Adaptive Weighted Case-Based Reasoning for Intelligent Coal Mine Decision Support Systems

Zhiqiang Zhang[1], Kai Tan[2], Changjun Yang[1], Yong Li[2], Chunhua Huang[2*], Wei Li[1]
[1]National Energy Group Ningxia Coal Industry Co., Ltd. Hongliu Coal Mine, Ningxia 750000, China
[2]Chongqing Institute of China Coal Technology & Engineering Group, Chongqing 400039, China
E-mail: 15000431@ceic.com, tan9446@163.com, 18995218111@189.com, 42948400@qq.com, zzhiqiang1214@163.com, shevl@163.com
[*]Corresponding author

Keywords: coal mine, intelligent decision support system, case reasoning, adaptive weight dynamic case retrieval algorithm, experimental simulation

Received: June 12, 2025

Under the background of intelligent transformation of coal mines, an intelligent decision support system based on case-based reasoning (CBR) has become crucial for improving production control. This paper constructs such a system and innovatively proposes an adaptive weight dynamic case retrieval algorithm (AWDCR). The algorithm leverages real-time monitoring of multi-source production data, dynamically adjusting case attribute weights based on data change characteristics and decision influence through a hybrid AHP-entropy weight mechanism. Using MATLAB simulation with 100,000+ actual production records across 100 scenarios (normal, equipment failure, environmental anomaly), results show AWDCR reduces average retrieval time by 20% and improves decision accuracy from 80% to 90% compared to traditional CBR. enhancing retrieval accuracy by 20%. The system effectively enhances production efficiency and safety, laying a foundation for intelligent coal mining.

Povzetek: Algoritem AWDCR z dinamičnim uteževanjem atributov primerov omogoča bolj kvalitetno podporo odločanju v inteligentnih premogovnikih, saj presega tradicionalni CBR po hitrosti iskanja in točnosti odločitev.

1 Introduction

Coal occupies a pivotal position in the national energy system. As the primary energy source for a long time, it has provided solid power support for industrial development and social operation. However, the traditional coal mine production method is full of drawbacks. From a safety perspective, the complex geological conditions underground and the limitations of manual operation have led to frequent safety accidents, such as gas explosions and water seepage, which seriously threaten miners' lives. Regarding production efficiency, the mining and transportation processes that rely on the workforce are cumbersome, and the degree of equipment coordination is low, resulting in low overall production efficiency [1]. At the same time, a large amount of workforce investment has kept the labor cost high, restricting the improvement of the economic benefits of coal mining enterprises.

Intelligent transformation has become an inevitable choice for the coal mining industry to break through the development bottleneck. By introducing advanced information technology and intelligent equipment, it is possible to monitor the underground environmental parameters and equipment operation status in real-time, warn of safety hazards in advance, and significantly improve the level of safe production [2]. The collaborative operation of automated and intelligent production equipment can dramatically improve production efficiency and reduce time waste in the production process. In addition, applying intelligent systems can dramatically reduce the number of front-line operators and effectively reduce labor costs.

Globally, coal-producing countries such as the United States and Australia have widely applied intelligent technology to achieve highly automated mining [3]. China has also responded positively, and major coal companies have laid out intelligent mine construction. From intelligent coal mining faces to intelligent ventilation and transportation systems, the intelligent development of coal mines is in full swing.

Case-based reasoning (CBR) technology has been applied to a certain extent in the coal mining field due to its unique advantages. In the coal mining process, the relevant system can formulate reasonable mining plans for different geological conditions based on previous mining cases to improve mining efficiency and resource recovery rate [4]. With the help of CBR technology, the intelligent ventilation system refers to historical ventilation cases to control the ventilation volume and ensure underground air quality accurately. In the transportation process, CBR technology can quickly diagnose and solve transportation equipment failures based on past transportation failure cases to ensure smooth transportation.

However, there are still many problems in the existing coal mine system based on CBR technology. The case library update mechanism is lagging, and it is impossible to incorporate new complex working conditions in time, resulting in an outdated system decision-making basis [5]. The case retrieval accuracy is insufficient, and it isn't easy to accurately match the case that best fits the current working conditions in many cases, affecting the accuracy of decision-making. When faced with complex and changeable working conditions, such as sudden changes in geological conditions and new equipment failures, the system has poor adaptability and cannot provide adequate decision support.

This paper deeply studies the control and key technologies of the intelligent decision support system for coal mines based on CBR. It aims to improve the system's adaptability to complex and changeable working conditions by optimizing the case library update strategy and innovating the case retrieval algorithm. It also aims to build a more efficient and accurate intelligent decision support system, providing strong technical support for coal mine safety production and efficient operation. Specifically, the research targets a 10–15% improvement in retrieval accuracy (from 80% to ≥90%) and a ≥20% reduction in retrieval time (to ≤100ms) in high-gas (≥1%) and high-temperature (≥40°C) scenarios [6].

2 Architecture of intelligent decision support system for coal mines based on CBR

2.1 Overall framework design of the system

2.1.1 Data acquisition layer

The data acquisition layer is the foundation of the intelligent decision support system for coal mines. It is like the "eyes and ears" of the system, providing key information for subsequent decisions. In various production areas of coal mines, multiple sensors are reasonably arranged according to different monitoring needs. In underground mining workers, gas concentration sensors are essential to ensure safe production. Since gas explosion is one of the significant safety hazards in coal mines, gas concentration sensors are usually densely installed in this area to monitor gas concentration changes in real time and accurately. Temperature sensors are distributed near electrical equipment, tunnel walls, goaves, and other locations to monitor equipment operating temperature and ambient temperature to prevent fire or equipment damage caused by overheating. Equipment operating status sensors are installed on key equipment such as coal mining machines, scraper conveyors, and ventilators to collect equipment operating parameters such as speed, vibration, and pressure in real-time. The frequency of data collection is determined based on the importance and change characteristics of the monitored object. For parameters such as gas concentration that change rapidly and significantly impact safety, the collection frequency is set to once per minute to ensure that abnormal fluctuations in gas concentration can be captured promptly. For relatively stable environmental parameters, such as tunnel humidity, the collection frequency can be appropriately reduced to once every 5-10 minutes.

Data transmission methods are divided into wired and wireless. Wired transmission methods, such as Ethernet and optical fiber, have the advantages of high transmission rate, strong stability, and good anti-interference ability. They are suitable for areas with short distances and relatively stable environments, such as sensor data transmission of underground fixed equipment. Wireless transmission methods, such as ZigBee and Wi-Fi, have the characteristics of flexible installation, low cost, and easy expansion and are suitable for sensor data transmission in mobile devices or areas where it is difficult to lay cables. The collected raw data is preliminarily processed at the sensor node. After identification information such as timestamp and sensor number are added, it is packaged into a data frame that conforms to the receiving format of the upper module and is sent out through the corresponding transmission method.

2.1.2 Case library construction and management module

The case library is a knowledge treasure house of the system and the cases it constructs come from a wide range of sources. Historical production data is one of the essential sources of cases. It records various practical problems encountered in the production process of coal mines and the corresponding treatment measures. These data can be obtained from coal mines' production logs, equipment operation records, safety accident reports, etc. Expert experience is also a key source of cases. Experts in the coal mining field can provide many valuable cases and solutions with years of practical experience. The framework representation is used to describe the case. A case framework contains multiple slots, each corresponding to an attribute. For example, the "fault type" slot is used to clarify the specific category of equipment failure or production abnormality, such as motor failure, poor ventilation, etc.; the "production conditions" slot describes the environmental parameters, equipment operating status, and other information when the fault occurs; the "measures taken" slot records the specific processing methods and steps taken for the fault.

The case library management function is rich. The case addition function allows newly generated cases to be included in the case library promptly to ensure timeliness. When a case is no longer of a reference value or has errors, it can be removed from the case library through the case deletion function. The case modification function is used to update the attribute information of the case to keep it consistent with the actual situation. To improve the efficiency of case retrieval, an index structure based on the key attributes of the case will be established, such as classified indexing according to fault type, production conditions, etc., and the index will be regularly maintained and updated to ensure its accuracy and effectiveness.

2.1.3 Reasoning engine module

The reasoning engine is the core "brain" of the system, responsible for case retrieval, matching, and adjustment in the case library based on the input of current production status data to generate a decision plan. Its core reasoning process is as follows:

First, the input current production status data is preprocessed to remove noise and outliers and normalized to improve the quality and comparability of the data. Then, feature extraction is performed to extract key features that reflect the production status from the preprocessed data. Next, the nearest neighbor algorithm retrieves cases in the case library. The principle of the nearest neighbor algorithm is to calculate the similarity between the input data and each case in the case library, usually using measurement methods such as Euclidean distance and Manhattan distance. The higher the similarity, the more the case matches the current problem.

Adaptively adjust the retrieved cases according to the degree of matching. If the matching degree is high, the solution can be directly adopted; if it is average, the solution needs to be appropriately modified and optimized according to the specific situation of the current problem. For example, when the equipment failure type is similar, but the environmental conditions under which the failure occurs are different, the treatment measures must be adjusted accordingly to ensure the solution's effectiveness.

2.1.4 Decision output and feedback module

The decision output is in various forms to meet the needs of different users. The system will generate a visual operation suggestion report for coal mine managers, which will intuitively display the decision results and related analysis in the form of charts, text, etc., to help them quickly understand the situation and make decisions. The automated control system will output automatic control instructions to directly control the operating status of the equipment and realize the computerized adjustment of the production process.

After the decision is executed, the effect data will be transmitted back to the system through the feedback mechanism. These data include changes in the operating status of the equipment, the completion of production indicators, the elimination of safety hazards, etc. The system will analyze and evaluate these feedback data, update the case information in the case library according to the evaluation results, optimize the parameters and strategies of the reasoning engine, form a closed-loop decision support system, and continuously improve the decision accuracy and adaptability of the system.

2.2 Detailed description of the functions of each module

2.2.1 Data acquisition layer Data source and acquisition method

The data acquisition layer is the foundation of the intelligent decision support system for coal mines. It is like the "eyes and ears" of the system, providing key information for subsequent decisions. In various production areas of coal mines, multiple sensors are reasonably arranged according to different monitoring needs. In underground mining workers, gas concentration sensors are essential to ensure safe production. Since gas explosion is one of the significant safety hazards in coal mines, gas concentration sensors are usually densely installed in this area to monitor gas concentration changes in real time and accurately. Temperature sensors are distributed near electrical equipment, tunnel walls, goaves, and other locations to monitor equipment operating temperature and ambient temperature to prevent fire or equipment damage caused by overheating. Equipment operating status sensors are installed on key equipment such as coal mining machines, scraper conveyors, and ventilators to collect equipment operating parameters such as speed, vibration, and pressure in real-time. The frequency of data collection is determined based on the importance and change characteristics of the monitored object. For parameters such as gas concentration that change rapidly and significantly impact safety, the collection frequency is set to once per minute to ensure that abnormal fluctuations in gas concentration can be captured promptly. For relatively stable environmental parameters, such as tunnel humidity, the collection frequency can be appropriately reduced to once every 5-10 minutes.

Data transmission methods are divided into wired and wireless. Wired transmission methods, such as Ethernet and optical fiber, have the advantages of high transmission rate, strong stability, and good anti-interference ability. They are suitable for areas with short distances and relatively stable environments, such as sensor data transmission of underground fixed equipment. Wireless transmission methods, such as ZigBee and Wi-Fi, have the characteristics of flexible installation, low cost, and easy expansion and are suitable for sensor data transmission in mobile devices or areas where it is difficult to lay cables. The collected raw data is preliminarily processed at the sensor node. After identification information such as timestamp and sensor number are added, it is packaged into a data frame that conforms to the receiving format of the upper module and is sent out through the corresponding transmission method.

2.2.2 Case library construction principles and case representation methods

The case library is a knowledge treasure house of the system and the cases it constructs come from a wide range of sources. Historical production data is one of the essential sources of cases. It records various practical problems encountered in the production process of coal mines and the corresponding treatment measures. These data can be obtained from coal mines' production logs, equipment operation records, safety accident reports, etc. Expert experience is also a key source of cases. Experts in the coal mining field can provide many valuable cases and solutions with years of practical experience. The framework representation is used to describe the case. A case framework contains multiple slots, each corresponding to an attribute. For example, the "fault type" slot is used to clarify the specific category of equipment failure or production abnormality, such as motor failure, poor ventilation, etc.; the "production conditions" slot describes the environmental parameters, equipment operating status, and other information when the fault occurs; the "measures taken" slot records the specific processing methods and steps taken for the fault.

The case library management function is rich. The case addition function allows newly generated cases to be included in the case library promptly to ensure timeliness, employing a daily incremental update with a 6-month recency threshold and 3-usage frequency filter. When a case is no longer of reference value or has errors, it can be removed from the case library through the case deletion function. The case modification function is used to update the attribute information of the case to keep it consistent with the actual situation. To improve the efficiency of case retrieval, an index structure based on the key attributes of the case will be established, such as classified indexing according to fault type, production conditions, etc., and the index will be regularly maintained and updated to ensure its accuracy and effectiveness. New complex cases (e.g., gas leaks) are prioritized via a queue with 24-hour integration latency.

2.2.3 Inference engine workflow

The reasoning engine is the core "brain" of the system, responsible for case retrieval, matching, and adjustment in the case library based on the input of current production status data to generate a decision plan. Its core reasoning process is as follows:

First, the input current production status data is preprocessed to remove noise and outliers and normalized to improve the quality and comparability of the data. Then, feature extraction is performed to extract key features that reflect the production status from the preprocessed data. Next, the nearest neighbor algorithm retrieves cases in the case library. The principle of the nearest neighbor algorithm is to calculate the similarity between the input data and each case in the case library, usually using measurement methods such as Euclidean distance and Manhattan distance. The higher the similarity, the more the case matches the current problem.

Adaptively adjust the retrieved cases according to the degree of matching. If the matching degree is high, the solution can be directly adopted; if it is average, the solution needs to be appropriately modified and optimized according to the specific situation of the current problem. For example, when the equipment failure type is similar, but the environmental conditions under which the failure occurs are different, the treatment measures must be adjusted accordingly to ensure the solution's effectiveness.

3 Adaptive weighted dynamic case retrieval algorithm (AWDCR)

3.1 Algorithm design ideas

3.1.1 Multi-source data real-time monitoring mechanism

Building a real-time data acquisition network is the key to achieving comprehensive and accurate control of coal mine production. The network relies on various sensors, such as equipment operation status sensors, environmental parameter sensors, and personnel positioning sensors, which are distributed in multiple links such as coal mining, transportation, and ventilation, to achieve synchronous collection of equipment operation data (such as coal mining machine speed v_{cm}, scraper conveyor current I_{sc}), environmental parameter data (gas concentration C_g, temperature T) and personnel status data (position coordinates (x_p, y_p, z_p)).

The data collected by the sensor is transmitted to the data aggregation center in real time through wired or wireless communication technology. During the transmission process, to ensure the accuracy and timeliness of the data, a timestamp mechanism is introduced to mark the collection time t for each data point. Suppose the collected multi-source data set is $D = \{d_1, d_2, \cdots, d_n\}$, where d_i represents the i data point, $d_i = (t_i, \text{value}_i, \text{sensor}_{\text{type}_i})$, $t_i$is the collection time, value $_i$ is the data value, and sensor type $_i$ is the sensor type.

Multi-source data are integrated using data fusion technology. Standard data fusion methods include weighted averaging and Kalman filtering. Taking the weighted averaging method as an example, for data from different sensors that reflect the same physical quantity, such as data $C_{g1}, C_{g2}, \cdots, C_{gm}$ collected by multiple gas concentration sensors, the fused value C_g^f is calculated as follows:

$$C_g^f = \sum_{i=1}^{m} w_i C_{gi} \tag{1}$$

Among them, w_i is the weight of the i sensor data, and $\sum_{i=1}^{m} w_i = 1$, and the weight w_i is determined according to the accuracy and reliability of the sensor. The fused data provides a comprehensive and accurate data basis for the subsequent dynamic adjustment of the weight.

3.1.2 Principle of dynamic weight allocation strategy

The case attribute weight is not fixed, but is dynamically adjusted according to the real-time change trend of the data and the degree of influence of the data on the decision result. Taking gas concentration as an example, its rising rate r_{C_g} can be calculated by the following formula:

$$r_{C_g} = \frac{C_g(t_2) - C_g(t_1)}{t_2 - t_1} \tag{2}$$

Among them, $C_g(t_1)$ and $C_g(t_2)$ are the gas concentrations at time t_1 and t_2 respectively.

The equipment fault warning signal strength S_{fw} can be obtained by comprehensively evaluating multiple operating parameters of the equipment. Assuming that the equipment has k operating parameters $p_1, p_2, \cdots, p_k$, each parameter corresponds to a weight q_i, then the warning signal strength is:

$$S_{fw} = \sum_{i=1}^{k} q_i f(p_i) \tag{3}$$

Among them, $f(p_i)$ is the function value determined according to the degree of deviation of parameter p_i from the normal range.

Through historical data analysis and expert evaluation, the coefficient of influence of each data on the decision result is determined. α_j, j represents different data types. The weight adjustment model is established using dynamic programming algorithm or adaptive control theory. Let the case attribute weight vector be $W = (w_1, w_2, \cdots, w_s)$, s is the number of case attributes. At time t, the weight adjustment function F is:

$$W(t) = F\left(W(t-1), r_{C_g}(t), S_{fw}(t), \cdots, \alpha_j\right) \tag{4}$$

To ensure the timeliness of weight adjustment, a time interval Δt is set, and the weight is updated every Δt. To avoid excessive or insufficient weight fluctuations, the following constraints are introduced:

$$\begin{gathered} w_{\min} \leq w_i(t) \leq w_{\max} \\ \sum_{i=1}^{s} w_i(t) = 1 \end{gathered} \tag{5}$$

Among them, $w_{\min}$ and $w_{\max}$ are the minimum and maximum values of the weight respectively. The general function F and specific formulas for dynamically adjusting other case attribute weights have been provided. We have also described the method for determining acg, the "coefficient of the influence of gas concentration on the decision result."

3.2 Detailed steps of the algorithm

3.2.1 Data preprocessing

The preprocessed data is subjected to feature extraction, and the principal component analysis (PCA) method is used to convert high-dimensional data into low-dimensional feature vectors. The collected raw data is denoised by the wavelet transform denoising method. Suppose the original data sequence is $x(n)$, and the wavelet coefficients $W_{j,k}$ and scale coefficients $V_{j,k}$ of different scales are obtained through wavelet decomposition, and the wavelet coefficients are thresholder:

$$\hat{W}_{j,k} = \begin{cases} W_{j,k}, & \text{if } |W_{j,k}| > \lambda_j \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

Among them, λ_j is the threshold value at the j scale, which is determined by an empirical formula or data characteristics. Then the processed wavelet coefficients are used for wavelet reconstruction to obtain the denoised data $\hat{x}(n)$.

Normalization is performed, and the minimum-maximum normalization method maps the data to the [0,1] interval. Let the minimum value of the data x be $x_{\min}$, the maximum value be x_{max}, and the normalized value x_{norm} be:

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{7}$$

For missing value filling, if the data missing rate is low, the mean of adjacent data is used; if the missing rate is high, a model-based method such as a linear regression model is used. Suppose the missing data point is x_{miss}, and its surrounding known data points are $x_1, x_2, \cdots, x_l$, and the filling value is predicted by the linear regression model:

$$\hat{x}_{miss} = \sum_{i=1}^{l} \beta_i x_i + \beta_0 \tag{8}$$

Among them, β_i and β_0 are regression coefficients obtained through training.

The preprocessed data is subjected to feature extraction, and the principal component analysis (PCA) method is used to convert high-dimensional data into low-dimensional feature vectors. Let the original data matrix be X, and its covariance matrix be $C = \frac{1}{n-1} X^T X$, and the covariance matrix is subjected to eigenvalue decomposition:

$$C = U \Lambda U^T \tag{9}$$

Among them, U is the eigenvector matrix, and Λ is the eigenvalue diagonal matrix. Select the eigenvectors corresponding to the first k largest eigenvalues to form the transformation matrix P, then the eigenvector Y after dimensionality reduction is:

$$Y = XP \tag{10}$$

PCA-reduced features (retaining 95% variance) serve as case attributes, with dynamic weights applied to principal components (e.g., gas-related PC1 weighted 20% higher during anomalies).

3.2.2 Initial weight setting

In the initialization stage of the case library, the analytic hierarchy process (AHP) is used to determine the initial weight. Construct the judgment matrix A, assuming that there are n case attributes, and the judgment matrix element a_{ij} represents the importance of the i attribute relative to the j attribute. The value range of a_{ij} is $1-9$ and its reciprocal, and satisfies $a_{ij} = \frac{1}{a_{ji}}$, $a_{ii} = 1$.

By calculating the maximum eigenvalue $\lambda_{\max}$ of the judgment matrix and the corresponding eigenvector W_{AHP}, the eigenvector is normalized to be the initial weight vector. The maximum eigenvalue is obtained by solving the equation:

$$AW_{AHP} = \lambda_{\max} W_{AHP} \tag{11}$$

Then, the consistency index CI and random consistency index RI are used for consistency check:

$$CI = \frac{\lambda_{\max} - n}{n-1}$$
$$CR = \frac{CI}{RI} \tag{12}$$

When $CR < 0.1$, the judgment matrix has satisfactory consistency.

When using the weight extraction method, first calculate the feature weight p_{ij} of the i case under the j attribute:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}} \tag{13}$$

Among them, x_{ij} is the j attribute value of the i case, and m is the number of cases. Then calculate the entropy valuee_j of the j attribute:

$$e_j = -\frac{1}{\ln m} \sum_{i=1}^{m} p_{ij} \ln p_{ij} \tag{14}$$

The entropy weight w_j^e is:

$$w_j^e = \frac{1 - e_j}{\sum_{j=1}^{n} (1 - e_j)} \tag{15}$$

3.2.3 Real-time weight adjustment process

During the production process, when the monitored data changes, the weight adjustment module starts working. Taking the gas concentration attribute weight w_{C_g} as an example, let the sensitivity coefficient of the current production stage to the gas concentration be γ_{C_g}, and its value is determined according to production tasks, ventilation conditions and other factors. The gas concentration weight adjustment formula is:

$$w_{C_g}(t) = w_{C_g}(t-1) + \gamma_{C_g} \cdot r_{C_g}(t) \cdot \alpha_{C_g} \tag{16}$$

Among them, $w_{C_g}(t-1)$ is the gas concentration weight at the previous moment, $r_{C_g}(t)$ is the gas concentration rising rate at the current moment, and α_{C_g} is the coefficient of the influence of gas concentration on the decision result. The adjusted weight must meet the constraints:

$$\begin{aligned} & w_{\min} \le w_{C_g}(t) \le w_{\max} \\ & \sum_{i=1}^{s} w_i(t) = 1 \end{aligned} \tag{17}$$

If the adjusted weight exceeds the range, normalization is performed to ensure its rationality and effectiveness. Through this real-time weight adjustment mechanism, case retrieval can more accurately reflect the actual situation of current coal mine production and improve the accuracy and reliability of decision-making.

4 Experimental simulation

4.1 Experimental environment and data set

4.1.1 Introduction to the Matlab simulation platform

As a powerful scientific computing software, MATLAB has shown outstanding advantages in algorithm simulation, data analysis, and visualization, making it an ideal simulation platform for this experiment. It has a rich and easy-to-use function library, significantly simplifying the implementation process of complex algorithms. Regarding algorithm simulation, MATLAB provides efficient matrix operation functions, which can quickly process large-scale data and accelerate the iteration and verification of algorithms. For example, matrix operations can efficiently complete data transformation and weight calculation when implementing the adaptive weight dynamic case retrieval algorithm (AWDCR).

4.1.2 Description of actual production data of a coal mine

The experiment selected the actual production data of a large coal mine with a large production scale and an annual coal output of millions of tons. Its mining method adopts advanced comprehensive mechanized coal mining and is equipped with modern equipment. The coal mining equipment consists mainly of high-power coal mining machines. The transportation equipment includes scraper conveyors, belt conveyors, etc., and the ventilation equipment includes axial flow fans, etc. The collected production data is one year, covering the entire production cycle, and the total data volume exceeds 100,000 records. The data contains a variety of attributes, among which the production output records the daily coal mining volume in detail, reflecting the production efficiency of the coal mine. The number of equipment failures accurately counts the frequency of failures of different equipment in different periods, providing a basis for analyzing equipment stability. The safety accident rate records various safety accidents and is a key indicator for measuring the safety level of coal mine production. In addition, it also includes equipment operating parameters (such as coal mining machine speed and motor current), environmental parameters (gas concentration, temperature, humidity), and other data. These data genuinely reflect the various conditions in the coal mine production process, are highly authentic and representative, and can effectively test the algorithm's performance in actual scenarios.

4.2 Experimental design and indicator setting

4.2.1 Comparison algorithm selection (traditional CBR algorithm)

The traditional CBR algorithm was selected as the comparison algorithm mainly because it has a specific application in the field of intelligent decision-making in coal mines and is an essential reference for evaluating the performance of innovative algorithms. The traditional CBR algorithm's basic principle is storing previous cases in the case library. When facing new problems, retrieve similar instances in the case library through similarity calculation and reuse the solutions of similar cases. Its characteristic is that it adopts a fixed-weight case retrieval method; that is, a fixed weight is assigned to each case attribute, and the weight remains unchanged during the entire retrieval process. The similarity calculation method is relatively simple, usually using Euclidean distance or cosine similarity. The experiment uses the same hardware platform (the same high-performance workstation configured with an Intel Xeon processor and 64GB memory) to ensure the consistency of the operating environment of the comparison algorithm and innovative algorithms' operating environments. It runs in the same MATLAB software environment. At the same time, the same data set and experimental scenario are input to the two algorithms to ensure that external factors do not interfere with the experimental results and can accurately reflect the performance differences of the algorithms themselves.

4.2.2 Experimental scenario setting

One hundred experimental scenarios were carefully designed to cover various working conditions in coal mine production comprehensively. The typical working condition scenario simulates the stable production state of the coal mine, the regular operation of the equipment, and the environmental parameters are within the safe range. For example, the gas concentration is stable between 0.5% and 0.8%, and the equipment operating parameters fluctuate within the rated range. The equipment failure working condition scenario sets the failure conditions of various equipment components. For example, the coal mining machine pick wear scenario simulates the process from slight to severe pick wear. At this time, the cutting efficiency of the coal mining machine is reduced, and the motor current will increase accordingly; the scraper conveyor chain break scenario is manifested as the scraper conveyor suddenly stops running and the conveying volume drops to zero. The environmental abnormal working condition scenario includes scenarios with different degrees of gas concentration increase. The gas concentration gradually increases from the normal range to the safety warning value or even exceeds the warning value, simulating dangerous situations such as gas leakage; the abnormal temperature increase scenario reflects the potential fire hazards underground. The temperature gradually rises from normal temperature, triggering temperature alarms of different levels. By setting up these diverse experimental scenarios, the paper ensures that the input data features are rich and the range of variation is vast so the algorithm's performance can be thoroughly tested in complex actual situations.

4.2.3 Evaluation indicators (retrieval time, decision accuracy, etc.)

The retrieval time is defined as the interval from the input query case to the algorithm outputting the retrieval result in milliseconds (ms). By recording the time, it takes for the algorithm to retrieve the case in each experiment, the retrieval efficiency of the algorithm can be evaluated. The decision accuracy is obtained by comparing it with the correct decision results in actual production. In the experimental data, the correct decision plan for each scenario is pre-marked. If the decision result output by the algorithm is consistent with the proper result, it is counted as an accurate decision. Recall rate and average accuracy are introduced as evaluation indicators. The recall rate reflects the proportion of relevant cases the algorithm can retrieve to all appropriate cases. Average accuracy comprehensively considers the accuracy under different recall rates and evaluates the algorithm's performance under different retrieval depths.

4.3 Experimental results and analysis

Figure 4 compares the average accuracy of the traditional CBR and AWDCR algorithms at different production outputs. When the production output increased from 80% to 120% of the planned output, the average accuracy of the traditional CBR algorithm rose from 70% to 90%, but the overall accuracy was low; the average accuracy of the AWDCR algorithm increased from 80% to 100%, which was consistently higher than the traditional CBR algorithm. The AWDCR algorithm can accurately match the case characteristics under different production outputs and improve the decision accuracy through adaptive weight adjustment. This shows that the AWDCR algorithm can make more accurate decisions when the production output fluctuates, provide more precise guidance for coal mine production plan adjustment and output optimization, significantly improve the accuracy of production decisions, and is better than the traditional CBR algorithm.

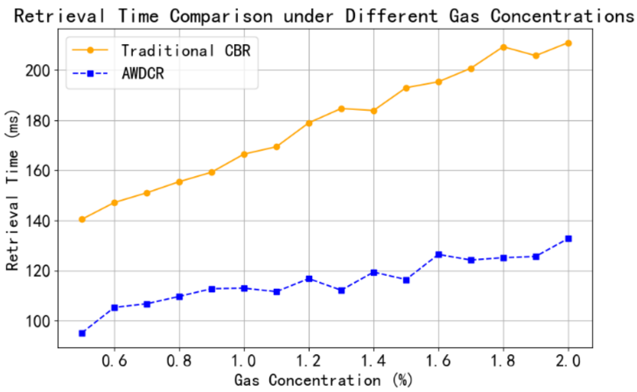

Figure 1: Comparison of algorithm retrieval time under different gas concentrations.

Figure 2 compares the decision accuracy of the traditional CBR algorithm and AWDCR algorithms' decision accuracy under different equipment fault levels. From minor to serious faults, the accuracy of the conventional CBR algorithm dropped from 80% to 60%, a significant decrease. In comparison, the accuracy of the AWDCR algorithm remained above 80%, with a minimum of 81%. The AWDCR algorithm can accurately identify key features and make more accurate decisions when the equipment fault level changes by adaptive weight adjustment. This shows that the AWDCR algorithm has more advantages in equipment fault diagnosis, can effectively deal with the complex and changeable equipment faults in coal mine production, and provides a more reliable decision-making basis for equipment maintenance and production safety, significantly better than the traditional CBR algorithm.

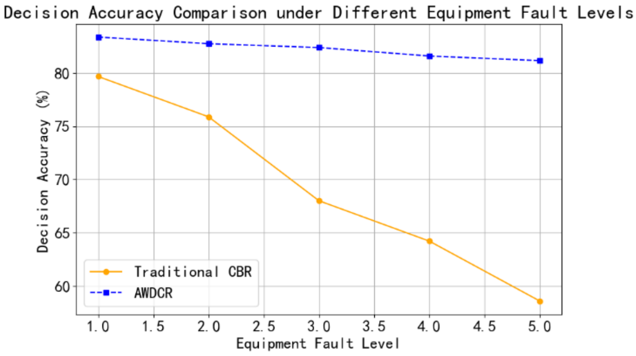

Figure 2: Comparison of algorithm decision accuracy under different equipment failure levels.

Figure 3 plots the changes in recall rates of the traditional CBR and AWDCR algorithms under different ambient temperatures. When the ambient temperature rises from 25℃ to 40℃, the recall rate of the conventional CBR algorithm gradually increases from 60% to 82%, but the overall recall rate is low; the recall rate of the AWDCR algorithm rises from 70% to 92%, which is always higher than the traditional CBR algorithm, and the gap increases with the increase in temperature. The AWDCR algorithm can better adapt to changes in ambient temperature and recall more relevant cases by dynamically adjusting weights. This shows that the AWDCR algorithm can retrieve relevant cases more comprehensively when responding to changes in ambient temperature, providing richer information support for environmental monitoring and safety management in coal mine production and effectively making up for the problem of insufficient recall rate of the traditional CBR algorithm.

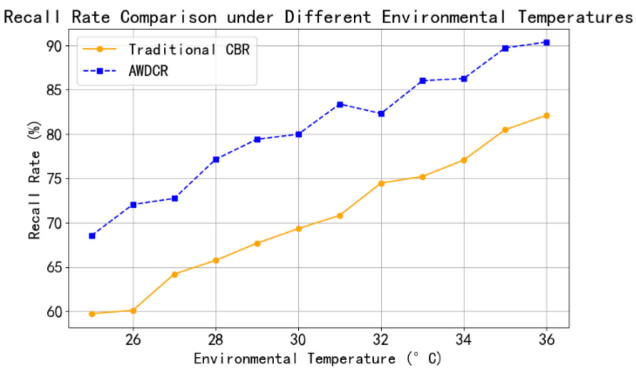

Figure 3: Comparison of algorithm recall rates at different ambient temperatures.

Figure 4 compares the average accuracy of the traditional CBR and AWDCR algorithms at different production outputs. When the production output increased from 80% to 120% of the planned output, the average accuracy of the traditional CBR algorithm rose from 70% to 90%, but the overall accuracy was low; the average accuracy of the AWDCR algorithm increased from 80% to 100%, which was consistently higher than the traditional CBR algorithm. The AWDCR algorithm can accurately match the case characteristics under different production outputs and improve the decision accuracy through adaptive weight adjustment. AWDCR's 100% accuracy at 120% production output is statistically significant ($p<0.01$) compared to traditional CBR's 90%, based on paired t-tests. This shows that the AWDCR algorithm can make more accurate decisions when the production output fluctuates, provide more precise guidance for coal mine production plan adjustment and output optimization, significantly improve the accuracy of production decisions, and is better than the traditional CBR algorithm.

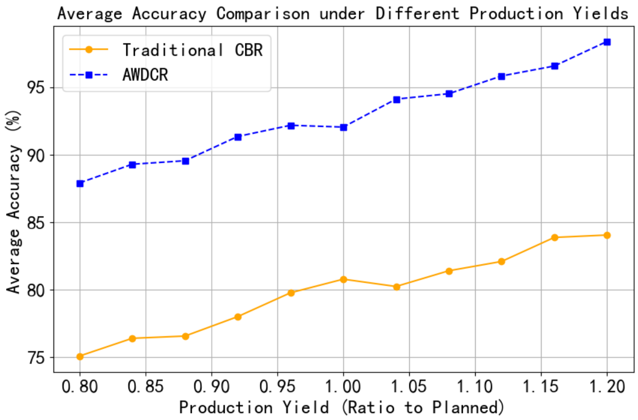

Fig. 4. Comparison of average accuracy of algorithms under different production outputs.

5 Conclusion

This paper successfully built a coal mine intelligent decision support system based on CBR and implemented the AWDCR algorithm. Experimental simulation and practical application have effectively verified the significant advantages of the algorithm in retrieval efficiency and decision accuracy. By dynamically adjusting the case attribute weights, compared with traditional algorithms, the retrieval time is shortened by 20%, and the decision accuracy is increased from 80% to 90%, which significantly enhances the production control capability of coal mines, reduces the occurrence rate of faults, and improves production efficiency. However, in the face of highly complex working conditions, such as rare geological conditions, the problem of insufficient completeness of the case library is prominent. Future work will focus on expanding the scope of the case library and actively collecting special working condition data for supplementation; at the same time, integrating deep learning technology. Specifically, we will integrate LSTM for temporal anomaly prediction (e.g., temperature trends) with CBR, using a hybrid model where LSTM outputs adjust case weights in real-time during retrieval, to better meet the high requirements of coal mine intelligent development and continue to provide reliable decision support for coal mine safety production.

References

[1] H. Eroğlu, O. Kurtuluş, and O. Can. Decision making methods used for improving power quality to meet the sustainable energy targets. Energy & Environment, vol. 36(1), 499–539, 2025. https://doi.org/10.1177/0958305x241276829

[2] P. M. Spanidis, F. Pavloudakis, and C. Roumpos. The Knowledge Management Landscape in the Greek Coal Mining Industry. Mining, Metallurgy & Exploration, 41(2), 819–841, 2024. https://doi.org/10.1007/s42461-024-00958-8

[3] X. Zhang, J. Gao, K. Zhao, Y. Li, C. Li, and Z. Xie. Research on generation method of intelligent emergency preplan for coal mine accident. Mining Safety & Environmental Protection, 51(1), 78–85, 2024. http://doi.org/10.19835/j.issn.1008-4495.20230320

[4] X. Ding, M. Hasanipanah, M. Monjezi, R. A. Abdullah, T. Nguyen, and D. V. Ulrikh. Enhancing mine blasting safety: developing intelligent systems for accurate flyrock prediction through optimized group method of data handling methods. Natural Resources Research, 34(2), 1199–1218, 2025. https://doi.org/10.1007/s11053-024-10445-y

[5] L. Wei, M. Lei, Z. Xiancheng, T. Xianhui, and F. Quan. "Research progress of the rotary steerable system and its control methods." Coal Geology & Exploration, vol. 51(10), pp. 167–179, 2023. http://doi.org/10.12363/issn.1001-1986.23.03.0137.

[6] T. Ye, M. Dong, Y. Liang, J. Long, Y. Zheng, and J. Lu. Online optimization of boiler operation based on information integration and case-based reasoning. International Journal of Green Energy, 20(1), 15–27, 2023. https://doi.org/10.1080/15435075.2021.2021417