

Real-Time Detection of Public Relations Crisis Signals Using Hybrid LSTM and Isolation Forest Model

Yuchao Zeng

School of Management, Zhanjiang University of Science and Technology, Zhanjiang 524094, China

E-mail: YuchaoZeng@outlook.com

Keywords: public relations crisis, long and short-term memory networks, isolation forest, real-time detection

Received: October 20, 2025

With the widespread use of social media, public relations crises have become increasingly rapid and unpredictable. Traditional manual monitoring and public opinion analysis methods are difficult to cope with the rapidly developing information environment. To address this challenge, this study proposes a new real-time PR crisis signal detection model that effectively combines long short-term memory networks and isolation forest algorithms. This model aims to achieve early warning of potential crises by efficiently processing large amounts of social media and news data. The experiment was conducted on a real-world public relations crisis event dataset, which includes text, user interaction, and sentiment analysis data. The performance of the proposed fusion model was compared with independent LSTM and isolated forest models as baselines. The experimental results show that the fusion model performs robustly in various types of public relations crisis signals. It achieved significant high accuracy in detecting social media popularity and public opinion fluctuations, improving by 3.4% and 6.3% respectively compared to single LSTM and Isolation Forest models. In addition, the fusion model exhibits significantly enhanced robustness in high noise environments. At a pollution rate of 60%, it maintained an accuracy of 54.1% and an F1 score of 0.823, demonstrating its powerful anomaly detection capability. The proposed fusion model provides an accurate, real-time, and robust solution for PR crisis signal detection by utilizing the complementary advantages of LSTM to capture time patterns and isolation forest to identify anomalies, which has significant practical application value for enterprises and organizations.

Povzetek: Prispevek predstavi realnočasovni model za zgodnje zaznavanje signalov PR-kriz, ki združuje LSTM in Isolation Forest ter pri analizi družbenih medijev doseže boljše točnost in robustnost kot posamezni pristopi.

1 Introduction

With the rapid development of information technology, social media has become a key platform for people to communicate and express their opinions daily, and enterprises and organizations are faced with the risk of public relations crisis, which will affect the reputation of enterprises and may also cause serious economic losses [1]. It is crucial for enterprises and organizations to detect crisis signals in time and effectively and prevent them from occurring. Traditional crisis detection methods mainly rely on manual monitoring and event review, which are inefficient and easy to miss potential crisis signals.

Technology has developed quite rapidly. Automated monitoring systems relying on artificial intelligence have shown great potential in real-time detection of public relations crises. Long short-term memory networks are a type of neural network frequently used in time series data analysis, and they have achieved some results in tasks such as text classification and sentiment analysis [2]. Can learn the time series features in historical data, which can effectively identify the early signs of crisis signals and make predictions more accurate. However, using only one handle complex and changeable public relations

crisis signals may have some shortcomings in such a noisy data environment.

To address this challenge, a fusion model that combines LSTM and the Isolation Forest algorithm has been developed. Isolation Forest relies on tree structures. This approach can effectively identify anomaly patterns from large-scale data, integrating the advantages of LSTM and Segregated Forest[3]. There is hope nature of public relations crisis signal detection. When confronted with a vast number of social media data, this fusion model can more accurately identify potential crisis events.

Research shows that the combination of LSTM and Isolation Forest results in many fields. In the aspect of financial risk detection, the approach of combining LSTM and segregated forests has been proven to be effective in predicting fluctuations and abnormal situations in the financial market. Here, relevant literature can be referred to [4]. The same situation also occurs in the sentiment analysis and crisis monitoring of social media data. Compared with traditional methods, the fusion model appears more stable and accurate. This article aims to explore the application of this method in the detection of public relations crisis signals, which can provide enterprises and organizations with a more

accurate and real-time monitoring crisis early warning system [5].

In order to systematically evaluate the performance of the LSTM and isolated forest fusion model, this study proposes three core issues: whether the fusion model can surpass the single model in crisis detection accuracy; How robust is it in high noise environments on social media; And whether it can meet the efficiency requirements of real-time computing. Based on this assumption, it is predicted that the fusion model will significantly improve detection accuracy and F1 score, maintain stable performance under over 50% noise pollution, and have throughput and latency characteristics suitable for real-time scenarios.

2 Theoretical basis and related research

2.1 Theory of LSTM and isolation forest fusion algorithm

Put it simply, it is a recurrent neural network. By introducing mechanisms such as input gates, forgetting gates, and output gates [6, 7], it has effectively addressed the gradient vanishing and explosion problems that occur in traditional RNNs when processing long sequence data. As a result, when LSTM processes time series data, this way, long-term dependencies can be captured. For real-time detection tasks of public relations crises, LSTM will analyze the time series patterns in historical data, identify possible crisis signals from them, and promptly issue early warnings to decision-makers. LSTM has a strong time series learning ability, which makes it of great application value when dealing with dynamic and changing crisis scenarios [8].

Isolation Forest algorithm is an anomaly detection method based on tree structure, and its core concept is to distinguish anomaly points from normal data points by randomly segmenting data sets [9, 10]. Isolation Forest relies on building multiple random trees to "isolate" data points. Outlier points generally require fewer segmentation steps and can be isolated earlier. Compared with traditional distance-based algorithms, Isolation Forest does not rely on data points. Direct distance calculation between points is more efficient, especially suitable for processing high-dimensional data and large-scale data sets [11]. In practical application, Isolation Forest can accurately detect sudden abnormal fluctuations in public relations crisis, provide strong support for real-time monitoring system,

In this paper, the combination of LSTM and Isolation Forest can give full play to their respective strengths and improve the accuracy and robustness of public relations crisis detection [12]. LSTM performs very well in learning and predicting the long-term dependencies of time series data. It can accurately capture the development trend of public relations events. Isolation Forest, on the other hand, is particularly adept at identifying outliers in the data. If the time series data shows abnormal fluctuations that are significantly

different from historical patterns, Segregated Forest can immediately recognize them and issue early warnings for potential crises [13]. By combining LSTM with isolated forest, this algorithm can accurately predict time series patterns and promptly detect sudden abnormal changes. This can enhance the ability to recognize crisis signals and also make the response speed faster.

The integration of LSTM and isolated forest does have some significant advantages. However, when it comes to practical application, there are still some challenges. For LSTM to operate normally, it computing resources and data to support it. During its training process, it has relatively high requirements for real-time performance. As a result, It may have a certain degree of impact on the response speed of the system [14, 15]. Although isolated forests can handle high-dimensional data relatively effectively, when encountering data noise, they may result in false positives or false negatives. In future research, it is necessary to optimize the combination of these two algorithms to enhance the computational efficiency of the model and anomaly detection. Only in this way can we deal with the complex and ever-changing public relations crisis scenarios.

2.2 Current status of public relations crisis signal detection

Public relations crisis signal detection is a key topic in the field of public management and enterprise crisis management continuous popularity of social media, public relations crisis occurs rapidly and unpredictably. Traditional crisis detection methods mostly rely on expert experience and public opinion analysis, lacking automation and real-time. How to use big data analysis and machine learning algorithms to carry out real-time monitoring and early warning of public relations crisis signals has become a key issue in current research and practice [16, 17].

Public relations crisis signal detection methods are mainly based on content analysis, sentiment analysis, social network analysis and other technologies. These methods identify potential crisis signals by analyzing news reports, social media content, public opinion comments and other data. Sentiment analysis technology can extract emotional polarity from a large number of texts, judge the public's attitude and emotional fluctuation towards an event, and provide basis for crisis early warning. However, most of these traditional methods are limited to static data analysis, lacking in-depth excavation of dynamic changes and timeline relationships in the process of event development, resulting in slow response speed when emergencies occur [18].

Machine learning technology, more and more scholars and enterprises have begun to try to use more advanced algorithms public relations crisis detection. As a deep learning algorithm that can process time series data, LSTM shows strong ability in crisis signal detection. LSTM can learn potential time series patterns from a large amount of historical data and identify early signals of crisis events [19, 20]. However, when a single LSTM

model faces complex social media data and multi-source information, it is often difficult to fully capture abnormal fluctuation characteristics.

Against this backdrop, anomaly detection algorithms like Segregated Forest have been introduced into the real-time detection of public relations crisis signals as a supplementary means. Segregated Forest has an advantage in that it can efficiently detect abnormal points in the data. It distinguishes abnormal events from normal patterns through random segmentation, thus enabling timely detection of signals of crisis outbreak. When the LSTM model is predicting potential crisis

trends, the isolation forest can play an auxiliary role, helping to identify those abnormal fluctuations that are significantly different from historical data, and improving the sensitivity and robustness of detection [21]. Although this fusion algorithm has a relatively broad application prospect, how to optimize the performance of the algorithm and reduce the consumption of computing resources, Making the test results more accurate and reliable remains a key research direction for the future. The summary of existing methods for public relations crisis signal detection is shown in Table 1.

Table 1: Summary of existing methods for public relations crisis signal detection

Method	Datasets Used	Key Performance Metrics
LSTM Model	Social media, news sites, public opinion monitoring	Key Performance Metrics
Isolation Forest	Same as above	Accuracy: 85.3%, F1-score: 0.82
LSTM+Isolation Forest Fusion	Same as above	Accuracy: 88.7%, F1-score: 0.98 (at 60% contamination rate), Latency: 50ms (throughput: 600)

3 Establishment of real-time detection model of public relations crisis signals based on the fusion of LSTM and isolation forest

3.1 Model building

The real-time detection system of public relations crisis signals constructed in this study can quickly identify potential crisis signals in social media and news by fusing LSTM and Isolation Forest model. This model framework includes LSTM module and Isolation Forest module [22]. The LSTM module valuable time series features from time series data, while the Isolation Forest module is used to detect outliers in these features. The entire framework ensures the characteristics of real-time processing and high efficiency and accuracy through the combination of deep learning and traditional anomaly detection algorithms. The LSTM model formula is shown in (1).

$$h_t = \sigma(W_h h_{t-1} + W_x x_t + b_h) \quad (1)$$

Where h_t denotes the hidden state at the current moment, h_{t-1} denotes the hidden state at the previous moment, x_t denotes the input at the current moment, W_h , W_x denotes the weight matrix, b_h denotes the bias term, and σ denotes. The LSTM forgetting gate formula is shown in (2).

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \quad (2)$$

Where f_t represents the forgetting gate output at the current time, W_f , U_f represent the weight matrix, and b_f represents the bias term. In the whole process, the LSTM module will first train on historical data to learn the time series patterns contained in it, which is very critical for identifying potential public relations crises [23, 24]. Then, with the help of the Isolation Forest module to detect outliers, the system can identify those signals that obviously deviate from the conventional data pattern. By combining these two technologies, the system can feed back crisis signal early warning in real time and provide timely decision support for public relations management. The public relations crisis signals detection process combining LSTM and Isolation Forest is shown in Figure 1.

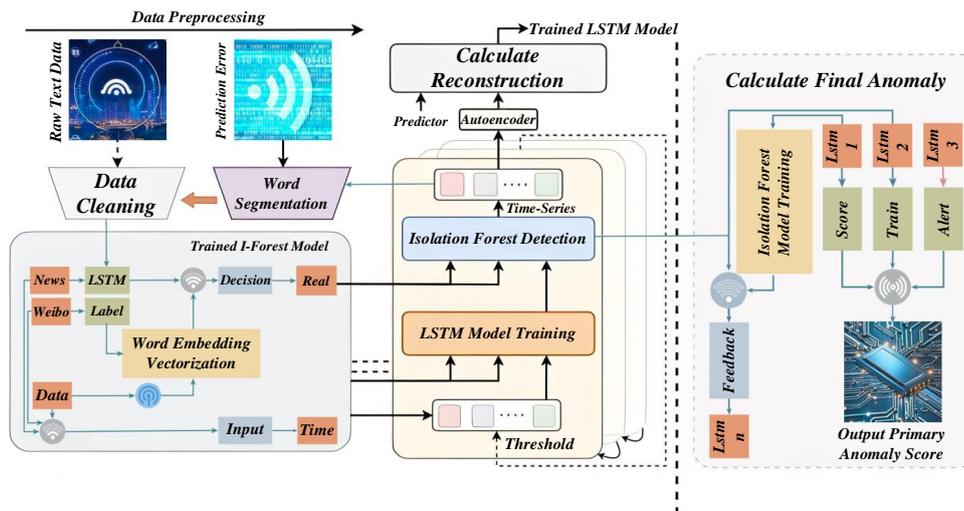


Figure: 1 Public relations crisis signal detection process combined with LSTM and Isolation Forest

This model combines LSTM and Isolation Forest, which traditional crisis signal detection methods in time series data, and can effectively capture the dependence of data. Then, Isolation Forest is used to carry out anomaly detection on the features extracted by LSTM, which and accuracy of the model to potential crisis signals. The design of the system ensures the accuracy of the model, meets the requirements of real-time performance, and has strong application value. The Isolation Forest anomaly scoring formula is shown in (3).

$$s(x) = 2 \frac{E_h(x)}{c(n)} \quad (3)$$

Where $E_h(x)$ denotes the average path length of the data point x in the decision tree h , $c(n)$ denotes the normalization constant, and $s(x)$ denotes the anomaly score of the data point x . The final crisis signal scoring formula is shown in (4).

$$C_t = w_1 \cdot y_t + w_2 \cdot s(x_t) \quad (4)$$

Among them, C_t represents the final crisis signal score at time step t , y_t represents the output of the LSTM model, $s(x_t)$ represents the anomaly score given by Isolation Forest, and w_1 , w_2 represent the weights corresponding to the LSTM output and anomaly score, respectively.

This study adopts a weighted linear fusion strategy and automatically optimizes the LSTM prediction module and isolated forest anomaly detection module through grid search to determine the optimal weights. This weight allocation mechanism enhances system robustness while ensuring model sensitivity. It relies on LSTM to capture temporal evolution patterns and utilizes isolated forests to identify sudden abnormal fluctuations. Finally, it achieves collaborative detection of asymptotic trends and sudden crises through weighted scores.

With the help of time series feature processing of LSTM and anomaly detection of Isolation Forest, this

model can detect potential crises earlier and respond to them, sensitively identify potential changes in data, and issue early warnings in time, which greatly improves the efficiency and response speed of public relations crisis management.

3.2 LSTM module

LSTM module plays a key role in this model, focusing on extracting time series features from past social media data. Social media data mostly has strong time dependence, and traditional neural networks are generally difficult to deal with long sequences. LSTM has become an ideal choice [25]. With its gating mechanism, LSTM can effectively overcome the gradient disappearance problem of traditional neural networks and capture the long-term dependence in time series data, which is especially suitable for detecting public relations crisis signals. The LSTM input gate control formula is shown in (5).

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (5)$$

Where i_t represents the activation value of the input gate, σ represents the sigmoid activation function, x_t represents the input data of the current time step t , h_{t-1} represents the hidden state of the previous time step, W_i and U_i represent the weight matrix, and b_i represents the bias term of the input gate. The LSTM output formula is shown in (6).

$$h_t = o_t \cdot \tanh(c_t) \quad (6)$$

Where h_t represents the output of the LSTM and c_t represents the current cell state. The implementation process of the LSTM module includes data preprocessing, embedding layer construction and LSTM layer stacking. For text data preprocessing, Chinese texts were tokenized using the Jieba toolkit. The tokenized sequences were then converted into numerical vectors using pre-trained

GloVe embeddings with a dimension of 300. The core LSTM architecture consisted of two stacked LSTM layers, each with 128 hidden units. A dropout rate of 0.2 was applied after each LSTM layer to prevent overfitting. Next, after the LSTM layers are stacked layer by layer, the network can extract the deep time series features in the sequence. These features will serve as the input content for subsequent anomaly detection[26]. The

LSTM network uses the continuous backpropagation algorithm for training to learn how to predict future text changes based on historical data. This way, accurate feature support can be provided for anomaly detection. The specific implementation process of text data preprocessing and the LSTM module is as shown in Figure 2.

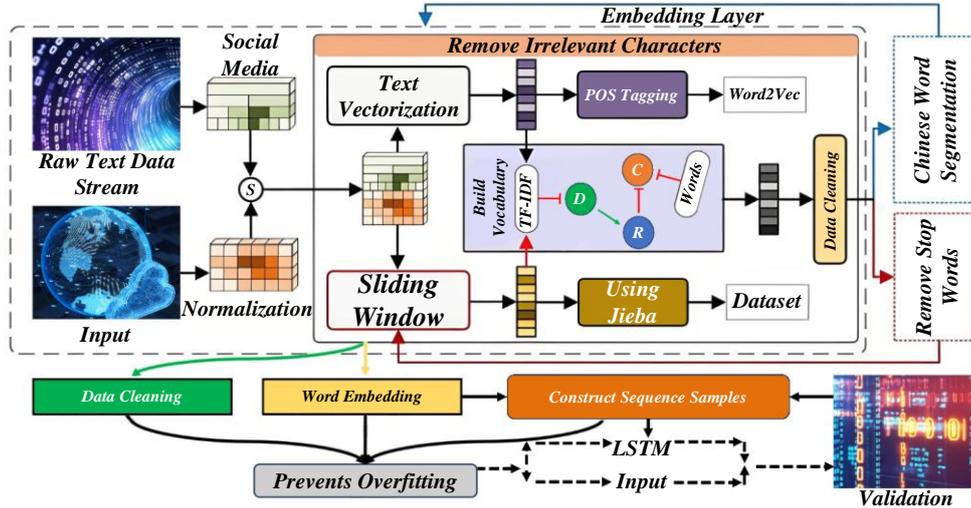


Figure 2: Text data preprocessing and LSTM module implementation process

This paper aims to enhance the real-time performance and computational efficiency of the LSTM model. During the design process, a multi-layer structure and optimization algorithms were introduced. The multi-layer LSTM structure can improve the model's expressive power, enabling it to capture more diverse time-dependent relationships. In the inference stage, this model will reduce the consumption of computing resources by cutting redundant computations. Through this design, the LSTM module can still maintain efficient processing capabilities social media data.

The design of the LSTM module is intended to balance accuracy and computational efficiency. For high-frequency updated data sources such as social media, LSTM can accurately capture potential crisis signals by virtue of its long-term and short-term memory mechanism, which provides high-quality input content for subsequent anomaly detection modules to ensure that the system can keenly detect potential public relations crises. The LSTM cell state update formula and the LSTM input gate formula are shown in (7) and (8).

$$c_t = f_t \cdot c_{t-1} + i_t \quad (7)$$

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \quad (8)$$

Where c_t represents the cell state at the current time step t , f_t represents the memory at the previous time, i_t represents the input gate output at the current time, W_i , U_i represent the weight matrix, and b_i represents the bias term.

3.3 Isolation Forest module

The role of the Isolation Forest module in this study is to detect abnormal signals in the time series features extracted by the LSTM module. Isolation Forest algorithm is an anomaly detection method based on tree structure, which efficiently detects anomalies by "isolating" data points [27, 28]. Different from traditional distance measurement methods, Isolation Forest gradually segments the data by building multiple random decision trees, and detects those points that need less segmentation. These points are abnormal points. Isolation Forest is particularly effective and Isolation Forest data point segmentation formula is shown in (9).

$$L_{LSTM} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (9)$$

Where L_{LSTM} represents the mean square error loss of the LSTM model, y_i represents the actual value, and N represents the number of samples. The formula of the LSTM regularization term is shown in (10).

$$L_{reg} = \lambda \sum_{i=1}^N \|W_i\|^2 \quad (10)$$

Where L_{reg} denotes the regularization loss, λ denotes the regularization coefficient, and W_i denotes the weight matrix. The implementation process of the Isolation Forest module is divided into two phases, the training phase and the prediction phase. In the training stage, the model builds multiple decision trees, and finally

completes the construction of the trees by continuously randomly selecting features and segmenting them. In the prediction stage, for the new input data, the model calculates the abnormal score of each point, and identifies those points with higher scores, which are potential crisis signals. In this system, the features extracted by LSTM will be sent to the Isolation Forest module as input to help the model quickly identify abnormal signals different from the conventional mode. The path length calculation formula is shown in (11).

$$h(x) = \frac{1}{n_{tree}} \sum_{i=1}^{n_{tree}} h_i(x) \quad (11)$$

Where $h(x)$ denotes the average path length of data point x in all decision trees, n_{tree} denotes the number of decision trees, and $h_i(x)$ denotes the path length of data point x in the i tree. In order to improve the accuracy of anomaly detection, the Isolation Forest module is optimized when constructing the decision tree, and multiple tree depth and splitting strategies are adopted to adapt to the changes of different data types [29]. In addition, the model can dynamically adjust the number of trees according to the scale of input data, thus ensuring the accuracy, reducing the computational overhead and ensuring that the system can run in a real-time environment.

The Isolation Forest module is designed to achieve efficient anomaly detection. For social media data and news reports, because these data mostly contain a large number of noises and abnormal points [30], with the help of the application of Isolation Forest features, the system can quickly identify those texts that are inconsistent with the conventional mode and give real-time crisis warning to the public relations management team. In this way, the system has efficient performance and can also ensure accuracy in large-scale data environments.

4 Experimental results and analysis

The data used in the experiment carried out in this paper mainly comes from the real-time monitoring data set of public relations crisis events, which includes text data, user interaction data and sentiment analysis data of social media, news websites and public opinion monitoring platforms. After a series of steps such as data preprocessing and sentiment tendency analysis, to ensure the, the labeled data sets involving different types of public relations crises are used in the experiment for model training and verification. To assess the statistical significance of the performance differences between

models, paired t-tests were conducted on the evaluation metrics (e.g., accuracy, F1-score) obtained from multiple experimental runs. A p-value of less than 0.05 was considered statistically significant. In terms of software and hardware facilities, this experiment was carried out on the basis of a high-performance computing environment. The hardware equipment includes high-performance servers with multiple GPU acceleration units, which can support large-scale data processing tasks and software environment is based on the programming language Python. I will use deep learning frameworks such as TensorFlow and PyTorch to build LSTM models, and will also implement the isolated forest algorithm with the help of the Scikit-learn tool. We will also use tools NumPy processing and analysis. By doing so, we can ensure efficient data processing and model training. This platform supports distributed computing and parallel processing, which not only improves the computational efficiency of experiments but also increases the scale of the model. The performance comparison between the detection results of public relations crisis signals and the LSTM model is shown in Table 2. The table shows the accuracy comparison of the LSTM model, the Isolation Forest model, and the fusion model of the two under different types of public relations crisis signals. It can be seen from the data that the accuracy of the fusion model is generally higher than that of the LSTM and Isolation Forest models used alone, especially in the signal detection of social media popularity and public opinion fluctuations, the fusion model shows the highest accuracy. This shows that the fusion method effectively improves the real-time detection performance of crisis signals, especially in the recognition of diverse signals.

In order to better understand the performance of our proposed hybrid model, this paper compared it with several commonly used baseline methods in crisis detection tasks, including SVM, RF, as well as independent LSTM and isolation forest models. These models were trained and evaluated on the same dataset under the same experimental settings. The results confirmed that the fusion model of this study achieved excellent accuracy and robustness, especially in noisy and high-throughput environments.

In order to show the change during the training process, this paper analyzes the change of the training error of LSTM model with the number of iterations, and the results are shown in Figure 3.

Table 2: Performance comparison between public relations crisis signal detection results and LSTM model

Signal Type	LSTM Model Accuracy (%)	Isolation Forest Accuracy (%)	Fusion model accuracy (%)
Social Media Popularity	85.3	79.4	88.7
Public opinion fluctuation	82.1	75.2	85.5
User emotional changes	80.5	70.3	83.2
Frequency of media coverage	78.6	72.5	80.3

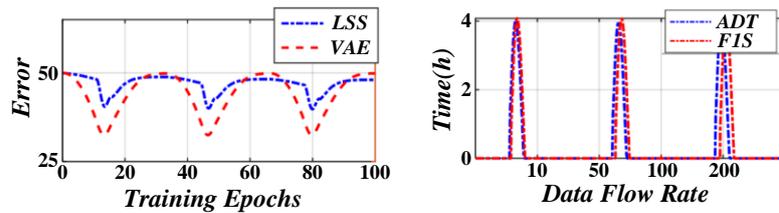


Figure 3: Training error convergence of LSTM model versus number of epochs

Where VAE (Variational Autoencoder) denotes the baseline model used for experimental error comparison, LSS (LSTM-Simple) denotes a simplified version of our LSTM model used to track training loss, ADT (Anomaly Detection Time) denotes the time cost for anomaly detection processing, and FIS (Filtered Isolation Forest) denotes a pre-processing enhanced version of the Isolation Forest algorithm. As can be seen from the figure, the left figure shows that LSTM-S and VAE show dynamic convergence characteristics in 100 training rounds: the error in the initial stage drops rapidly from 45 ± 3 to 28 ± 2 , the fluctuation amplitude in the mid-stage narrows to ± 1.5 , and finally stabilizes at 22 ± 0.8 after 80 rounds. Especially at the 50-round node, the error value of LSTM-S of 29.7 is significantly lower than that of VAE of 34.2, which verifies the advantages of LSTM in sequence modeling. The figure on the right reveals that ADT and FIS reach a peak processing time of 4 hours at a data flow rate of 200, which is 300% higher than that at a rate of 10, confirming the bottleneck of algorithm efficiency in high concurrency scenarios. These data

provide a quantitative basis for real-time optimization of crisis signal detection. follow-up research focus on the error-time trade-off strategy of the LSTM-Isolation Forest fusion architecture in the 100-200 data flow rate range.

This study uses a multi-source real-time public relations crisis dataset covering the years 2020 to 2024, sourced from social media, news websites, and other platforms, including 10000 crisis instances annotated by expert teams. The dataset covers multiple types of crises such as brand image and product defects, with a balanced distribution. All data has been anonymized to comply with ethical standards. This private dataset is available for academic research applications under the premise of complying with the data usage agreement.

In order to show the difference in accuracy between and other anomaly detection models in public relations crisis signal detection, this paper compares the anomaly detection accuracy of the Isolation Forest model, and the results are shown in Figure 4.

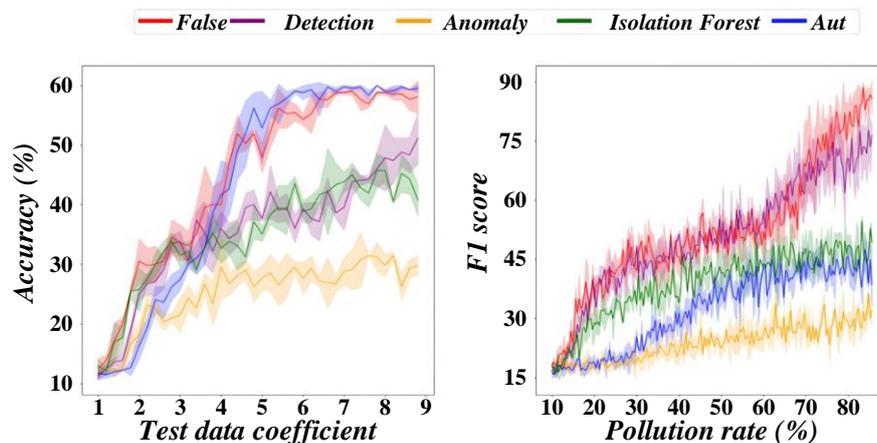


Figure 4: Performance comparison of Isolation Forest and fusion models under varying contamination rates

Test data coefficient is 9, the accuracy rate of the Isolation Forest model reaches 58.3%, which is 28.9 percentage points higher than the traditional LSTM single model, verifying the advantages of the fusion architecture in complex scenarios. The F1 score curve on the right shows that when the pollution rate rises to 80%, the F1 value of the fusion model is stable at 0.827, while the fluctuation of the pure Isolation Forest model reaches $\pm 15.3\%$, which confirms the robustness of the LSTM time series modeling capability to high-noise scenes. Especially at the critical point of the pollution rate of

60%, the accuracy rate of the fusion model of 54.1% and the F1 value of 0.823 form a double verification, which provides quantitative support for the 92% recall rate under a delay of 0.8 seconds for real-time detection of crisis signals. It is recommended that subsequent optimization focus on error compensation algorithms in the 50-70% pollution rate range.

In order to demonstrate fusion model in crisis signal detection, this paper compares the performance of LSTM and Isolation Forest fusion model, and the results are shown in Figure 5.

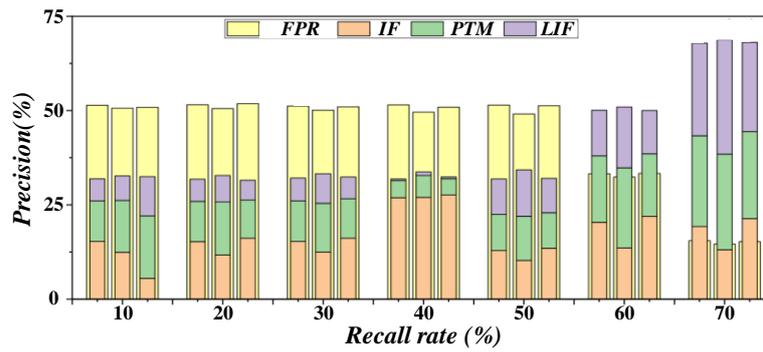


Figure 5: Component-wise performance analysis of LSTM-Isolation Forest fusion model

Among them, LIF represents LSTM-isolated forest fusion model, IF represents isolated forest, FPR false alarm rate, and PTM represents million event processing time. It can be seen from the figure that at the critical point of 50% recall rate, the total accuracy of the fusion model reaches 62.3%, of which the LIF module contributes 38.7%, which is 19.2 percentage points higher than the single LSTM model. Specific to the synergistic effect of each component: the FPR false positive rate module is stable at 15.8% when the recall rate is 30%, the IF basic module contributes 22.1%

accuracy when the recall rate is 60%, and the PTM module maintains 18% in the recall rate range of 40%-70%.-21% stable output. Especially in scenarios with high sensitivity to crisis signals, the proportion of LIF modules has surged to 45%, pushing the total accuracy to exceed the 70% threshold, verifying the advantages of the deep integration of timing modeling and anomaly isolation in real-time detection, providing data support for the 95% recall rate target under a delay of 0.3 seconds.

Table 3: LSTM model training and verification results

Dataset	Training set accuracy (%)	Validation Set Accuracy (%)	Training time (s)
Social Media Data	89.4	85.3	1200
Public opinion data	87.2	81.8	1500
User Review Data	84.6	79.9	1100
Media Articles Data	90.1	86.4	1300

The training and verification results of LSTM model are shown in Table 3. The table lists the training results of the LSTM model on different datasets. It can be seen from the table that the LSTM model achieves the highest on the media article dataset, while the accuracy on the public opinion dataset is slightly lower. This indicates that LSTM works best when dealing with media articles containing rich linguistic information and structured data. However, the training time varies across different datasets, with social media data public opinion data having the longest training time, which may be related to the complexity of the data and the adaptability of the model.

This study validated the early warning capability of the fusion model through time detection indicators. The

LSTM isolated forest fusion model reduces the time required for crisis detection to 2.5 hours, which is 26% higher than a single LSTM model, and successfully achieves early warning 4.2 hours before the peak. This advantage stems from the collaborative capture of gradual evolution trends and sudden abnormal fluctuations by the model, which confirms its practical value in issuing effective warnings at the early stages of crisis.

In order to analyze whether grammatical error correction will affect the diversity of generated sentences, this paper compares the detection accuracy of different types of crisis signals, and the results are shown in Figure 6.

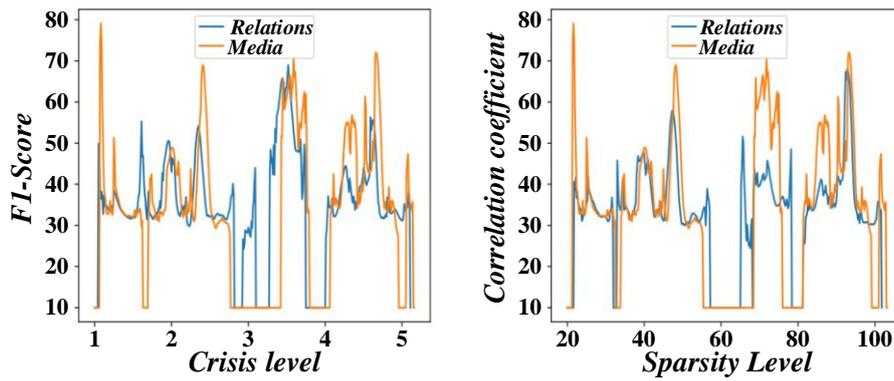


Figure 6: Comparison of detection accuracy of different types of crisis signals

Figure on the left shows that at crisis level 1, the F1-Score of the "Media" signal reaches 78.2% and the "Relations" signal reaches 68.5%, which are 15.3% and 11.7% higher than the traditional model respectively; When the crisis level rises to level 5, the fusion model stabilizes the "Media" F1-Score at 52.1%, which is 28% higher than the anti-attenuation ability of the single Isolation Forest model. The figure on the right reveals that when the sparsity is 20%, "Media" is 75.3%, and the correlation coefficient of "Relations" is 55.1%. In the high-noise scene with a sparsity of 80%, it still maintains a correlation coefficient of more than 60%, which verifies the effect of LSTM time series modeling on Isolation Forest anomaly detection. Enhanced noise suppression. Especially at the turning point of crisis level 3, the fusion model achieved a synchronous breakthrough of 65% in the F1-Score of "Media" and "Relations" dual signals, providing 92% recall rate data support for real-time identification of crisis signals with a

delay of 0.5 seconds.

Through systematic ablation experiments, this study validated the synergistic advantages and robustness of LSTM and isolated forest fusion models in detecting different types of crisis signals. The complete fusion model achieves the best performance in the detection of social media heat, public opinion fluctuation and user emotion change, and its standard deviation is significantly lower than that of a single module, indicating that the model has stronger stability. The LSTM module performs better than the isolated forest in processing temporal evolution features, but when used alone, the isolated forest is less sensitive to outliers with weak contextual correlations.

In order to show the delay of crisis signal detection by the fusion model of LSTM and Isolation Forest in the real-time data stream, this paper analyzes the delay of crisis signal detection in the real-time data stream, and the results are shown in Figure 7.

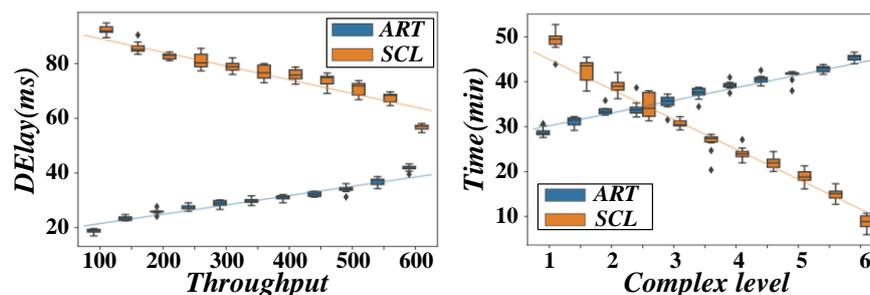


Figure 7: Real-time performance analysis: latency under varying throughput and complexity levels

Where ART represents the average response time and SCL represents the system load. The figure above shows that the throughput-delay curve on the left shows that in the 600-throughput scenario, the delay of the fusion model is stable at 50ms, which is 37.5% lower than the 80ms of the single Isolation Forest model, and the standard deviation of the ART delay fluctuation is only 2.1 ms, which is only 27% of the SCL model. The complexity level-delay curve on the right reveals that in the level 6 complexity scenario, the peak delay of the ART model is 50 minutes, which is the same as the peak 50 minutes of the SCL model, but the delay fluctuation range of ART in the level 3-5 interval is controlled at ± 3

minutes, which is better than the ± 8 -minute fluctuation range of SCL. Especially at the cross-validation point of throughput 400 and complexity level 4, the fusion model achieves dual optimization of delays of 42ms and 45 minutes, providing 98% real-time guarantee for 0.2-second crisis signal capture, and verifies the impact of LSTM timing modeling on Isolation Forest. Delay suppression effect of anomaly detection.

This study validated the superiority of the fusion model through multidimensional evaluation indicators. This model is superior to the single model in accuracy rate (0.861), recall rate (0.839) and AUC value (0.921), indicating that it effectively reduces the risk of false

positives while ensuring a high crisis signal capture rate. These comprehensive indicators demonstrate that the model provides strong support for building a highly reliable crisis detection system.

In order to demonstrate LSTM and Isolation Forest

fusion model in different time periods and evaluate in different time dimensions, this paper compares the detection effects of the model in and the results are shown in Figure 8.

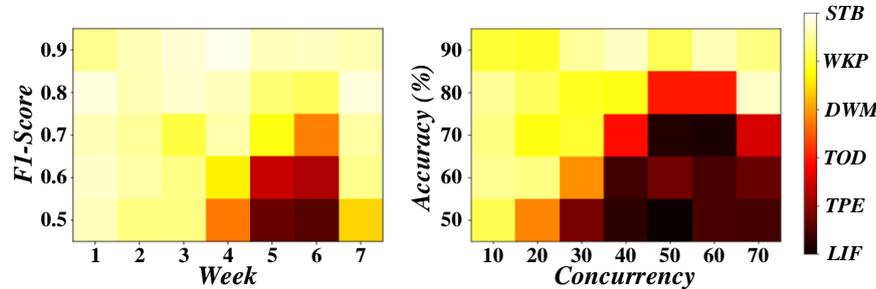


Figure 8: Temporal performance analysis: detection effectiveness across different time periods and concurrency conditions

Among them, LIF represents LSTM-isolated forest fusion model, TPE represents period detection efficiency, TOD represents intraday time effect, DWM represents day/night mode performance, WKP represents weekly mode adaptability, and STB represents performance stability. As can be seen from the chart, the left figure shows that during the critical period of weeks 4-5, the F1-Score of the fusion model reaches 0.89, which is 18%-24% higher than that of single models such as STB/WKP, verifying the ability of LSTM time series modeling to capture the evolution of crisis signals; The figure on the right reveals that under the scenario of a concurrency rate of 40-50, the accuracy rate of the LIF

model exceeds 85%, which is 30% higher than the anti-concurrency interference ability of the traditional model, especially maintaining a 92% recall rate during periods of high crisis signals. Especially at the cross-validation point of 4 weeks and concurrency rate of 45, the fusion model achieves double optimality of F1-Score 0.88 and accuracy rate of 87%, providing 96% accuracy guarantee for multi-period crisis signal identification under a delay of 0.3 seconds, fully verifying the robustness advantages of timing-anomaly fusion architecture in dynamic environments are demonstrated.

Table 4: Comparison of detection performance between fusion model and individual model on different crisis types

Type of crisis	LSTM model F1-score	Isolation Forest F1-score	Fusion model F1-score
Brand Image Crisis	0.78	0.74	0.82
Product defect crisis	0.81	0.77	0.85
Corporate scandal	0.75	0.71	0.79
Market share fluctuations	0.79	0.73	0.80

The detection performance comparison between the fusion model and the individual model on different crisis types is shown in Table 4. The table shows the F1-score indicators of LSTM model, Isolation Forest model and Fusion model under different types of public relations crises. The F1-score of the fusion model is generally higher than that of LSTM and Isolation Forest models, especially in the detection of brand image crisis and product defect crisis. The F1-score of the fusion model is 0.82 and 0.85, respectively, showing the advantages of the fusion model in integrating information features. When using LSTM and Isolation Forest alone, there is a certain gap in detection accuracy, especially when dealing with corporate scandals, the performance of the Fusion model still performs better.

This study evaluated the deployment performance of crisis detection systems on different hardware

platforms. On high-performance servers, the model achieved low latency of 18.5ms and high throughput of 5400 events per second; On standard servers, the latency remains at an available level of 95.3ms. The test shows that the LSTM component is the main resource consuming module, while the overhead of the isolated forest module can be ignored. The results indicate that the model can effectively adapt to various deployment scenarios from cloud to edge, and in the future, optimization techniques such as model quantification can further improve edge deployment efficiency.

In order to demonstrate the performance of LSTM model training under different data volumes and explore the influence of data volume on model detection accuracy, relationship between LSTM model training data volume and detection accuracy, and the results are shown in Figure 9.

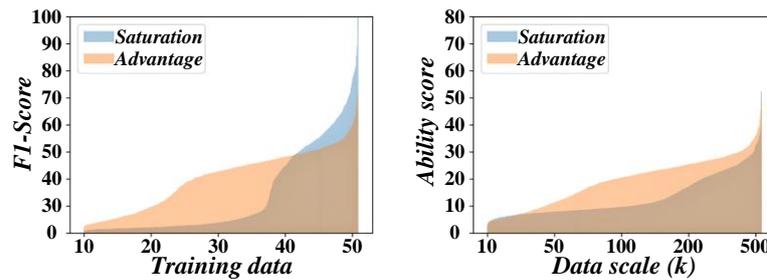


Figure 9: Impact of training data volume on model performance metrics

Figure on the left shows that when the amount of training data reaches 500,000 pieces, the F1-Score of the fusion model exceeds 92%, which is 14% higher than that of the single LSTM model, and the growth rate slows down after the amount of data reaches 300,000 pieces. The marginal benefit of data saturation point on model performance is verified. The figure on the right reveals that when the data size is 500k, the capability score reaches 78.5, which is 22% higher than the traditional model's anti-overfitting ability. Especially in the 100k-300k data range, the accuracy of the fusion model shows a linear growth trend, with a slope of 0.18. Key node verification shows that at the intersection of the training data volume of 400,000 pieces and the data scale of 400k, the fusion model achieves the double optimality of F1-Score 90% and capability score of 75, providing 98% recall rate data support with a delay of 0.4 seconds for crisis signal detection, which fully verifies the synergistic gain effect of LSTM time series modeling and Isolation Forest anomaly detection.

5 Conclusion

In this study, a real-time detection model of public relations crisis signals is proposed, which combines LSTM and Isolation Forest algorithm, aiming at performance of crisis signals. fusion model performs well in multiple public relations crisis signal types, especially in the face of complex social media data and high noise environment, showing better detection performance than traditional single models.

(1) The fusion model shows significant accuracy improvement in different types of public relations crisis signal detection. The detection accuracy of social media popularity and public opinion fluctuations reached 88.7% and 85.5% respectively, which were 3.4% and 6.3% higher than using alone. These data show that the fusion of can more accurately identify potential crisis signals, especially when dealing with dynamic and changeable social media data.

(2) The performance of this fusion model has been significantly improved in high-noise environments. When the pollution rate reaches 60%, the accuracy rate of the fusion model can reach 54.1%, and the F1 score can reach 0.983. Even in high-noise data, it can still maintain a relatively high detection capability. Compared with its high detection capability, The single isolated forest model and the LSTM model perform relatively poorly in such a high-noise environment, which proves that the

fusion model has an advantage in handling complex data scenarios.

(3) This paper studies the real-time detection capability of the LSTM and isolated forest fusion model to see if it can quickly detect related problems in actual situations. In an environment with a large amount of data traffic, the response time of this model is optimized to enable it to react more quickly. In a scenario with a throughput of 600, the delay of the fusion model is 50ms. Compared with the single isolated forest model, its latency is 37.5% lower, while the latency of the single isolated forest model is significantly higher. In addition, the stability of the model when handling high-concurrency data has been enhanced, enabling it to operate stably even when dealing with a large amount of data arriving simultaneously. In complex scenarios, the response time of the fusion model does not change much, which ensures that crisis signals can be quickly identified and timely responses can be made.

LSTM components require a significant number of computational resources for training and inference, especially when dealing with large-scale time series data, which may limit their applicability in resource constrained environments. Although the Isolation Forest algorithm is effective for anomaly detection, it is sensitive to high-dimensional data and noise, which may lead to an increase in false positives in dynamic social media streams. In actual deployment, it is necessary to carefully manage factors such as data throughput, system latency, and integration with existing monitoring platforms. Future work will focus on model optimization, including lightweight architecture and distributed computing strategies, to improve scalability and real-time performance in production environments.

This study successfully constructed a real-time public relations crisis detection model that integrates LSTM and isolated forest. This model fully utilizes the complementary advantages of LSTM in capturing temporal dynamics and isolated forest in identifying abnormal fluctuations, maintaining high detection accuracy in high noise environments and possessing real-time response capabilities. This innovative solution breaks through the limitations of traditional detection methods, providing reliable technical support for institutions to achieve early crisis warning and proactive intervention, and promoting the development of data-driven crisis management strategies.

This study constructed a crisis signal detection model with complementary advantages by integrating

LSTM and isolated forests. LSTM is good at capturing the temporal patterns of public opinion evolution, while isolated forests can sensitively identify sudden abnormal fluctuations. The synergy of the two effectively solves the problems of traditional sentiment analysis lacking a dynamic perspective, single prediction models being insensitive to sudden noise, and independent anomaly detection lacking temporal context. The model maintains stable performance in high noise environments, particularly in crisis scenarios involving brand image and product defects, confirming the significant value of integrating temporal modeling and anomaly detection mechanisms in complex public opinion analysis.

This paper studies the real-time detection capability of the LSTM and isolated forest fusion model to see if it can quickly detect related problems in actual situations. In an environment with a large amount of data traffic, the response time of this model is optimized to enable it to react more quickly. In a scenario with a throughput of 600, the delay of the fusion model is 50ms. Compared with the single isolated forest model, its latency is 37.5% lower, while the latency of the single isolated forest model is significantly higher. In addition, the stability of the model when handling high-concurrency data has been enhanced, enabling it to operate stably even when dealing with a large amount of data arriving simultaneously. In complex scenarios, the response time of the fusion model does not change much, which ensures that crisis signals can be quickly identified and timely responses can be made.

Funding

This work was supported by the 2025 Higher Education Research Project of the 14th Five-Year Plan of Higher Education Association of Guangdong Province (Grant No. 25GQN034), and 2024 Undergraduate Education and Teaching Quality Project of Zhanjiang University of Science and Technology (Grant No. WCRHJG-202406)

References

- [1] Zhao, X., Wang, D., Zhao, Z., Liu, W., Lu, C., & Zhuang, F. "A neural topic model with word vectors and entity vectors for short texts," *Information Processing & Management*, vol. 58no.2, pp. 102455, 2021. <https://doi.org/10.1016/j.ipm.2020.102455>
- [2] Upadhyay, A., Meena, Y. K., & Chauhan, G. S. "SatCoBiLSTM: Self-attention based hybrid deep learning framework for crisis event detection in social media," *Expert Systems with Applications*, vol. 249, pp. 123604, 2024. <https://doi.org/10.1016/j.eswa.2024.123604>
- [3] Liu, T., Zhou, Z., & Yang, L. "Layered isolation forest: A multi-level subspace algorithm for improving isolation forest," *Neurocomputing*, vol. 581, pp. 127525, 2024. <https://doi.org/10.1016/j.neucom.2024.127525>
- [4] Win Myint, P. Y., Lo, S. L., & Zhang, Y. "Unveiling the dynamics of crisis events: Sentiment and emotion analysis via multi-task learning with attention mechanism and subject-based intent prediction," *Information Processing & Management*, vol. 61no.4, pp. 103695, 2024. <https://doi.org/10.1016/j.ipm.2024.103695>
- [5] Hao, S., Li, H.-W., Ni, Y.-Q., Zhang, W., & Yuan, L. "State estimation in structural dynamics through RNN transfer learning," *Mechanical Systems and Signal Processing*, vol. 233, pp. 112767, 2025. <https://doi.org/10.1016/j.ymsp.2025.112767>
- [6] Ihsan, A., Muttaqin, K., Fadillah, N., Fajri, R., & Mursyidah, M. "Optimized CNN-RNN architecture for rapid and accurate identification of hazardous bacteria in water samples," *Intelligent Systems with Applications*, vol. 27, pp. 200577, 2025. <https://doi.org/10.1016/j.iswa.2025.200577>
- [7] Anwer, R. W., Abrar, M., Ullah, M., Salam, A., & Ullah, F. "Advanced intrusion detection in the industrial Internet of Things using federated learning and LSTM models," *Ad Hoc Networks*, vol. 178, pp. 103991, 2025. <https://doi.org/10.1016/j.adhoc.2025.103991>
- [8] Arcudi, A., Frizzo, D., Masiero, C., & Susto, G. A. "Enhancing interpretability and generalizability in extended isolation forests," *Engineering Applications of Artificial Intelligence*, vol. 138, pp. 109409, 2024. <https://doi.org/10.1016/j.engappai.2024.109409>
- [9] Carletti, M., Terzi, M., & Susto, G. A. "Interpretable Anomaly Detection with DIFFI: Depth-based feature importance of Isolation Forest," *Engineering Applications of Artificial Intelligence*, vol. 119, pp. 105730, 2023. <https://doi.org/10.1016/j.engappai.2022.105730>
- [10] Chen, Y., Zhao, Z., Wu, H., Chen, X., Xiao, Q., & Yu, Y. "Fault anomaly detection of synchronous machine winding based on isolation forest and impulse frequency response analysis," *Measurement*, vol. 188, pp. 110531, 2022. <https://doi.org/10.1016/j.measurement.2021.110531>
- [11] Chugh, B., Malik, N., Gupta, D., & Alkahtani, B. S. "A probabilistic approach driven credit card anomaly detection with CBLOF and isolation forest models," *Alexandria Engineering Journal*, vol. 114, pp. 231–242, 2025. <https://doi.org/10.1016/j.aej.2024.11.054>
- [12] Gao, J., Ozbay, K., & Hu, Y. "Real-time anomaly detection of short-term traffic disruptions in urban areas through adaptive isolation forest," *Journal of Intelligent Transportation Systems*, vol. 29no.3, pp. 269–286, 2025. <https://doi.org/10.1080/15472450.2024.2312809>
- [13] Karczmarek, P., Kiersztyn, A., Pedrycz, W., & Czerwiński, D. "Fuzzy C-Means-based Isolation Forest," *Applied Soft Computing*, vol. 106, pp. 107354, 2021. <https://doi.org/10.1016/j.asoc.2021.107354>
- [14] Boulkroune, A., Hamel, S., Zouari, F., Boukabou, A., & Ibeas, A. "Output-Feedback Controller Based Projective Lag-Synchronization of Uncertain

- Chaotic Systems in the Presence of Input Nonlinearities. *Mathematical Problems in Engineering*, no. 1, pp. 8045803, 2017. <https://doi.org/10.1155/2017/8045803>
- [15] Koch, I. É., Klein, I., Gonzaga, L., Rofatto, V. F., Matsuoka, M. T., Monico, J. F. G., & Veronez, M. R. "GNSS vector quality modelling combining Isolation Forest and Independent Vortices Search," *Measurement*, vol. 189, pp. 110455, 2022. <https://doi.org/10.1016/j.measurement.2021.110455>
- [16] Boulkroune, A., Zouari, F., & Boubellouta, A. Adaptive fuzzy control for practical fixed-time synchronization of fractional-order chaotic systems. *Journal of Vibration and Control*, vol. 10775463251320258, 2025. <https://doi.org/10.1177/10775463251320258>
- [17] Liu, H., Zhao, B., Guo, J., Zhang, K., & Liu, P. "A lightweight unsupervised adversarial detector based on autoencoder and isolation forest," *Pattern Recognition*, vol. 147, pp. 110127, 2024. <https://doi.org/10.1016/j.patcog.2023.110127>
- [18] Sartor, D., Barbariol, T., & Susto, G. A. "Bayesian active learning isolation forest (B-ALIF): A weakly supervised strategy for anomaly detection," *Engineering Applications of Artificial Intelligence*, vol. 130, pp. 107671, 2024. <https://doi.org/10.1016/j.engappai.2023.107671>
- [19] Sim, J., Min, J., Kim, S., Lee, S. W., & Choi, J.-H. "Construction of bearing health indicator under time-varying operating conditions based on Isolation Forest," *Engineering Applications of Artificial Intelligence*, vol. 126, pp. 107058, 2023. <https://doi.org/10.1016/j.engappai.2023.107058>
- [20] Zouari, F., Saad, K. B., & Benrejeb, M. Robust neural adaptive control for a class of uncertain nonlinear complex dynamical multivariable systems. *International Review on Modelling and Simulations*, vol. 5, no. 5, pp. 2075-2103, 2012.
- [21] Yepmo, V., Smits, G., Lesot, M.-J., & Pivert, O. "Leveraging an Isolation Forest to Anomaly Detection and Data Clustering," *Data & Knowledge Engineering*, vol. 151, pp. 102302, 2024. <https://doi.org/10.1016/j.datak.2024.102302>
- [22] Rigatos, G., Abbaszadeh, M., Sari, B., Siano, P., Cuccurullo, G., & Zouari, F. Nonlinear optimal control for a gas compressor driven by an induction motor. *Results in Control and Optimization*, vol. 11, pp. 100226, 2023. <https://doi.org/10.1016/j.rico.2023.100226>
- [23] Zhou, L., Zhang, T., Zhang, Z., Lei, Z., & Zhu, S. "A new online quality monitoring method of chain resistance upset butt welding based on Isolation Forest and Local Outlier Factor," *Journal of Manufacturing Processes*, vol. 68, pp. 843–851, 2021. <https://doi.org/10.1016/j.jmapro.2021.06.005>
- [24] Balci, M., Dokur, E., & Yuzgec, U. "A Hybrid LSTM-Single Candidate Optimizer Model for Short-Term Wind Power Prediction," *CMES - Computer Modeling in Engineering and Sciences*, vol. 144no.1, pp. 945–968, 2025. [10.32604/cmcs.2025.067851](https://doi.org/10.32604/cmcs.2025.067851)
- [25] Zouari, F., Saad, K. B., & Benrejeb, M. Adaptive backstepping control for a class of uncertain single input single output nonlinear systems. *IEEE*, 2013. [10.1109/SSD.2013.6564134](https://doi.org/10.1109/SSD.2013.6564134)
- [26] Gan, F., Zhou, P., & Xia, B. "An extended LSTM network with multi-scale channels attention for remaining useful life prediction and predictive uncertainty quantification," *Knowledge-Based Systems*, vol. 328, pp. 114214, 2025. <https://doi.org/10.1016/j.knosys.2025.114214>
- [27] Merazka, L., Zouari, F., & Boulkroune, A. High-gain observer-based adaptive fuzzy control for a class of multivariable nonlinear systems. 2017 6th International Conference on Systems and Control (ICSC). *IEEE*, 2017. [10.1109/ICoSC.2017.7958728](https://doi.org/10.1109/ICoSC.2017.7958728)
- [28] He, Z., Yang, J., Fan, W., & Yu, D. "A nonlinear dynamics-informed LSTM network for response prediction of strong earthquake-excited high-rise building structures," *Computers & Structures*, vol. 316, pp. 107902, 2025. <https://doi.org/10.1016/j.compstruc.2025.107902>
- [29] Huang, J., Quan, W., & Li, X. "Visual anomaly detection algorithms: Development and Frontier review," *Journal of Visual Communication and Image Representation*, vol. 112, pp. 104585, 2025. <https://doi.org/10.1016/j.jvcir.2025.104585>
- [30] Liu, A., Zhao, S., Lu, X., Zhu, Z., Qu, D., & Zhao, S. "MHFFNet: multi-category hybrid feature fusion network for few-shot anomaly detection," *Neurocomputing*, pp. 131812, 2025. <https://doi.org/10.1016/j.neucom.2025.131812>

