ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Design of big data traceability and risk prevention and control system for aviation safety incidents

Yu Cheng[1,*]

[1] Information Engineering College, Chongqing Industry and Information Vocational College, Changshou High-tech Zone, Chongqing 401233, P.R.China

SUMMARY: *Aiming at the problems of scattered data, difficulty in tracing risk sources and lagging prevention and control of aviation safety incidents, this paper designs a big data traceability and risk prevention and control system for aviation safety incidents. The system integrates flight operation parameters, safety incident reports, maintenance records, meteorological information, airport operation logs and air traffic control interaction data. Through data cleaning, time alignment, semantic coding and correlation graph modeling, it realizes event element extraction, risk chain tracking and level assessment. The experimental results show that the event recognition Accuracy of the proposed method reaches 95.1%, Macro-F1 reaches 94.6%, AUC reaches 97.1%, traceability hit rate reaches 92.6%, Top-3 risk source coverage reaches 96.4%, and risk level consistency rate reaches 93.1%. The end-to-end response time is 154 ms. The research shows that the system can provide data support for active early warning, cause review and hierarchical prevention and control of aviation safety events.*

KEYWORDS: *aviation safety incident; Big data provenance; Risk prevention and control; Association analysis; Risk level assessment*

1 Introduction

With the rapid development of the air transport industry, the aviation safety problem has gradually shifted from single equipment reliability control to comprehensive risk management under the joint action of multi-systems, multi-agents and multi-scenarios. Aviation safety incidents are not only related to the safety of passengers' lives and property, but also directly affect the efficiency of airport operation, the management cost of airlines and the stability of civil aviation supervision system. With the continuous improvement of information technology in flight operation, air traffic control, maintenance, airport support and aviation security, aviation safety management has accumulated a large number of structured and unstructured data, including flight parameters, fault records, maintenance logs, incident reports, weather data, flight operation data, control communication texts and safety inspection records. These data provide a new technical basis for security incident identification, cause tracking and risk warning.

In the traditional mode of aviation safety management, incident analysis mostly relies on manual investigation, expert experience and post-hoc attribution. This method can form a relatively complete explanation of typical accidents, but it is easily limited by information lag, sample dispersion and subjective judgment differences in the face of massive low-level events,

---

*chengyucqiivc@126.com

https://doi.org/10.65102/is2026914

Published: 30 April, 2026

near-loss events and hidden risk signals. Especially in the modern aviation operation environment, risk is often not directly triggered by a single factor, but is formed by the superposition of multiple factors such as personnel operation, equipment status, operating environment, organizational management and external disturbance. If the isolated event analysis method is still used, it is difficult to discover the potential correlation between different data sources in time, and it is difficult to reveal the evolution chain of security events from risk signs to events.

In recent years, machine learning, natural language processing, knowledge graph and big data association analysis techniques have been widely used in aviation safety research. Related studies have shown that building prediction models based on aviation incident reports, flight operation data and maintenance records can improve the accuracy of incident severity classification, anomaly detection and risk trend identification. At the same time, Bayesian network, structural topic model and text semantic analysis methods are also used to mine the key factors and causal relationships in aviation accident symptoms. These studies have promoted the transformation of aviation safety management from empirical judgment to data-driven, but there are still some shortcomings in the existing work. On the one hand, some studies focus more on a single task, such as accident classification, runway departure prediction or flight anomaly detection, and the integration of multi-source evidence chains before and after the event is insufficient. On the other hand, there are still some problems between different business systems, such as inconsistent data standards, missing fields, differences in semantic expression and difficulties in cross-department sharing, which make it difficult to form a continuous, interpretable and traceable analysis framework for the traceability process of aviation safety incidents.

Based on the above problems, this paper focuses on the needs of big data traceability and risk prevention and control of aviation safety events, and constructs an analysis system for multi-source aviation safety data. The system is based on flight operation data, aviation safety reports, maintenance support records, meteorological and airport operation information. After data cleaning, time alignment, semantic coding and anomaly feature extraction, a unified event data representation is formed. On this basis, this paper uses the big data association analysis and risk feature identification method to establish an aviation safety incident traceability model, excavate the propagation relationship, trigger conditions and risk chain between event elements, and further design the prevention and control strategy output mechanism for different risk levels. Different from the simple prediction model, this paper emphasizes the closed-loop process of "event identifie-cause trace-level assessment-policy feedback", so that the system can not only judge the state of the risk, but also explain the source of the risk, and provide an executable basis for the subsequent security management.

The core question this paper intends to answer is how to use multi-source aviation safety data to construct an interpretable incident traceability model, and on this basis, form a dynamic prevention and control system oriented to risk level. Around this problem, this paper studies from four levels of data processing, correlation modeling, risk assessment and system design, aiming to improve the real-time performance, accuracy and prevention and control initiative of aviation safety incident analysis. The research results can provide data safety governance tools for airlines, airport operation departments and civil aviation regulators, and also provide reference for the transformation of aviation safety management from post-disposal to pre-warning and process control.

2 Related Research

2.1 Research on identification and analysis technology of aviation safety incidents

Aviation safety event recognition technology has experienced the process of transformation from manual experience judgment to intelligent analysis. Early aviation safety management mainly relied on accident investigation reports, pilot feedback, maintenance records and expert review to manually classify incident types, causes and consequences. These methods can retain strong professional interpretation, but the processing cycle is long, and it is difficult to capture near-loss events, hidden errors and cross-system risk signals in time. With the continuous expansion of aviation operation data, researchers have begun to introduce machine learning, natural language processing and statistical inference methods into the aviation safety event recognition process to improve the ability of event classification, anomaly detection and risk pattern discovery. Demir et al. [1] pointed out through a systematic review that artificial intelligence has been gradually applied to aviation safety evaluation, accident identification and operational risk analysis, and promoted aviation safety research from experience induction to data modeling.

In terms of incident recognition methods, Bayesian inference, support vector machine, random forest and deep learning models are widely used for aviation accident and symptom classification. Cankaya et al. [2] introduced Bayesian inference into aviation event management decision-making, and showed that probabilistic inference could assist in identifying event occurrence mechanisms under uncertain conditions. Caetano[3] combined accident investigation information and meteorological data to carry out aviation accident and event prediction, and proved that multi-source variables can enhance the situational adaptability of event recognition. Omrani et al. [4] evaluated the aviation accident severity prediction dataset and pointed out that data quality, category distribution and feature selection can significantly affect the recognition effect of machine learning models. Haselein et al. [5] built a variety of machine learning models around close air encounters, indicating that probabilistic modeling is helpful to identify low-frequency but high-risk aviation safety scenarios. The existing research has proved the effectiveness of intelligent recognition technology, but most of the work is still focused on the classification of event results, and the description of the multi-step evolution process before the event occurs is insufficient.

2.2 Research on traceability of aviation safety incidents driven by big data

The traceability of aviation safety events emphasizes the restoration of the process of risk generation, transmission and amplification from multi-source data. The research object is no longer a single event label, but the evidence chain, causal chain and risk propagation chain behind the event. Modern aviation operation system involves flight operation, maintenance, air traffic management, airport security, meteorological environment and information system. The data generated by different links have the characteristics of various formats, inconsistent time granularity and different semantic expression. Therefore, big data-driven traceback research needs to complete multi-layer processing such as data aggregation, time alignment, semantic parsing, correlation modeling and causal inference, so as to transform scattered event records into traceable risk clues. Bauranov and Rakas[6] constructed a Bayesian network model for aviation safety to analyze the impact of new communication technologies on the risk of air collisions, indicating that the network probability structure can express the dependence between event elements.

In specific applications, flight operation data and text reports are an important basis for aviation safety traceability. Odisho et al. [7] used machine learning methods to improve runway safety management capabilities, and linked runway deviation risk mitigation with operational data characteristics. Silva and Murca [8] proposed a data analysis framework for flight operation anomaly detection, and showed that continuous flight parameters can reveal early signals that deviate from the normal operation state. KAIDI et al. [9] improved aircraft accident severity classification through feature selection, feature extraction and machine learning models, reflecting the basic value of feature engineering in accident traceability. Ghaderi and Saghafi[10] combined flight data and continuous performance test to identify pilot attention loss, indicating that personnel status data can be used as an important supplement to trace the source of safety incidents. In general, the research on big data provenance has expanded from the analysis of single operating parameters to the comprehensive modeling of personnel, equipment, environment and management factors. However, there is still room for further improvement in cross-system data sharing, evidence chain interpretation and causality verification.

2.3 Research on aviation safety risk prediction and prevention and control strategy

The goal of aviation safety risk prediction and prevention and control research is to identify the risk trend before the event causes serious consequences, and convert the prediction results into executable intervention strategies. Different from traditional post-rectification, risk prevention and control emphasizes dynamic monitoring, hierarchical early warning and closed-loop feedback. Mendes et al. [11] systematically reviewed aviation maintenance risk management and pointed out that organizational processes, personnel capabilities and equipment status in the maintenance process would all affect aviation safety risks. Patriarca et al. [12] discussed the application of business intelligence and machine learning in air traffic management safety, indicating that data democratization and intelligent decision-making tools can improve the efficiency of risk governance. Rose et al. [13] applied the structural topic model to aviation safety data analysis and proved that text topic mining could extract potential risk patterns from a large number of incident reports. Jonk et al. [14] studied the natural language processing method of aviation incident reports and showed that text semantic analysis can assist safety management departments to find high-frequency risk factors and abnormal expressions.

With the advancement of the digitalization of the aviation system, network security, onboard intelligence and pilot cognitive state have gradually entered the research scope of risk prevention and control. Ukwandu et al. [15] reviewed the cybersecurity challenges in the aviation industry, and pointed out that aviation digital infrastructure faces new risks such as data attacks, system intrusion and communication interference. Habler et al. [16] proposed an aircraft safety assessment method system, emphasizing that safety prevention and control should cover the airborne system, communication link and operation environment. Tim et al. [17] carried out research on the future air traffic management network security risk assessment, indicating that the intelligent air traffic management system needs a more fine-grained risk identification mechanism. Faye et al. [18] used case study method to analyze aviation network security preparedness capability, reflecting the role of emergency response mechanism in risk prevention and control system. Luettig et al. [19] discussed the certification challenges of airborne artificial intelligence, suggesting that the aviation safety prediction model must take into account accuracy, interpretability and verifiability in the actual deployment. Masse et al. [20] used explainable artificial intelligence to identify alert-related electrophysiological features in flight simulators, indicating that human-computer interaction data can improve the accuracy of risk perception. Somon et al. [21] conducted experiments on the classification of attention deficit in flight simulation, which provided physiological signal basis for pilot cognitive risk identification. Ziakkas and Pechlivanis[22] explored the application of artificial intelligence in aviation accident classification, further illustrating the application potential of intelligent models in risk identification and safety governance. In order to further sort out the differences of existing studies in data sources, task objectives, technical methods and research limitations, this paper summarizes the representative results related to aviation safety incident identification, traceability and risk prevention and control, and the results are shown in Table 1.

Table 1: Summary of related research on aviation safety incident identification, traceability and risk prevention and control

Research Direction	Representative Literature	Main Method	Research Contribution	Remaining Problem
Intelligent aviation safety identification	Demir et al. [1]	Review of artificial intelligence systems	Summarized application trends of AI in aviation safety identification and evaluation	Insufficient modeling of specific traceability chains
Probabilistic inference of aviation events	Cankaya et al. [2]	Bayesian inference	Supported event management decision-making under uncertain conditions	Limited consideration of multi-source real-time data fusion
Joint prediction of accidents and weather	Caetano [3]	Machine learning prediction	Combined accident investigation information with meteorological variables	Insufficient explanation of event evolution paths
Accident severity classification	Omrani et al. [4]	Dataset evaluation and machine learning	Analyzed the impact of data quality on severity prediction	Generalization ability is limited by data distribution
Near mid-air collision analysis	Haselein et al. [5]	Multi-model probabilistic reasoning	Improved the identification capability for low-frequency high-risk events	Insufficient connection with subsequent prevention and control strategies
Runway safety risk mitigation	Odisho et al. [7]	Machine learning and operational data analysis	Supported runway excursion risk identification	Focuses on a single scenario, with limited system scalability
Flight operation anomaly detection	e Silva and Murça [8]	Data analysis framework	Detected abnormal patterns in flight parameters	Insufficient integration with textual reports and maintenance data

In summary, the existing research has formed a rich technical foundation in aviation safety incident recognition, flight anomaly detection, text report mining and risk prediction. However, most studies still focus on a single data source, a single model task or a single risk scenario, and pay insufficient attention to the cross-domain traceability, hierarchical assessment and closed-loop design of prevention and control of aviation safety incidents. Based on this, this paper further builds a big data traceability and risk prevention and control system for the requirements of cross-source evidence integration and hierarchical disposal of aviation safety incidents, so as to improve the active identification ability and process interpretation ability of safety governance.

3 Research Methods

The big data traceability and risk prevention and control system of aviation safety incidents constructed in this paper aims to identify incident symptoms, track risk sources, and form a hierarchical prevention and control basis from multi-source heterogeneous aviation operation data. Flight operation parameters, aviation safety incident reports, maintenance support records, meteorological observation information, airport operation logs and air traffic management data are integrated into a unified analysis framework. After data cleaning, time alignment, semantic coding, feature fusion and risk labeling, a standardized data set that can be called by the traceability model and the risk assessment model is formed. Different from the single event classification method, the proposed method emphasizes the continuous processing from "original data access" to "event chain recovery", so that the security information scattered in different business systems can be uniformly organized and calculated.

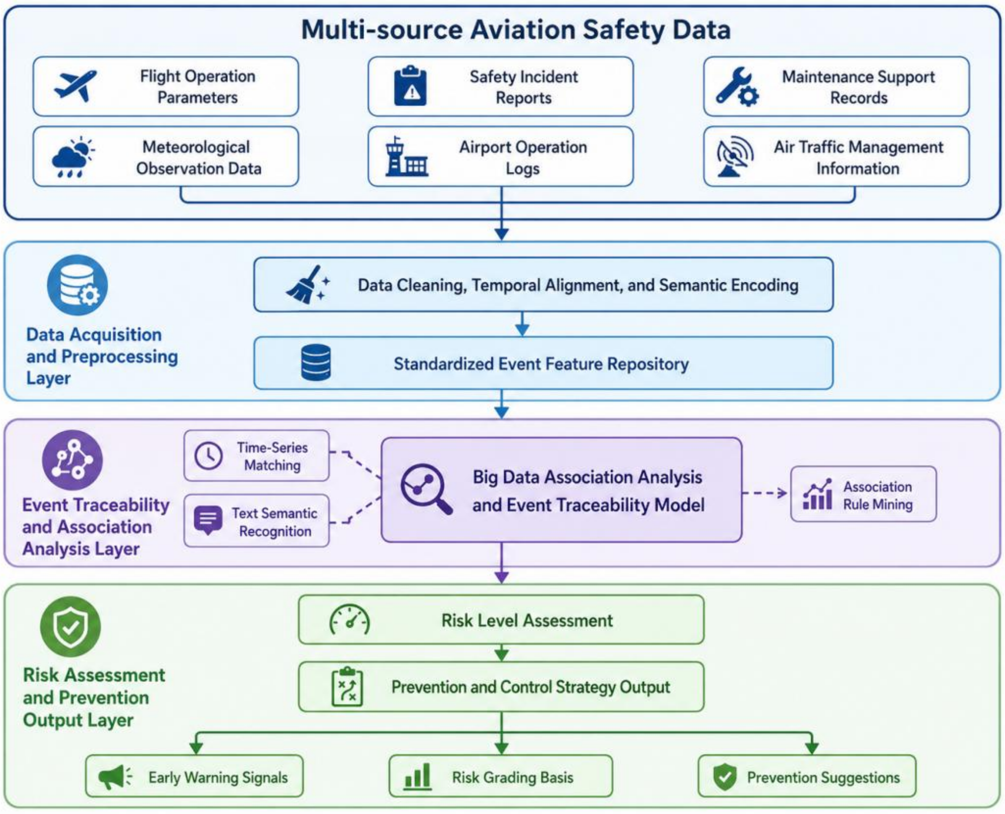

Figure 1: Overall framework of big data traceability and risk prevention and control of aviation safety incidents

As shown in Figure 1, the research method in this paper consists of three layers: data collection and preprocessing layer, event traceability and correlation analysis layer, and risk assessment and prevention and control output layer. The data acquisition and preprocessing layer is responsible for multi-source aviation safety data access, field normalization and quality control. The event provenance and association analysis layer extracted event triggers through time series matching, text semantic recognition and association rule mining. The risk assessment and prevention and control output layer generates early warning signals and prevention and control suggestions according to the risk level, event type and historical disposal results. The architecture can transform post-analysis of aviation safety management into process-oriented risk identification and closed-loop control.

3.1 Multi-source aviation safety data acquisition and preprocessing

The formation of aviation safety incidents usually involves many factors such as personnel operation, equipment status, operating environment and organizational management, so it is difficult for a single data source to completely explain the event process. The data collected in this paper mainly includes six categories: flight operation data is used to record altitude, speed, heading, vertical overload, engine status and automatic driving mode changes. Aviation safety incident report is used to describe the incident process, personnel disposal and preliminary reasons; Maintenance support records are used to reflect component failures, delayed maintenance and repeated defects; Meteorological data are used to characterize environmental conditions such as wind shear, low visibility, thunderstorms, and crosswinds. The airport operation log is used to record runway status, apron support and taxiing conflicts. Atc data is used to reflect interval changes, instruction changes, and communication anomalies. Different data have obvious differences in sampling frequency, field format and semantic expression, which need to be uniformly processed before modeling.Let the original multi-source data of a flight or event sample be as follows.

$$\mathrm{X_i} = \{\mathrm{F_i}, \mathrm{R_i}, \mathrm{M_i}, \mathrm{W_i}, \mathrm{A_i}, \mathrm{T_i}\} \tag{1}$$

where, $\mathrm{F_i}$ represents flight operation parameters, $\mathrm{R_i}$ represents safety incident report, $\mathrm{M_i}$ represents maintenance support record, $\mathrm{W_i}$ represents meteorological observation data, $\mathrm{A_i}$ represents airport operation log, and $\mathrm{T_i}$ represents air traffic control operation information. Through this representation, aviation safety incidents are no longer regarded as isolated text records, but are organized as comprehensive samples containing operational status, environmental conditions, and management behaviors.

In the data cleaning phase, this paper removes duplicate records, invalid fields, and timestamp abnormal samples. For the short-time missing values of flight operation parameters, the interpolation method of adjacent time Windows was used to complete the missing values. For missing fields in maintenance records and incident reports, the null value tags are retained according to business rules to avoid semantic misjudgment caused by forced filling. For continuous variables, this paper uses standardization to eliminate dimensional differences, which is calculated as follows.

$$\mathrm{z_{ij}} = \frac{\mathrm{x_{ij}} - \mu_\mathrm{j}}{\sigma_\mathrm{j} + \varepsilon} \tag{2}$$

where, $\mathrm{x_{ij}}$ represents the original value of the ith sample on the JTH continuous feature, μ_j and σ_j represent the mean and standard deviation of the feature, ε is the minimum constant to prevent the denominator from being zero, and $\mathrm{z_{ij}}$ is the normalized feature value. This processing can reduce the bias of altitude, speed, temperature, wind speed and other variables caused by different dimensions on model training.

In the time alignment stage, this paper establishes a unified time window centered on the event occurrence time, and maps the key data before, during and after the event to the same time axis. For high frequency data such as flight parameters, fixed time granularity is used for resampling. For low-frequency data such as maintenance records, incident reports and air traffic control texts, the association matching is performed according to the event number, flight number, aircraft registration number, airport code and time range. Text data preprocessing includes removing irrelevant symbols, unifying professional abbreviations, word segmentation, stop word filtering and semantic vector coding, so that safety semantics such as "runway intrusion", "communication misunderstanding", "go-around" and "unstable approach" can be transformed into feature representations that can be recognized by the model. In order to ensure the reliability of the subsequent traceability model, this paper sets up quality control indicators for different source data, including completeness rate, consistency rate, time matching rate and abnormal proportion. The collection contents and preprocessing methods of various types of data are shown in Table 2.

Table 2: Multi-source aviation safety data collection content and preprocessing methods

Data Type	Main Fields	Data Characteristics	Preprocessing Method	Output Result
Flight operation data	Altitude, speed, heading, vertical rate, engine status	High-frequency continuous data	Missing value interpolation, resampling, standardization	Flight state features
Safety event reports	Event type, event process, handling process, preliminary cause	Unstructured text	Text cleaning, word segmentation, semantic encoding	Event semantic features
Maintenance support records	Faulty component, maintenance action, repeated defect, release status	Semi-structured data	Field mapping, categorical encoding, missing value marking	Equipment risk features
Meteorological observation data	Wind speed, visibility, precipitation, thunderstorm, cloud base height	Spatiotemporal environmental data	Time matching, outlier filtering, unit unification	Environmental disturbance features
Airport operation logs	Runway status, taxiing conflict, apron support, flight delay	Event-based data	Flight matching, time-window association	Surface operation features
Air traffic control operation information	Instruction change, separation variation, communication abnormality, route adjustment	Mixed sequence and text data	Semantic extraction, instruction labeling, time alignment	Air traffic control interaction features

3.2 Aviation safety incident traceability model based on big data association analysis

In aviation safety incident analysis, incident results can usually be described by report records, operation logs and investigation conclusions, but the risk formation process is often hidden among multi-source data. A runway incursion, air gap reduction, unplanned return flight, or maintenance related event may be simultaneously affected by changes in flight status, crew operations, ATC instructions, equipment health, weather disturbances, and airport operating conditions. If the attribution is only based on a single event report, it is easy to compress the complex risk into a single direct cause, and ignore the transmission relationship of risk in different business links. Therefore, on the basis of the standardized event feature library formed in Section 3.1, this paper constructs an aviation safety event traceability model based on big data association analysis, and realizes the transformation from result description to process tracking of aviation safety events through event element extraction, association graph construction, path search and risk source sorting. The overall flow of the model is shown in Figure 2.

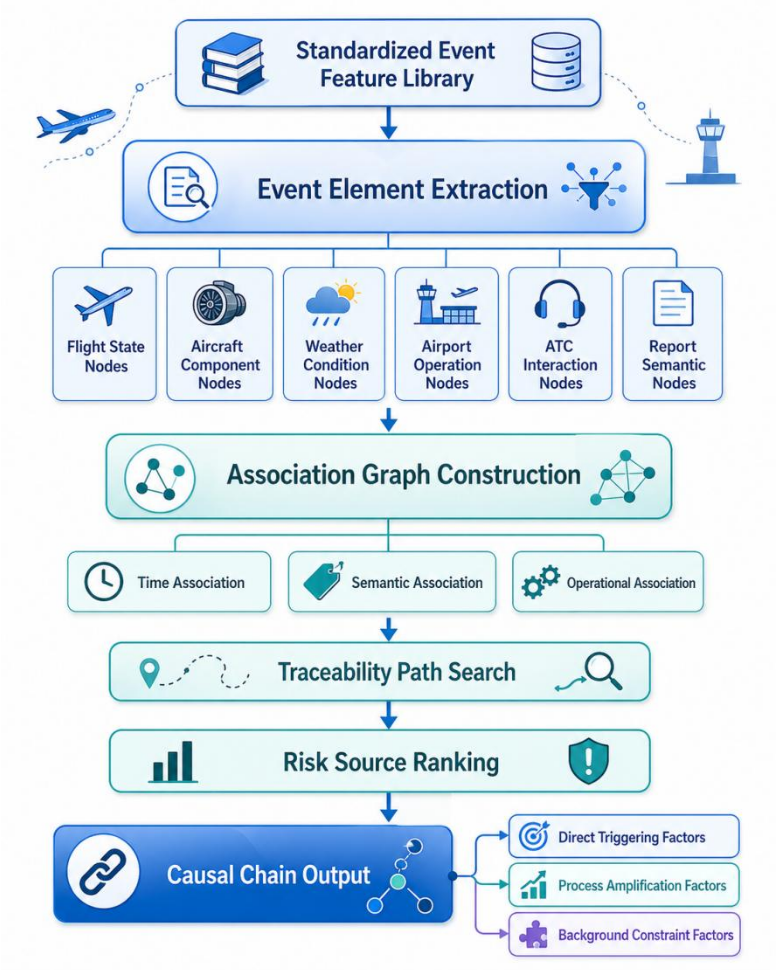

Figure 2: Big data association traceability model of aviation safety incidents

It can be seen from Figure 2 that the aviation safety incident traceability model does not directly make a static discrimination of the event category, but establishes a correlation relationship around the data nodes before and after the event. Firstly, the model extracts elements such as flights, aircraft, components, airports, runways, weather conditions, air traffic control instructions, maintenance actions and event texts from multi-source data, and then converts these elements into event nodes. Let the aviation safety incident correlation graph be as follows.

$$G = (V, E, \Omega) \tag{3}$$

where, V represents the set of event element nodes, E represents the associated edges formed between nodes, and Ω represents the set of edge weights. The graph structure can preserve the connection relationship between different security elements, so that the model can not only identify "which category" the event belongs to, but also further analyze "which risk factors gradually promote the formation of the event".

In the process of constructing correlation edges, this paper comprehensively considers temporal correlation, semantic correlation and business correlation. Time correlation is used to measure whether two event elements are in the same risk evolution window, for example, whether there is a time proximity relationship between meteorological abrupt changes and round-trip events. Semantic association is used to measure whether the expressions in incident reports, maintenance records and ATC texts have similar meanings. Business association is used to determine whether there is a logical connection between two elements in the aviation operation process, such as the relationship between component failure and maintenance release, runway status and approach disposal. The comprehensive association weight of node i and node j is defined as follows.

$$\omega_{ij} = \alpha e^{-\frac{|t_i - t_j|}{\tau}} + \beta \cos(s_i, s_j) + \gamma b_{ij} \tag{4}$$

where, ω_{ij} represents the comprehensive association strength between two nodes, t_i and t_j represent the corresponding time markers of nodes, τ represents the time attenuation coefficient, s_i and s_j represent the text semantic vector, b_{ij} represents the business rule association value, and α, β, γ are the weight coefficients. Through this formula, the model can simultaneously absorb three types of information: time order, text meaning and operation logic, and reduce misjudgment caused by simply relying on keyword matching or time proximity.

After completing the construction of the association graph, the model takes the confirmed aviation safety event node as the target node and searches for candidate risk sources from its lead time window. For the target event q, the traceability contribution of the candidate risk source node k is defined as follows.

$$\rho_k = \sum_{p \in P(k,q)} \left(\prod_{(i,j) \in p} \omega_{ij} \right) a_k \eta_p \tag{5}$$

where, ρ_k represents the traceability contribution of node k to the target event q, P(k,q) represents the set of candidate paths from node k to node q, a_k represents the abnormal strength of node itself, and η_p represents the path continuity correction factor. When a node appears repeatedly in multiple high-weight paths and its abnormal intensity is high, the node will be identified as a key risk source.

In order to improve the interpretability of traceability results, this paper divides the output results into direct trigger factors, process amplification factors and background constraint factors. The direct trigger factors mainly include abnormal operation, equipment failure, ATC instruction conflict and sudden weather change. Process amplification factors are mainly reflected in the risk diffusion in flight operation, maintenance support, airport surface and air traffic control coordination links. Background constraints include flight density, support resources, organizational management and historical repeated defects. After the above processing, the output of the model is no longer a single cause label, but a risk chain with weight ranking. The results can provide a computable basis for subsequent risk level assessment, prevention and control strategy matching and security management closed-loop.

3.3 Feature identification and level assessment of aviation safety events for risk prevention and control

After completing the association and traceability of aviation safety events in Section 3.2, the system has been able to identify potential risk sources and event evolution chains from multi-source data. However, the traceability results still need to be further transformed into risk levels that can be used for prevention and control decision-making, otherwise it is difficult to support the airport operation department, airline safety management department and air traffic control unit to carry out differentiated disposal. The risk of aviation safety incident does not completely depend on whether the incident has occurred, but also is related to the severity of the incident, the probability of occurrence, the scope of transmission, the difficulty of disposal and the historical repeatability. For example, short-time poor communication, runway incursion signs, unsteady approach and repeated failure of critical components can be classified as safety incidents on the surface, but their risk levels and prevention and control priorities are significantly different. Therefore, based on the risk chain output by the traceability model, this paper constructs an aviation safety incident feature recognition and level assessment method for risk prevention and control.

The aviation safety event feature recognition module takes the standardized event feature library and traceability contribution results as input, and focuses on extracting four types of risk features: operation deviation features, equipment status features, environmental disturbance features and management response features. The operational deviation feature is used to describe the abnormal degree of flight altitude, speed, heading, vertical rate, interval change and approach stability. Equipment state features are used to describe faulty components, repeated defects, maintenance delays and release records. The environmental disturbance characteristics reflect external effects such as crosswind, low visibility, thunderstorms, runway smoothness and traffic flow density. The management response characteristics focus on incident discovery time, disposal time, emergency coordination and rectification closed-loop situation. In order to avoid a single indicator dominating the risk judgment, this paper combines various features into a risk assessment vector:

$$\mathrm{r_i} = [\mathrm{u_i}, \mathrm{d_i}, \mathrm{e_i}, \mathrm{m_i}, \rho_\mathrm{i}] \tag{6}$$

where, $\mathrm{r_i}$ represents the risk assessment vector of the ith aviation safety incident, $\mathrm{u_i}$ represents the operation deviation feature, $\mathrm{d_i}$ represents the equipment status feature, $\mathrm{e_i}$ represents the environmental disturbance feature, $\mathrm{m_i}$ represents the management response feature, and ρ_i represents the aggregated value of traceability contribution obtained in Section 3.2. The performance of events and the intensity of risk sources are included in the assessment process at the same time, so that the risk level judgment no longer stays at the result description level.

In the stage of risk level assessment, this paper uses the combination of weighted comprehensive scoring and classification mapping. Let the weights of various risk features be $\theta = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5]$, then the comprehensive risk score of an event is defined as follows.

$$\mathrm{R_i} = \theta_1\mathrm{u_i} + \theta_2\mathrm{d_i} + \theta_3\mathrm{e_i} + \theta_4\mathrm{m_i} + \theta_5\rho_\mathrm{i} \tag{7}$$

In the formula, R_i represents the comprehensive risk score of the event, and the weight parameters can be obtained by training based on historical event samples, or corrected by combining with the evaluation of aviation safety experts. The design can take into account both data-driven results and industry safety experience, and avoid underestimation of risk due to insufficient samples in low-frequency high-risk events.

In order to transform continuous risk scores into executable prevention and control levels, this paper further constructs a level mapping function:

$$L_i = \begin{cases} I, R_i < \delta_1 \\ II, \delta_1 \le R_i < \delta_2 \\ III, \delta_2 \le R_i < \delta_3 \\ IV, R_i \ge \delta_3 \end{cases} \tag{8}$$

where, L_i represents the risk level of the ith event, and $\delta_1, \delta_2, \delta_3$ are the level thresholds. Level I corresponds to events of general concern, level II corresponds to low and medium risk events that need to be tracked and handled, level III corresponds to high risk events that need cross-departmental collaborative intervention, and level IV corresponds to high risk events that may cause serious consequences. Through the level mapping, the model output can directly serve the subsequent prevention and control strategy matching. The identification basis and prevention and control implications corresponding to different risk levels are shown in Table 3. Table 3 links the model calculation results with aviation safety management actions, so that the risk assessment results have a clear business interpretation.

Table 3: Risk level assessment criteria and prevention and control implications of aviation safety incidents

Risk Level	Main Identified Features	Risk Status Description	Prevention and Control Focus
Level I	Slight deviation in a single indicator, without continuous abnormalities	General operational fluctuation with limited short-term impact	Include in routine monitoring and record event features
Level II	Mild abnormalities in multiple indicators, with a local risk chain	May develop into a recurring safety hazard	Conduct cause investigation and track changes in similar events
Level III	Abnormal key operational indicators, with a relatively complete traceability chain	Potential cross-stage risk transmission	Initiate interdepartmental coordination and develop targeted rectification measures
Level IV	Concentrated high-risk factors, with significant abnormal intensity and propagation path	May lead to serious aviation safety consequences	Trigger high-level warning, implement immediate intervention and review

3.4 Design of aviation safety risk prevention and control system

After the aviation safety event feature recognition and risk level assessment are completed, the system needs to convert the model output into executable prevention and control actions. Aviation safety risk prevention and control is not a single warning release, but a closed-loop process of continuous update. In the initial stage, an event may be only a low-level operational deviation, but if it is accompanied by repeated failures, bad weather, runway resource constraints, or frequent changes in ATC instructions, the risk status will increase for a short time. Therefore, based on the risk level assessment results in Section 3.3, this paper designs an aviation safety risk prevention and control system, so that the system can dynamically adjust the prevention and control strategy according to real-time data changes, traceability contribution and historical disposal effects.

The risk prevention and control system constructed in this paper is composed of four modules: risk status update, strategy matching, disposal and execution, and feedback correction. The risk status update module receives the comprehensive risk score of the event, the risk level and the traceability chain results to determine whether the current risk is expanding. The strategy matching module selects the corresponding prevention and control actions according to the risk level, including continuous monitoring, special verification, cross-departmental collaborative disposal and high-level early warning. The disposal and execution module pushes the strategy results to the airline safety management department, the airport operation control department, the maintenance department and the air traffic control coordination unit. The feedback correction module updates the model parameters according to the time effectiveness of disposal, the degree of risk reduction and the recurrence of similar events. The overall structure is shown in Figure 3.

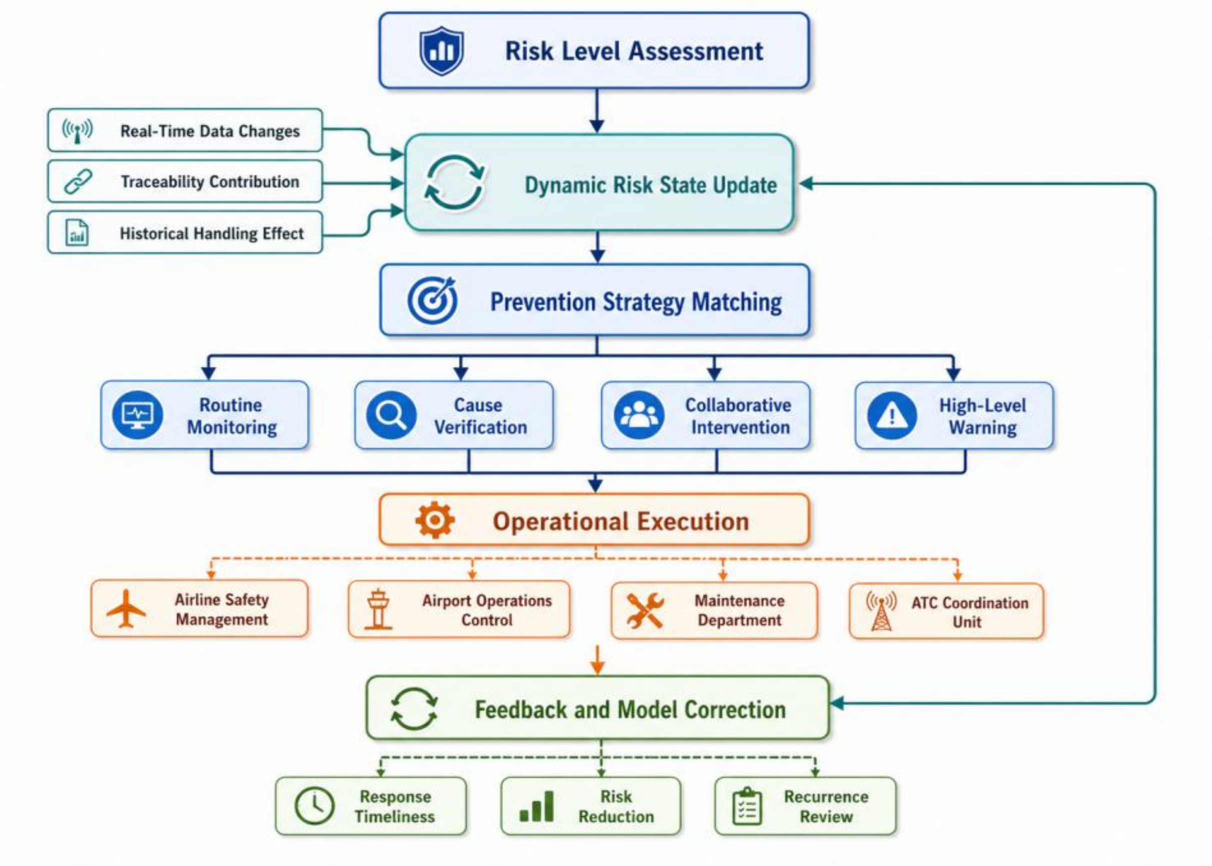

Figure 3: Operation process of aviation safety risk prevention and control system

It can be seen from Figure 3 that the aviation safety risk prevention and control system has an obvious circular structure. The system does not terminate after the risk level output, but continues to receive newly generated operational data and disposal feedback to reassess the risk status. In order to depict the dynamic change of aviation safety risk, this paper defines the dynamic risk state of event i at time t as $H_i(t)$, and its update process is expressed as follows.

$$H_i(t+1) = \lambda H_i(t) + (1-\lambda)R_i(t) + \kappa C_i(t) - \phi Q_i(t) \tag{9}$$

where $H_i(t+1)$ represents the risk state at the next moment, $R_i(t)$ represents the current comprehensive risk score, $C_i(t)$ represents the aggregation strength of high contribution nodes in the traceability chain, $Q_i(t)$ represents the quality of disposal feedback, λ is the historical risk memory coefficient, κ is the chain amplification coefficient, ϕ is the disposal inhibition coefficient. The formula reflects the continuity and intervention of the risk state. If the high contribution risk node continues to appear, the risk state will be amplified. If the disposal feedback is timely and effective, the risk status will gradually decrease.

In the strategy matching phase, the system needs to select the scheme with higher benefits and controllable costs from multiple optional prevention and control actions. Let the set of control actions be $A = \{a_1, a_2, \ldots, a_n\}$, These include data review, enhanced flight quality monitoring, key flight tracking, maintenance review, runway operation restrictions, air traffic control collaborative reminders and special safety review. The comprehensive utility of a control action A is defined as follows.

$$U(a) = \xi_1 D(a) + \xi_2 P(a) - \xi_3 K(a) - \xi_4 G(a) \tag{10}$$

where, D(a) represents the expected contribution of the action to the risk reduction, P(a) represents the matching degree between the action and the current risk type, K(a) represents the execution cost, G(a) represents the impact on the flight operation efficiency, and $\xi_1, \xi_2, \xi_3, \xi_4$ are the regulation coefficients. The design makes the prevention and control strategy not only be triggered mechanically according to the risk level, but also consider the prevention and control effectiveness, business adaptability and operation cost.

In actual operation, the system selects the optimal prevention and control action according to the dynamic risk state and the strategy utility, which is expressed as follows.

$$a_i^* = \arg\max_{a \in A} [U(a) \mid L_i, H_i(t), \rho_i] \tag{11}$$

where, a_i^* represents the optimal prevention and control action corresponding to event i, L_i represents the risk level, $H_i(t)$ represents the dynamic risk state, and ρ_i represents the traceability contribution degree. For level I events, the system mainly generates daily monitoring and trend record suggestions. For level II events, the system trigger cause verification and similar event tracking were performed. For level III events, the system pushed a cross-department collaborative disposal task. For level IV events, the system initiates high-level warning and requires immediate intervention, risk review and rectification closure.

4 Experimental Evaluation

4.1 Experimental Design

In order to verify the effectiveness of the aviation safety event big data traceability and risk prevention and control system constructed in this paper, the experiment is carried out around three tasks: event identification, traceability chain reduction and risk level assessment. The event recognition experiment was used to test whether the model could accurately distinguish the types of safety events from flight operation data, incident reports and maintenance records. The traceability experiment was used to analyze the ability of the model to identify the key risk sources, antecedent symptoms and propagation paths. The risk level evaluation experiment investigates whether the system can transform the event characteristics and traceability results into executable risk levels, which provides a basis for subsequent prevention and control strategy matching.

The experimental data consisted of public aviation safety reports, simulated data in the operation record field and desensitization samples. The time range was set from 2019 to 2024, and a total of 36,820 aviation safety-related records were collected. Among them, there are 18,600 flight operation parameter samples, 7,240 aviation safety incident reports, 5,860 maintenance support records, 3,420 meteorological and airport operation data, and 1,700 air traffic control interaction records. The event types cover scenarios such as unstable approach, runway operation conflict, communication anomaly, equipment failure, maintenance release deviation, and meteorological disturbance effects. All samples are divided into training set, validation set and test set according to the ratio of 7:2:1, and the distribution of different event types in the three subsets is basically the same.

The experimental environment uses Ubuntu 22.04 operating system, Python 3.11 as the development language, and the deep learning framework is PyTorch 2.2. The hardware platform is configured with an Intel Core i7-12700 processor, 32 GB memory, and an NVIDIA RTX 4080 16 GB GPU. In order to simulate the online deployment requirements of the actual aviation safety management system, the model inference time and the risk strategy generation time are recorded simultaneously. Logistic regression, random forest, support vector machine, BiLSTM and common graph association model were selected as the comparison models, and manual rule discrimination method was used as the traditional baseline. The evaluation metrics include Accuracy, Macro-F1, Recall, AUC, traceability hit rate, Top-3 risk source coverage, and average response time. See Table 4 for the composition of experimental data and task Settings.

Table 4: Composition of experimental data and task Settings

Experimental Task	Sample Size	Input Data	Comparative Methods	Evaluation Metrics	Objective Description
Aviation safety event identification	36,820 records	Flight parameters, event reports, maintenance records	Logistic Regression, Random Forest, SVM, BiLSTM	Accuracy, Macro-F1, Recall, AUC	Determine event types and abnormal states
Event traceability chain reconstruction	12,460 records	Standardized event features, time windows, textual semantic vectors	Manual rules, ordinary graph association model	Traceability hit rate, Top-3 risk source coverage	Identify key risk sources and propagation paths
Risk level assessment	36,820 records	Risk feature vectors, traceability contribution, historical handling records	Random Forest, SVM, BiLSTM	Accuracy, Macro-F1, level consistency rate	Output Level I–IV risk levels
Prevention and control response test	4,800 records	Risk levels, event types, strategy library records	Manual rule-based strategy	Average response time, strategy matching rate	Test strategy generation efficiency and adaptability

4.2 Analysis of experimental results

In order to fully verify the effectiveness of the aviation safety incident big data traceability and risk prevention and control system in this paper, the experiment was carried out from four aspects: incident recognition performance, traceability analysis ability, risk level assessment effect and system response efficiency. The comparison methods include BiLSTM, random forest, support vector machine, logistic regression, and artificial rule methods. The experimental evaluation indicators include Accuracy, Precision, Recall, Macro-F1, AUC, traceability hit rate, Top-3 risk source coverage rate, level consistency rate and average response time. The comprehensive results of each model on the test set are shown in Table 5.

Table 5: Comparison of comprehensive performance of different methods in aviation safety incident analysis task

Method	Accuracy (%)	Precision (%)	Recall (%)	Macro-F1 (%)	AUC (%)	Traceability Hit Rate (%)	Top-3 Risk Source Coverage (%)	Risk Level Consistency Rate (%)
Proposed Method	95.1	94.8	94.5	94.6	97.1	92.6	96.4	93.1
BiLSTM	91.3	90.9	90.4	90.8	93.8	84.7	89.6	88.5
Random Forest	88.4	87.6	88.1	87.9	91.0	80.3	85.2	84.7
Support Vector Machine	85.1	84.2	85.0	84.6	88.4	76.8	81.5	80.9
Logistic Regression	82.4	81.6	81.9	81.7	85.2	70.5	76.2	77.4
Manual Rule-based Method	75.8	73.6	74.9	74.2	78.1	62.4	69.1	68.5

It can be seen from Table 5 that the proposed method achieves 95.1% Accuracy, 94.6% Macro-F1 and 97.1% AUC in the aviation safety incident recognition task, and the overall performance is better than that of various baseline models. Compared with BiLSTM, the Accuracy of the proposed method is improved by 3.8 percentage points, Macro-F1 is improved by 3.8 percentage points, and AUC is improved by 3.3 percentage points. Compared with random forest, the Accuracy is improved by 6.7 percentage points, Macro-F1 is improved by 6.7 percentage points, and AUC is improved by 6.1 percentage points. The Macro-F1 of the manual rule method is only 74.2%, and the AUC is 78.1%, indicating that the fixed threshold and empirical rules are difficult to deal with the complex scene of multi-source heterogeneous data in aviation safety incidents.

Figure 4 further demonstrates the differences between the models in the two core metrics Macro-F1 and AUC. In Figure 4, the Macro-F1 of the proposed method is 94.6%, and the AUC is 97.1%. BiLSTM were 90.8% and 93.8%, respectively; Random forest is 87.9% and 91.0%; Support vector machine 84.6% and 88.4%; Logistic regression were 81.7% and 85.2%, respectively; The artificial rule method is 74.2% and 78.1%, respectively. From the numerical distribution, it can be seen that the proposed method not only has the highest comprehensive classification performance, but also maintains a high consistency between Macro-F1 and AUC, indicating that the model still has good discrimination stability under the condition of class imbalance.

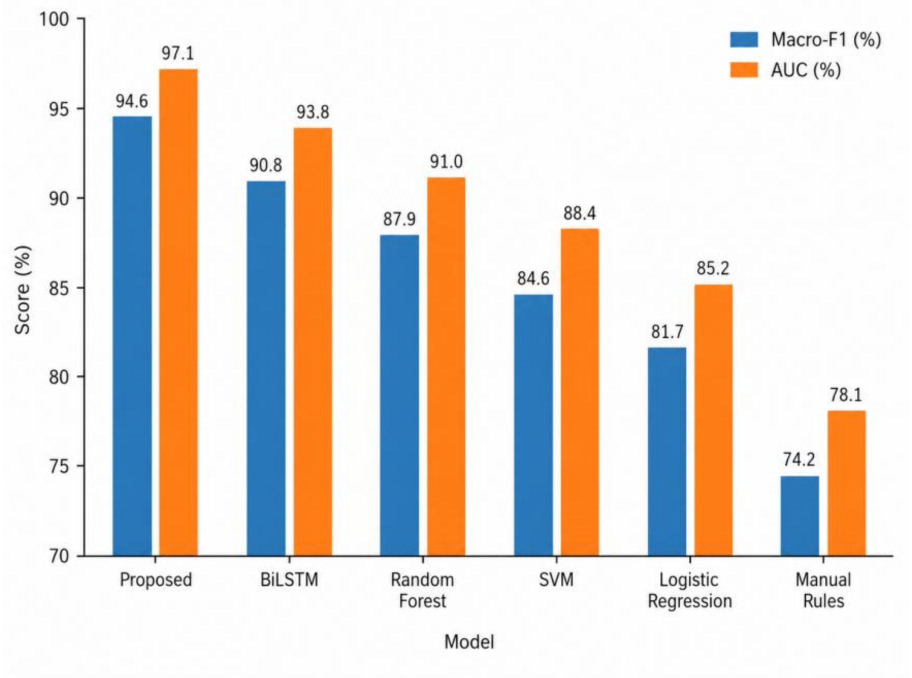

Figure 4: Comparison of aviation safety incident recognition performance of different models

In the event traceability task, the traceability hit rate of the proposed method reaches 92.6%, and the coverage rate of Top-3 risk sources reaches 96.4%, which is significantly higher than that of BiLSTM (84.7% and 89.6%). This result shows that the event correlation graph constructed in this paper can effectively integrate time proximity, semantic similarity and business correlation strength, thereby improving the ability to identify key risk sources. Different from common classification models, the proposed method does not only output event categories such as "unstable approach", "communication anomaly" or "runway conflict", but further gives high-contribution risk sources. For example, in the runway operation conflict sample, the system can simultaneously identify the association chain between runway occupancy status, ATC instruction change, flight density increase and low visibility conditions. In the maintenance release deviation sample, the system can link the duplicate defect records, component maintenance intervals and subsequent operation anomalies to form a more complete event traceability result.

In order to analyze the influence of data scale on model performance, this paper further sets five training data proportions of 20%, 40%, 60%, 80% and 100%, and statistics event identification Macro-F1, traceability hit rate and risk level evaluation Macro-F1 respectively. The results are shown in Figure 5. When the proportion of training data is 20%, the Macro-F1 of event identification is 86.3%, the hit rate of traceability is 78.4%, and the Macro-F1 of risk level assessment is 82.5%. When the proportion of training data increased to 40%, the three indicators increased to 90.1%, 83.6% and 86.4%, respectively. When the proportion of training data reaches 60%, the three indicators further increase to 92.7%, 87.9% and 89.7%. The improvement in this stage is relatively large, indicating that the increase of multi-source aviation safety data can significantly improve the model's ability to learn incident patterns and risk chains.

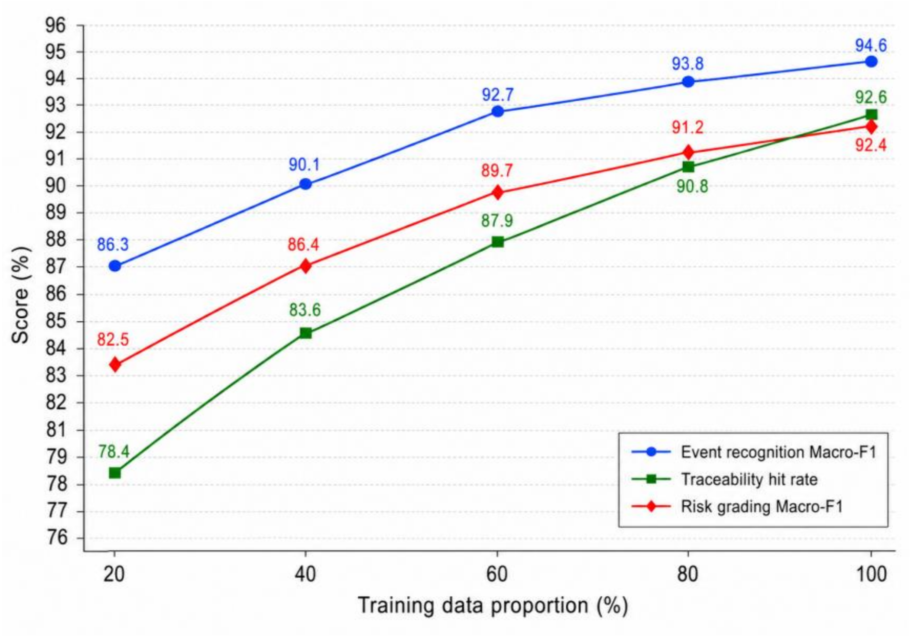

Figure 5: Changes in model stability for different training data sizes

When the proportion of training data is increased from 60% to 80%, the Macro-F1 of event recognition is increased from 92.7% to 93.8%, the hit rate of traceability is increased from 87.9% to 90.8%, and the Macro-F1 of risk level assessment is increased from 89.7% to 91.2%. When the proportion of training data reaches 100%, the three indicators are 94.6%, 92.6% and 92.4%, respectively. Compared with the condition of 20% training data, the event recognition Macro-F1 under the condition of complete data is increased by 8.3 percentage points, the traceability hit rate is increased by 14.2 percentage points, and the risk level assessment Macro-F1 is increased by 9.9 percentage points. It can be seen from Figure 5 that the traceability hit rate is the most sensitive to the data scale, and its improvement is the largest, indicating that the event traceability task relies more on the full connection between cross-system records, historical event samples and multiple types of risk nodes.

From the risk level evaluation results, the risk level consistency rate of the proposed method is 93.1%, which is higher than that of BiLSTM (88.5%), random forest (84.7%), support vector Machine (80.9%), logistic regression (77.4%) and artificial rule method (68.5%). This shows that after the traceability contribution is included in the risk level calculation, the model can more accurately distinguish between slight operational fluctuations and high-risk event chains. For the samples whose single index deviated for a short time but did not form a continuous risk chain, the system mostly judged as grade I or grade II. For the samples accompanied by repeated equipment failures, meteorological disturbances, and ATC interaction anomalies, the system is more likely to identify as class III or class IV. This assessment method can avoid simple classification only based on the severity of the event result, and make the risk judgment more in line with the process control requirements in aviation safety management.

The test results of the system response efficiency are shown in Figure 6. The data preprocessing module took 38 ms on average, the event recognition module took 22 ms on average, the traceability analysis module took 64 ms on average, the risk classification module took 18 ms on average, and the policy matching module took 12 ms on average. The average end-to-end response time of the system was 154 ms. Among them, the traceability analysis module consumes the most time, accounting for 41.6% of the end-to-end response time, mainly because this module needs to complete event node matching, comprehensive edge weight calculation and candidate risk path search. The data preprocessing module took 24.7% of the total time, the event recognition module took 14.3%, and the risk classification and strategy matching module took 19.5%. Although the provenance analysis increases a certain computational overhead, the end-to-end response time of 154 ms can still meet the near-real-time analysis requirements of the aviation safety management platform.

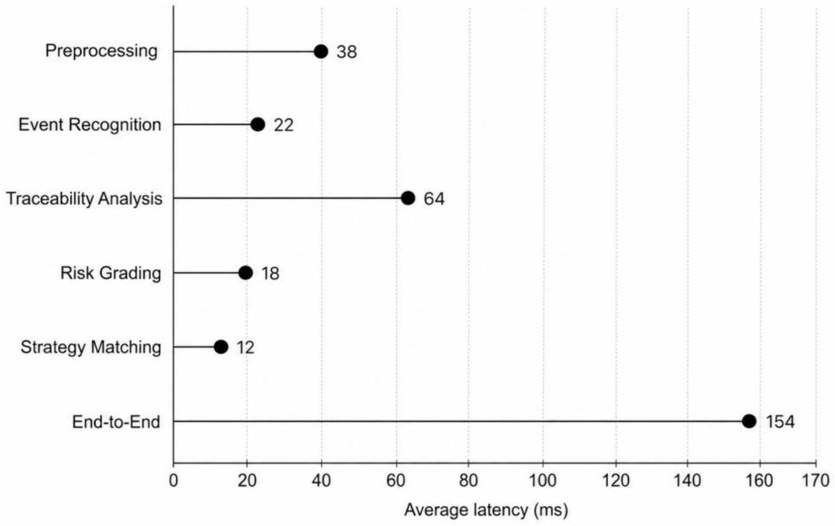

Figure 6: Average response time of each module of the aviation safety risk prevention and control system

5 Discussion

The big data traceability and risk prevention and control system of aviation safety incidents constructed in this paper shows good recognition accuracy, traceability interpretation ability and system response efficiency in the experimental evaluation. Compared with BiLSTM, random forest, support vector machine, logistic regression and artificial rule methods, the proposed method achieves better results in Accuracy, Macro-F1, AUC, traceability hit rate and risk level consistency rate. Among them, the event recognition Accuracy reaches 95.1%, Macro-F1 reaches 94.6%, and AUC reaches 97.1%. The hit rate of traceability reached 92.6%, and the coverage rate of Top-3 risk sources reached 96.4%. The average end-to-end response time of the system is 154 ms. This shows that the unified modeling of multi-source aviation safety data can provide more stable computational support for event identification, cause tracking and risk prevention and control. The above results show that the advantages of the proposed method are mainly reflected in three aspects: multi-source data fusion, correlation traceability modeling and risk closed-loop disposal.

This performance advantage mainly comes from the combination of multi-source data fusion and associated provenance mechanism. Traditional aviation safety analysis often focuses on a single incident report or a single operating parameter. Although it is convenient for manual review, it is difficult to reveal the multi-factor coupling relationship behind the incident. In this paper, flight operation data, incident reports, maintenance records, meteorological information, airport operation logs and air traffic control interaction data are incorporated into the same feature space, and the event correlation graph is constructed through time correlation, semantic correlation and business correlation, so that the system can identify risk sources from multiple dimensions. For example, in the runway operation conflict or unstable approach samples, the model can not only identify the event type, but also further track the relationship between meteorological disturbance, command change, operation deviation and disposal delay, thereby improving the interpretability of traceability results.

From the perspective of risk prevention and control, the value of the proposed method is not limited to improving the classification accuracy, but to transforming the model output into executable security management actions. The risk level assessment module integrates the abnormal intensity of the event, the traceability contribution degree and the historical disposal record into the calculation, so as to form a clear distinction between the low-grade operation fluctuation and the high-risk risk chain. The prevention and control system further matches the monitoring, verification, collaborative disposal and high-level early warning strategies according to the risk level, and promotes the transformation of aviation safety management from post-review to process intervention. However, there are still some limitations in this paper. Although the experimental data cover many types of aviation safety scenarios, there are still differences in data standards between different airlines, airports and air traffic control units, and the cross-regional migration ability of the model still needs to be further verified. Subsequent studies can introduce larger scale real operation data, and combine interpretable artificial intelligence methods to carry out finer-grained analysis of key risk sources, risk propagation paths, and prevention and control strategy effects, so as to improve the suitability and credibility of the system in actual aviation safety governance.

6 Conclusions

Focusing on the problems of lagging identification of aviation safety incidents, unclear traceability chain and lack of closed-loop linkage of risk prevention and control, this paper constructs a big data traceability and risk prevention and control system of aviation safety incidents. Based on flight operation parameters, safety incident reports, maintenance support records, meteorological information, airport operation logs and air traffic control interaction data, the system forms a standardized event feature library through data cleaning, time alignment, semantic coding and feature fusion. On this basis, the incident traceability model is constructed by using big data association analysis, and the risk feature identification and level assessment method are combined to realize the continuous analysis from incident discovery, cause tracking to prevention and control strategy output. Experimental results show that the proposed method is superior to the comparison models in terms of incident recognition Accuracy, Macro-F1, AUC, traceability hit rate and risk level consistency rate, indicating that the system can effectively improve the accuracy and interpretability of aviation safety incident analysis. There are still some shortcomings in this research. The model training relies on high-quality multi-source data, and the differences in data standards between different aviation operation subjects will affect the cross-scene transfer effect. The number of samples of some low-frequency high-risk events is limited, and the verification of composite risk scenarios still needs to be strengthened. Subsequent research can further introduce cross-agency data collaboration, digital twin simulation, and end-to-end joint optimization methods to enhance the generalization ability and application stability of the system in real aviation safety governance.

References

[1] Demir G, Moslem S, Duleba S. Artificial intelligence in aviation safety: systematic review and biometric analysis[J]. International Journal of Computational Intelligence

Systems, 2024, 17(1): 279.

[2] Cankaya B, Topuz K, Delen D, et al. Evidence-based managerial decision-making with machine learning: The case of Bayesian inference in aviation incidents[J]. Omega, 2023, 120: 102906.

[3] Caetano M. Aviation accident and incident forecasting combining occurrence investigation and meteorological data using machine learning[J]. Aviation, 2023, 27(1): 47-56.

[4] Omrani F, Etemadfard H, Shad R. Assessment of aviation accident datasets in severity prediction through machine learning[J]. Journal of Air Transport Management, 2024, 115: 102531.

[5] Haselein B Z, da Silva J C, Hooey B L. Multiple machine learning modeling on near mid-air collisions: An approach towards probabilistic reasoning[J]. Reliability Engineering & System Safety, 2024, 244: 109915.

[6] Bauranov A, Rakas J. Bayesian network model of aviation safety: Impact of new communication technologies on mid-air collisions[J]. Reliability Engineering & System Safety, 2024, 243: 109905.

[7] Odisho E V, Truong D, Joslin R E. Applying machine learning to enhance runway safety through runway excursion risk mitigation[J]. Journal of Aerospace Information Systems, 2022, 19(2): 98-112.

[8] e Silva L C, Murça M C R. A data analytics framework for anomaly detection in flight operations[J]. Journal of Air Transport Management, 2023, 110: 102409.

[9] KAIDI R, AL ACHHAB M, LAZAAR M, et al. Improving the Classification of Airplane Accidents Severity using Feature Selection, Extraction and Machine Learning Models[J]. International Journal of Advanced Computer Science & Applications, 2023, 14(12).

[10] Ghaderi A, Saghafi F. Enhancing pilot vigilance assessment: The role of flight data and continuous performance test in detecting random attention loss in short IFR flights[J]. Journal of Air Transport Management, 2024, 120: 102673.

[11] Mendes N, Vieira J G V, Mano A P. Risk management in aviation maintenance: A systematic literature review[J]. Safety science, 2022, 153: 105810.

[12] Patriarca R, Di Gravio G, Cioponea R, et al. Democratizing business intelligence and machine learning for air traffic management safety[J]. Safety science, 2022, 146: 105530.

[13] Rose R L, Puranik T G, Mavris D N, et al. Application of structural topic modeling to aviation safety data[J]. Reliability Engineering & System Safety, 2022, 224: 108522.

[14] Jonk P, de Vries V, Wever R, et al. Natural language processing of aviation occurrence reports for safety management[J]. arXiv preprint arXiv:2301.05663, 2023.

[15] Ukwandu E, Ben-Farah M A, Hindy H, et al. Cyber-security challenges in aviation industry: A review of current and future trends[J]. Information, 2022, 13(3): 146.

[16] Habler E, Bitton R, Shabtai A. Assessing aircraft security: A comprehensive survey and methodology for evaluation[J]. ACM Computing Surveys, 2023, 56(4): 1-40.

[17] Tim H S K, Frank M, Karin B, et al. Towards cybersecurity risk assessment for future ATM–first steps[C]//2024 AIAA DATC/IEEE 43rd Digital Avionics Systems Conference (DASC). IEEE, 2024: 1-8.

[18] Faye S, Abdulrahman J, Talb R A, et al. Cybersecurity in Aviation: A Case-Based Approach to Preparedness[J]. Int'l J. Info. Sec. & Cybercrime, 2024, 13: 33.

[19] Luettig B, Akhiat Y, Daw Z. ML meets aerospace: challenges of certifying airborne AI[J]. Frontiers in Aerospace Engineering, 2024, 3: 1475139.

[20] Massé E, Bartheye O, Fabre L. Classification of electrophysiological signatures with explainable artificial intelligence: The case of alarm detection in flight simulator[J]. Frontiers in Neuroinformatics, 2022, 16: 904301.

[21] Somon B, Giebeler Y, Darmet L, et al. Benchmarking cEEGrid and solid gel-based electrodes to classify inattentional deafness in a flight simulator[J]. Frontiers in Neuroergonomics, 2022, 2: 802486.

[22] Ziakkas D, Pechlivanis K. Artificial intelligence applications in aviation accident classification: A preliminary exploratory study[J]. Decision Analytics Journal, 2023, 9: 100358.