft, Supervision, Project administration, Funding acquisition. **Qiaozhen Zhu:** Validation, Data curation, Visualization, Formal analysis, Writing – review & editing. **Yanyan Liu:** Validation, Resources, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The url of the data source is attached in the manuscript.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (2022JBZY017), the National Natural Science Foundation of China (Grant Nos. 71942006, 72288101), and the Research Foundation of State Key Laboratory of Railway Traffic Control and Safety, China, Beijing Jiaotong University (Grant No.

RCS2021ZT001).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ress.2023.109136.

Appendix. Results Based on the Data from JEFFERSON

References

[1] Abioye OF, Dulebenets MA, Pasha J, Kavoosi M, Moses R, Sobanjo J, et al. Accident and hazard prediction models for highway–rail grade crossings: a state-of-the-practice review for the USA. Railw Eng Sci 2020;28:251–74. https://doi.org/10.1007/s40534-020-00215-w.

[2] Zhou X, Lu P, Zheng Z, Tolliver D, Keramati A. Accident prediction accuracy assessment for highway-rail grade crossings using random forest algorithm compared with decision tree. Reliab Eng Syst Saf 2020;200:106931. https://doi.org/10.1016/j.ress.2020.106931.

[3] Gao L, Lu P, Ren Y. A deep learning approach for imbalanced crash data in predicting highway-rail grade crossings accidents. Reliab Eng Syst Saf 2021;216:108019. https://doi.org/10.1016/j.ress.2021.108019.

[4] Noh B., Yeo H. A novel method of predictive collision risk area estimation for proactive pedestrian accident prevention system in urban surveillance infrastructure 2021. https://doi.org/10.48550/arXiv.2105.02572.

[5] Liu J, Schmid F, Li K, Zheng W. A knowledge graph-based approach for exploring railway operational accidents. Reliab Eng Syst Saf 2021;207:107352. https://doi.org/10.1016/j.ress.2020.107352.

[6] Yang H, Chew DAS, Wu W, Zhou Z, Li Q. Design and implementation of an identification system in construction site safety for proactive accident prevention. Accid Anal Prev 2012;48:193–203. https://doi.org/10.1016/j.aap.2011.06.017.

[7] Mathew J, Benekohal RF. Highway-rail grade crossings accident prediction using zero inflated negative binomial and empirical bayes method. J Safety Res 2021;79:211–36. https://doi.org/10.1016/j.jsr.2021.09.003.

[8] Pasha J, Dulebenets MA, Abioye OF, Kavoosi M, Moses R, Sobanjo J, et al. A comprehensive assessment of the existing accident and hazard prediction models for the highway-rail grade crossings in the state of florida. Sustainability 2020;12:4291. https://doi.org/10.3390/su12104291.

[9] Khan I, Lee E, Khan M. Developing a highway rail grade crossing accident probability prediction model: a north dakota case study. Safety 2018;4:22. https://doi.org/10.3390/safety4020022.

[10] Ma C, Hao W, Xiang W, Yan W. The Impact of aggressive driving behavior on driver-injury severity at highway-rail grade crossings accidents. J Adv Transp 2018;2018:1–10. https://doi.org/10.1155/2018/9841498.

[11] Singh P, Pasha J, Khorram-Manesh A, Goniewicz K, Roshani A, Dulebenets MA. A holistic analysis of train-vehicle accidents at highway-rail grade crossings in Florida. Sustainability 2021;13:8842. https://doi.org/10.3390/su13168842.

[12] Zhao Y, Xu T, Hai-feng W. Text mining based fault diagnosis of vehicle on-board equipment for high speed railway. In: 17th Int. IEEE Conf. Intell. Transp. Syst. ITSC, Qingdao: IEEE; 2014. p. 900–5. https://doi.org/10.1109/ITSC.2014.6957803.

[13] Cao T, Mu W, Gou J, Peng L. A study of risk relevance reasoning based on a context ontology of railway accidents. Risk Anal 2020;40:1589–611. https://doi.org/10.1111/risa.13506.

[14] Wan X, Li Q, Yuan J, Schonfeld PM. Metro passenger behaviors and their relations to metro incident involvement. Accid Anal Prev 2015;82:90–100. https://doi.org/10.1016/j.aap.2015.05.015.

[15] Wang X, Liu J, Khattak AJ, Clarke D. Non-crossing rail-trespassing crashes in the past decade: a spatial approach to analyzing injury severity. Saf Sci 2016;82:44–55. https://doi.org/10.1016/j.ssci.2015.08.017.

[16] Liu X. Statistical temporal analysis of freight train derailment rates in the United States: 2000 to 2012. Transp Res Rec J Transp Res Board 2015;2476:119–25. https://doi.org/10.3141/2476-16.

[17] Kyriakidis M, Hirsch R, Majumdar A. Metro railway safety: an analysis of accident precursors. Saf Sci 2012;50:1535–48. https://doi.org/10.1016/j.ssci.2012.03.004.

[18] Xing Y, Dissanayake S, Lu J, Long S, Lou Y. An analysis of escalator-related injuries in metro stations in China, 2013–2015. Accid Anal Prev 2019;122:332–41. https://doi.org/10.1016/j.aap.2017.10.007.

[19] Ouyang M, Hong L, Yu M-H, Fei Q. STAMP-based analysis on the railway accident and accident spreading: taking the China–Jiaoji railway accident for example. Saf Sci 2010;48:544–55. https://doi.org/10.1016/j.ssci.2010.01.002.

[20] Read GJM, Cox JA, Hulme A, Naweed A, Salmon PM. What factors influence risk at rail level crossings? A systematic review and synthesis of findings using systems thinking. Saf Sci 2021;138:105207. https://doi.org/10.1016/j.ssci.2021.105207.

[21] Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D. Complex networks: structure and dynamics. Phys Rep 2006;424:175–308. https://doi.org/10.1016/j.physrep.2005.10.009.

[22] Wang Z, Bhamra RS, Wang M, Xie H, Yang L. Critical hazards identification and prevention of cascading escalator accidents at metro rail transit stations. Int J Environ Res Public Health 2020;17:3400. https://doi.org/10.3390/ijerph17103400.

[23] Li K, Wang S. A network accident causation model for monitoring railway safety. Saf Sci 2018;109:398–402. https://doi.org/10.1016/j.ssci.2018.06.008.

[24] Wullems C., Dell Yvonne Toft G. Improving the Railway's understanding of accident causation through an integrated approach to human factors analysis and technical safety data recording, 2021. https://doi.org/10.54941/ahfe100653.

[25] Ding X, Yang X, Hu H, Liu Z. The safety management of urban rail transit based on operation fault log. Saf Sci 2017;94:10–6. https://doi.org/10.1016/j.ssci.2016.12.015.

[26] Brown DE. Text mining the contributors to rail accidents. IEEE Trans Intell Transp Syst 2016;17:346–55. https://doi.org/10.1109/TITS.2015.2472580.

[27] Yan F, Zhang S, Majumdar A, Tang T, Ma J. A failure mapping and genealogical research on metro operational incidents. IEEE Trans Intell Transp Syst 2020;21:3551–60. https://doi.org/10.1109/TITS.2019.2932628.

[28] Zulfiqar O, Chang Y-C, Chen P-H, Fu K, Lu C-T, Solnick D, et al. RISECURE: metro incidents and threat detection using social media. In: 2020 IEEEACM Int. Conf. Adv. Soc. Netw. Anal. Min. ASONAM. The Hague, Netherlands: IEEE; 2020. p. 531–5. https://doi.org/10.1109/ASONAM49781.2020.9381317.

[29] Chen H, Jiang B, Ding SX, Huang B. Data-driven fault diagnosis for traction systems in high-speed trains: a survey, challenges, and perspectives. IEEE Trans Intell Transp Syst 2022;23:1700–16. https://doi.org/10.1109/TITS.2020.3029946.

[30] Yan D, Li K, Gu S, Yang L. Network-based bag-of-words model for text classification. IEEE Access 2020;8:82641–52. https://doi.org/10.1109/ACCESS.2020.2991074.

[31] Fan L, Su H, Wang W, Zio E, Zhang L, Yang Z, et al. A systematic method for the optimization of gas supply reliability in natural gas pipeline network based on Bayesian networks and deep reinforcement learning. Reliab Eng Syst Saf 2022;225:108613. https://doi.org/10.1016/j.ress.2022.108613.

[32] Tian Y, Han M, Kulkarni C, Fink O. A prescriptive Dirichlet power allocation policy with deep reinforcement learning. Reliab Eng Syst Saf 2022;224:108529. https://doi.org/10.1016/j.ress.2022.108529.

[33] Yang H, Li W, Wang B. Joint optimization of preventive maintenance and production scheduling for multi-state production systems based on reinforcement learning. Reliab Eng Syst Saf 2021;214:107713. https://doi.org/10.1016/j.ress.2021.107713.

[34] Andriotis CP, Papakonstantinou KG. Managing engineering systems with large state and action spaces through deep reinforcement learning. Reliab Eng Syst Saf 2019;191:106483. https://doi.org/10.1016/j.ress.2019.04.036.

[35] Mohammadi R, He Q. A deep reinforcement learning approach for rail renewal and maintenance planning. Reliab Eng Syst Saf 2022;225:108615. https://doi.org/10.1016/j.ress.2022.108615.

[36] Yang A, Qiu Q, Zhu M, Cui L, Chen W, Chen J. Condition-based maintenance strategy for redundant systems with arbitrary structures using improved reinforcement learning. Reliab Eng Syst Saf 2022;225:108643. https://doi.org/10.1016/j.ress.2022.108643.

[37] Ying C, Chow AHF, Chin K-S. An actor-critic deep reinforcement learning approach for metro train scheduling with rolling stock circulation under stochastic demand. Transp Res Part B Methodol 2020;140:210–35. https://doi.org/10.1016/j.trb.2020.08.005.

[38] Šemrov D, Marsetič R, Žura M, Todorovski L, Srdic A. Reinforcement learning approach for train rescheduling on a single-track railway. Transp Res Part B Methodol 2016;86:250–67. https://doi.org/10.1016/j.trb.2016.01.004.

[39] Jiang Z, Fan W, Liu W, Zhu B, Gu J. Reinforcement learning approach for coordinated passenger inflow control of urban rail transit in peak hours. Transp Res Part C Emerg Technol 2018;88:1–16. https://doi.org/10.1016/j.trc.2018.01.008.

[40] Wang H., Han Z., Liu W., Wu Y. A reinforcement learning-based pantograph control strategy for improving current collection quality in high-speed railways. IEEE Trans NEURAL Netw Learn Syst n.d. https://doi.org/10.1109/TNNLS.2022.3219814.

[41] Yoshida Y, Arai S, Kobayashi H, Kondo K. Charge/discharge control of wayside batteries via reinforcement learning for energy-conservation in electrified railway systems. Electr Eng Jpn 2021;214. https://doi.org/10.1002/eej.23319.

[42] Jung H. An optimal charging and discharging scheduling algorithm of energy storage system to save electricity pricing using reinforcement learning in urban railway system. J Electr Eng Technol 2022;17:727–35. https://doi.org/10.1007/s42835-021-00881-8.

[43] Yang Z, Zhu F, Lin F. Deep-reinforcement-learning-based energy management strategy for supercapacitor energy storage systems in urban rail transit. IEEE Trans Intell Transp Syst 2021;22:1150–60. https://doi.org/10.1109/TITS.2019.2963785.

[44] Deng K, Liu Y, Hai D, Peng H, Lowenstein L, Pischinger S, et al. Deep reinforcement learning based energy management strategy of fuel cell hybrid railway vehicles considering fuel cell aging. ENERGY Convers Manag 2022;251. https://doi.org/10.1016/j.enconman.2021.115030.

[45] Xu J., Ai B., Chen L., Cui Y., Wang N. Deep reinforcement learning for computation and communication resource allocation in multiaccess mec assisted railway IoT networks. IEEE Trans Intell Transp Syst n.d. https://doi.org/10.1109/TITS.2022.3205175.

[46] Xu J, Ai B. When mmWave high-speed railway networks meet reconfigurable intelligent surface: a deep reinforcement learning method. IEEE Wirel Commun Lett 2022;11:533–7. https://doi.org/10.1109/LWC.2021.3135602.

[47] Cui L, Yuan Z, Ming Z, Yang S. Improving the congestion control performance for mobile networks in high-speed railway via deep reinforcement learning. IEEE Trans Veh Technol 2020;69:5864–75. https://doi.org/10.1109/TVT.2020.2984038.

[48] Xu J, Ai B. Deep Reinforcement Learning for Handover-Aware MPTCP Congestion Control in Space-Ground Integrated Network of Railways. IEEE Wirel Commun 2021;28:200–7. https://doi.org/10.1109/MWC.001.2100116.

[49] Yu C, Quan W, Liu K, Liu M, Xu Z, Zhang H. Deep reinforcement learning-based fountain coding for concurrent multipath transfer in high-speed railway networks. PEER–PEER Netw Appl 2022;15:2744–56. https://doi.org/10.1007/s12083-022-01321-8.

[50] Gao T, Li Z, Gao Y, Schonfeld P, Feng X, Wang Q, et al. A deep reinforcement learning approach to mountain railway alignment optimization. Comput-AIDED Civ Infrastruct Eng 2022;37:73–92. https://doi.org/10.1111/mice.12694.

[51] Khadilkar H. A scalable reinforcement learning algorithm for scheduling railway lines. IEEE Trans Intell Transp Syst 2019;20:727–36. https://doi.org/10.1109/TITS.2018.2829165.

[52] Li K, Yan D, Liu Y, Zhu Q. A network-based feature extraction model for imbalanced text data. Expert Syst Appl 2022;195:116600. https://doi.org/10.1016/j.eswa.2022.116600.

[53] Zhang P, Huang X, Wang Y, Jiang C, He S, Wang H. Semantic similarity computing model based on multi model fine-grained nonlinear fusion. IEEE Access 2021;9:8433–43. https://doi.org/10.1109/ACCESS.2021.3049378.

[54] Acheampong FA, Nunoo-Mensah H, Chen W. Transformer models for text-based emotion detection: a review of BERT-based approaches. Artif Intell Rev 2021;54:5789–829. https://doi.org/10.1007/s10462-021-09958-2.

[55] Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R. Indexing by latent semantic analysis. J Am Soc Inf Sci 1990;41:391–407. https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9.