EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Risk Early Warning for Police-Related Online Public Opinion Based on Deep Learning

Juan Wang[1,*], Yuxiang Guan[1], Nan Wang[1], Jie Pan[1] and Peng Zhang[1]

[1]Research Center for Network Public Governance, China People's Police University, Langfang 065000, China

Abstract

INTRODUCTION: Police-related online public opinion is highly sensitive and can easily have a negative impact on social stability.
OBJECTIVES: This paper aims to address the crucial need for early warning systems for the risks associated with police-related online public opinion to ensure social harmony and effectively prevent and resolve major social risks.
METHODS: Based on a literature review and deep learning methods, this research constructs an indicator system from four dimensions, analyzing data from representative police-related online public opinion incidents over the past five years. A CNN-BiLSTM sentiment classification model is built for sentiment analysis, and an optimized SSA-CNN-LSTM-Attention model is used for public opinion risk early warning.
RESULTS: The experimental results demonstrate that the SSA-CNN-LSTM-Attention model has the minimum error.
CONCLUSION: This research provides a theoretical reference for public security organs in responding to and preventing police-related online public opinion.

Keywords: Police-related online public opinion, Deep learning, Risk early warning, Optimization algorithm.

Received on 21 October 2025, accepted on 19 November 2025, published on 04 December 2025

doi: 10.4108/eetsis.10664

1. Research Status

With the popularization of social media, the rapid evolution of highly sensitive police-related online public opinion poses challenges for risk early warning. Traditional methods struggle with its dynamism and complexity, whereas deep learning techniques like Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and attention mechanisms offer new technological avenues for intelligent risk warning.

Traditional research on online public opinion often focuses on propagation [1], guidance [2], and evolution [3]. Early warning models have involved constructing indicator systems based on specific theories [4] and combining them with prediction methods [5]. Machine learning methods such as Support Vector Machines (SVM) [6], Bayesian Networks [7], Neural Networks [8], and Fuzzy Neural Networks [9] have also been applied to predict public opinion risks. For example, Huang et al [8]. used an accelerated genetic algorithm to improve a BP neural network for this purpose.

However, online public opinion data is typically high-dimensional and non-linear, making it difficult for traditional machine learning to capture long-term temporal dependencies. Deep learning, with its powerful feature extraction and sequence modeling capabilities, has become a research focus, applied to tasks like public opinion topic and sentiment analysis [10]. Specifically, CNNs can extract local features, LSTMs can handle long-range dependencies, and attention mechanisms can focus on key information. Intelligently optimizing these models further improves performance.

[*]Corresponding author. Email: wangjuan@cppu.edu.com

Regarding the specific domain of police-related online public opinion, existing research has focused on its characteristics, response strategies, and some preliminary early warning attempts [11]. However, there is still a need for research on combining advanced deep learning fusion models with optimization algorithms to conduct high-precision, automated risk early warning.
To address this gap, this paper constructs and optimizes a fusion model based on CNN, BiLSTM, and Attention to improve the accuracy and timeliness of early warning. The main work includes: first, utilizing a CNN-BiLSTM model for sentiment analysis; second, constructing a baseline CNN-LSTM risk early warning model; and finally, proposing and validating an SSA-CNN-LSTM-Attention model optimized via the Sparrow Search Algorithm (SSA). The performance of the proposed method is evaluated through empirical comparative analysis.

2. Construction of the Risk Indicator System for Police-Related Online Public Opinion

2.1. Analysis of Risk Influencing Factors

Police-related online public opinion stems from incidents involving police misconduct or negligence, attracting significant public attention due to its connection to safety, justice, and the public interest. Consequently, it often necessitates active intervention by public security organs. The associated risks, such as fostering prejudice, damaging reputation, and triggering social disorder, are amplified by social media. Analyzing the influencing factors of these risks is the foundation for constructing a risk indicator system.

Scholars have identified these influencing factors from various perspectives. Deng et al. [12] analyzed from three aspects: the incident, online media, and netizens. Cao and Hou. [13] categorized public opinion risk factors into information factors, information actor factors, and information environment factors. Zhang et al.[14] conducted research focusing on three elements: the public opinion subject, the public opinion object, and the public opinion content. Based on these perspectives, this paper identifies four key participants influencing police-related online public opinion risk: incidents, netizens, online media, and platforms. Based on this framework, a total of 23 quantifiable hourly indicators were selected to construct the risk early warning indicator system. The specific details of these indicators are presented in Table 1.

Table 1. The 23 Quantitative Indicators for Risk Early Warning

ID	Police-related incident	Dimension
I1	Duration of Public Opinion (in hours)	Incident
I2	Hourly Volume of All Related Information	Incident
I3	Presence of Casualties (Binary: 1=Yes, 0=No)	Incident
I4	Incident Sensitivity	Incident
N1	Number of Unique Participating Netizens	Netizens
N2	Hourly Volume of Original Posts	Netizens
N3	Hourly Volume of Retweets	Netizens
N4	Hourly Volume of Comments	Netizens
N5	Hourly Volume of Likes	Netizens
N6	Proportion of Negative Sentiment Posts	Netizens
N7	Proportion of Neutral Sentiment Posts	Netizens
N8	Proportion of Positive Sentiment Posts	Netizens
N9	Number of Key Opinion Leaders (KOLs) Involved	Netizens
N10	Average Follower Count of Participants	Netizens
N11	Emotional Intensity Score (Average)	Netizens

2.2. Construction of the Risk Early Warning Indicator System

The foundation of the police-related online public opinion risk early warning model is its indicator system. Based on a literature review, this paper constructs the system from four dimensions: police-related incidents, netizens, online media, and platforms.
Quantifiable hourly indicators were derived to measure risk, including:Incident Metrics: Duration, quantity of related information, presence of casualties.Netizen Engagement: Number of participants, follower counts, volume of posts, sentiment polarity, and interaction volume (retweets, comments, likes).Media and Platform Activity: Proportion of media posts and distribution across platforms.
Additionally, specific attribute indicators for police-related public opinion are incorporated, such as the existence of historical or similar public opinion incidents and the incident background.

3. Implementation of the Police-Related Online Public Opinion Risk Early Warning Model

3.1. Selection of Police-Related Incidents and Acquisition of Public Opinion Data

Police-related incidents are categorized into two types: those from official police duties (e.g., improper enforcement) and those from officers' personal lives. Representative incidents are listed in Table 2.

Table 2. Police Related Incidents and Their Time Frame

ID	Police-related incident	Time Frame
A1	Elderly woman attacks police with knife, shot and killed by police	September 10, 2023 to September 19, 2023
A2	Ningxia police officer accused of sheltering obscene places and involved gang leader	August 29, 2023 to September 5, 2023
A3	Shaanxi police report officer assaulting disabled person in foot bath shop after drinking	August 27, 2023 to September 11, 2023
A4	Auxiliary police officer hits elderly person with stool	July 19, 2023 to July 23, 2023
A5	Shandong auxiliary police officer threatens to kill citizen	July 3, 2023 to July 11, 2023
A6	Netizen posts alleging rape by police officer	June 28, 2023 to July 7, 2023
A7	Quzhou traffic police post accused of suspected gender discrimination	February 12, 2023 to February 19, 2023
A8	Anti-fraud Officer Lao Chen decides to resign from police duties	April 8, 2022 to April 15, 2022
A9	Starbucks expels on-duty police officers	February 13, 2022 to February 22, 2022
A10	Wife reports police officer husband for suspected illegal activities and disciplinary violations	January 5, 2022 to January 12, 2022
A11	Binzhou traffic police report loopholes in drunk driving checks	September 22, 2021 to September 29, 2021
A12	Anhui Wangjiang woman drowns	December 4, 2020 to December 11, 2020
A13	Shenzhen Bao'an police officer commits domestic violence against girlfriend	December 10, 2019 to December 18, 2019
A14	Fourth Chinese People's Police Day	January 10, 2024 to January 19, 2024
A15	Former Police Chief in Tangshan assault case sentenced to 12 years	January 22, 2024 to January 29, 2024

A total of 450,097 text data entries on police-related incidents were collected via crawlers from Zhihu, Weibo, and Toutiao. Subsequently, all collected public opinion text and multi-dimensional indicator data were standardized and vectorized for model input.

The data collection process for this study followed a systematic methodology to ensure comprehensiveness and relevance. The collection period spanned from December 2019 to February 2024, covering the entire evolution of the representative incidents listed in Table 2. We employed a custom-developed web crawler based on the Python language, integrating technical libraries such as Scrapy and Beautiful Soup, for targeted data scraping from three major Chinese social media platforms: Weibo, Zhihu, and Toutiao. The collection strategy was based on keyword combinations, including general terms like "police" and "officer," as well as entity-specific terms related to each incident (e.g., "Tangshan assault," "Officer Lao Chen"). To ensure the quality of the corpus, the raw crawled data was cleaned by removing irrelevant commercial advertisements and obvious duplicate content. Basic statistics of the corpus are as follows: of the 450,097 total entries, approximately 65% are from Weibo, 20% from Toutiao, and 15% from Zhihu. The entire corpus contains over 50 million tokens, with a vocabulary size of about 180,000 unique words.

3.2. Public Opinion Data Sentiment Classification and Analysis

To address the scarcity of a domain-specific sentiment dictionary for police-related online public opinion, this paper constructs one using word embeddings, as illustrated in Figure 1. The process begins by creating a foundational sentiment dictionary of 41,738 terms by merging mainstream Chinese lexicons like HowNet, DUT, and NTUSD. This dictionary is then expanded for the specific domain. After collecting and preprocessing a police-related text corpus, a Word2Vec (Skip-gram) model is trained. By matching this corpus against the foundational dictionary, 3,959 baseline terms were identified. Subsequently, new terms exhibiting high cosine similarity to these baselines were computed, and after filtering, 896 expansion terms were added to create the final domain-specific dictionary (see Table 3 for examples).

To account for the complex and diverse expressions of netizens, this paper modifies traditional sentiment calculation methods by formulating more reasonable semantic rules (detailed in Table 4). These rules consider the weight of adverbs of degree, the number of adverbs, the number of negation words (m), and the intensity of the sentiment wordto calculate the sentiment of a clause. The overall sentence sentiment is the sum of its clause sentiments, with a multiplier of 2 if an exclamation mark is present. This rule-based approach, including the use of multipliers for punctuation and negation, is adapted from established methodologies in lexicon-based sentiment analysis. The specific parameter values, such as the multiplier of 2 for exclamation marks, were empirically fine-tuned on a validation subset of our annotated data to best capture the intensified emotional expressions commonly found in police-related online discourse. This approach determines the sentiment polarity of police-related online public opinion text.

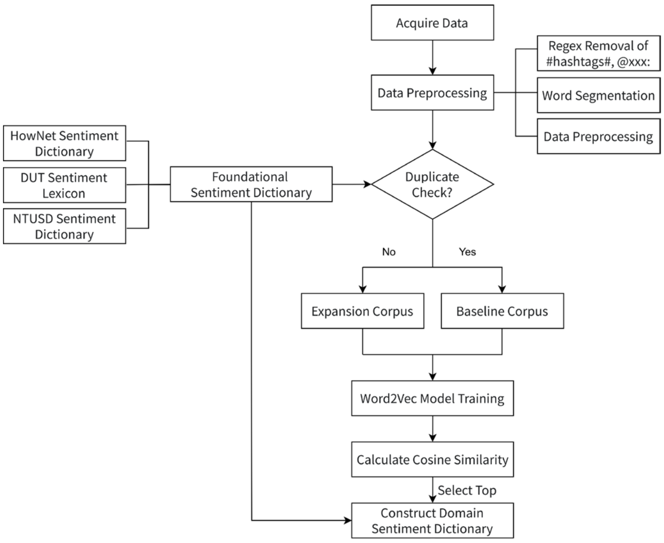

Figure 1. The Process of Constructing an Emotional Dictionary

Table 3. Police Related Incidents and Their Time Frame

Term	Polarity	Term	Polarity
Out of danger	Positive	Evil forces	Negative
Master of all trades	Positive	Black swan	Negative

Table 4. Calculation Rules

Combination Type	Calculation Formula
Sentiment word	$sen_i = p_i$
Negation word + Sentiment word	$sen_i = -p_i$
Adverb of degree + Sentiment word	$sen_i = w{*}p_i$
(Combination of adverb of degree and negation word)+ Negation word + Sentiment word	$sen_i = (-1)^m {*} w^n {*} p_i$
(Combination of adverb of degree and negation word)+ Adverb of degree + Sentiment word	$sen_i = (-1)^m {*} w^n {*} p_i {*} 0.2$

Sentiment Classification Based on CNN-BiLSTM

To provide a high-quality training dataset for the sentiment classification model, we constructed a balanced set of 5,888 manually annotated samples. The annotation process adhered to the following rigorous procedure:

Annotators and Guidelines: The annotation task was performed by three graduate students in computer science with a thorough understanding of public opinion. We established comprehensive annotation guidelines, defining sentiment polarity into three categories: Positive, Negative, and Neutral. "Positive" refers to remarks supporting or praising the police; "Negative" includes criticism, questioning, or expressions of anger; and "Neutral" pertains to objective statements or news reports.

Process and Quality Control: To ensure consistency, we implemented a double-blind independent annotation process. Each entry was classified by two annotators independently. Disagreements were resolved by a third, senior researcher acting as an arbiter. To quantify the reliability of the annotations, we calculated the Inter-Annotator Agreement (IAA), achieving a Cohen's Kappa score of 0.85. This value signifies a high level of agreement, establishing the dataset as a reliable gold-standard for training the model.

For sentiment classification, three methods were compared: a domain-specific sentiment dictionary, a standalone LSTM model, and a CNN-BiLSTM model. The experiment utilized a balanced dataset of 5,888 pre-annotated police-related public opinion entries, which were preprocessed into 200-length sequences and split into training (80%), validation (10%), and testing (10%) sets. As shown in Table 5, the dictionary approach yielded an accuracy of 0.5888 and the LSTM model achieved 0.8787, while the CNN-BiLSTM model attained a superior accuracy rate. Its superior performance is attributed to its use of word embeddings, which mitigate data sparsity, and its BiLSTM component, which effectively processes sequential features. Consequently, the CNN-BiLSTM model was selected for the subsequent sentiment analysis of all text data.

The architecture of our CNN-BiLSTM model for sentiment classification consists of an embedding layer, followed by a one-dimensional convolutional layer with filters and a kernel size of. The output of the CNN layer is then passed through a max-pooling layer before being fed into a Bidirectional LSTM layer with hidden units. Finally, a dense output layer with a softmax activation function classifies the sentiment into positive, negative, or neutral categories. A dropout rate of was applied after the BiLSTM layer to prevent overfitting.

Table 5. Model Accuracy

Model	Accuracy rate
Sentiment dictionary	0.5888
LSTM model	0.8787
CNN-BiLSTM model	0.9047

Data Analysis

Analysis of the “Former Police Chief in Tangshan assault case sentenced to 12 years” incident (Figures 2 and 3) reveals the sudden and sensitive nature of police-related public opinion. The incident showed rapid propagation, with an initial peak on the first day followed by a more sustained peak on the second. Weibo registered the highest discussion volume, confirming that such topics propagate quickly, attract significant societal attention, and feature a sharp initial volume increase before gradually declining.

Analysis of the discussion shows that most participating netizens have 0 to 10,000 followers (Figure 4). Keyword analysis identified "Tangshan," "assault," "sentenced," and "protective umbrella" as the most frequent terms. Using LDA topic modeling (Figure 5, Table 6), the data was classified into three topics: the "Tangshan assault case protective umbrella sentenced" news (Z1), netizen commentary on the incident (Z2), and the Guangyang District People's Court work report (Z3).

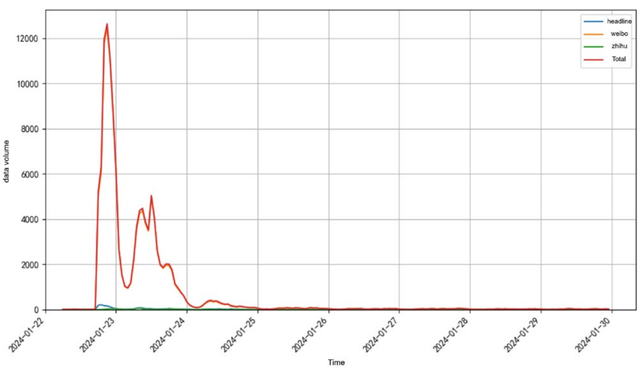

Figure 2. Line Chart of Hourly Public Opinion Data Volume

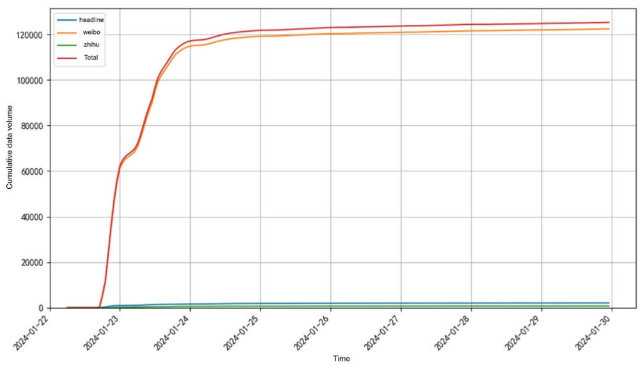

Figure 3. Hourly Cumulative Graph of Public Opinion Data Volume

To objectively identify the core topics within the unstructured text data, we applied the Latent Dirichlet Allocation (LDA) model. A key step in implementing this model is determining the optimal number of topics (K). We addressed this by calculating the Coherence Score for K values ranging from 2 to 9. As illustrated in Figure 6 (Theme Consistency Change Chart), the coherence score peaked at K=3. This indicates that a three-topic division provides the highest intra-topic semantic relevance and the clearest inter-topic distinction. Consequently, we set the final number of topics to three, with the corresponding results detailed in Table 6.

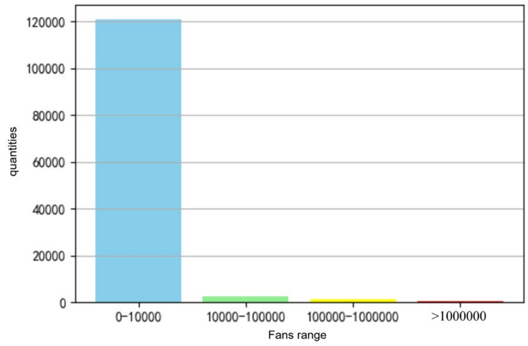

Figure 4. Fan Count Chart

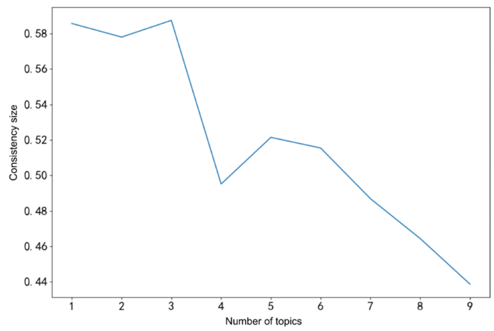

Figure 5. Theme Consistency Change Char

Table 6. Theme and Characteristic Words

Theme	Characteristic Words
Theme Z1	Tangshan, assault, protective umbrella, sentenced
Theme Z2	Black sheep, Evil forces, Justice
Theme Z3	Guangyang District, People's Court, Work, Report

Sentiment analysis of the “Former Police Chief in Tangshan assault case” (Figures 6-8) shows that negative sentiment predominated across all platforms. Sentiment trajectories varied, with Zhihu, Toutiao, and Weibo showing distinct peaks on different days. Notably, on Zhihu and Weibo, sentiment shifted towards neutral on the second day post-incident.

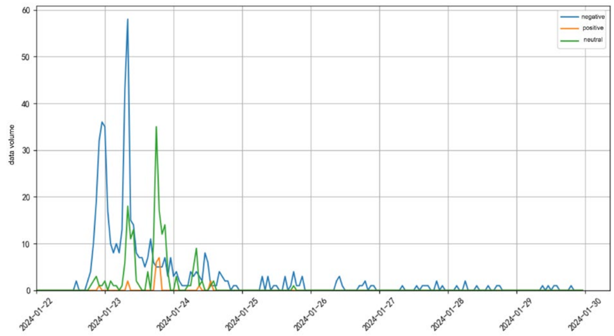

Figure 6. Emotional Changes on Zhihu Platform

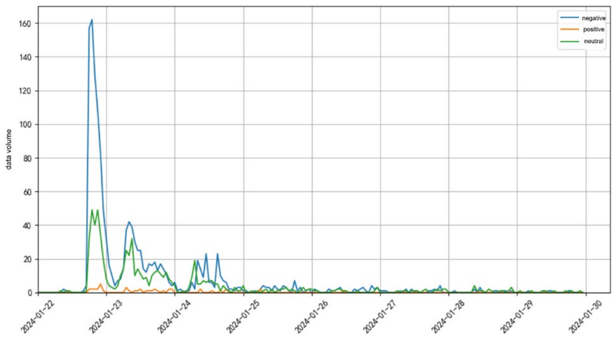

Figure 7. Emotional Changes on Headline Platform

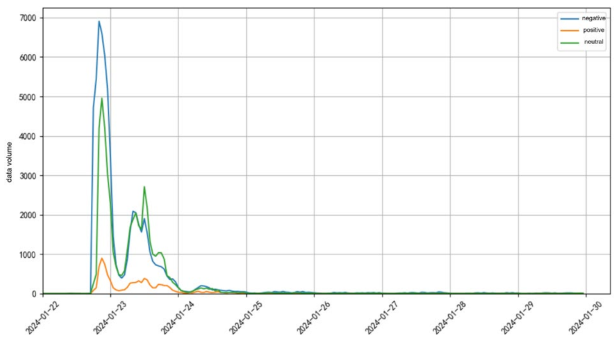

Figure 8. Emotional Changes on Weibo Platform

These cross-platform differences in sentiment evolution can likely be attributed to the distinct characteristics of each platform's ecosystem. For instance, Weibo, as a highly public and event-driven platform, often facilitates rapid emotional contagion and amplification through its 'hot search' mechanism, which may explain the sharp initial peak of negative sentiment. In contrast, Zhihu's community culture encourages more in-depth, rational analysis and multi-perspective discussions. The shift towards neutrality on this platform could be a result of users engaging in more factual verification and reasoned debate after the initial emotional outburst. Meanwhile, the sentiment trajectory on Toutiao might be heavily influenced by its powerful recommendation algorithms, which create personalized information feeds that could potentially reinforce or moderate users' initial emotional stances depending on their reading history. Understanding these platform-specific dynamics is crucial for developing tailored public opinion guidance strategies.

3.3. Setting of Risk Early Warning Levels

Hourly public opinion risk indicators are used with the entropy weight method and TOPSIS to calculate a comprehensive hourly evaluation index for public opinion incidents. Risk levels are then established using a clustering algorithm. This paper employs the objective entropy weight method to determine indicator weights for the TOPSIS framework, avoiding the subjectivity of other methods. These weights and information entropy values, retained to four decimal places, are presented in Table 7.

The specific implementation of the TOPSIS method in this research proceeds as follows: First, we use the objective weights derived from the entropy weight method (presented in Table 6) to create a weighted decision matrix from the hourly data of the 23 quantifiable indicators. Second, based on this matrix, we identify the 'positive ideal solution' (the vector of the best performance values) and the 'negative ideal solution' (the vector of the worst performance values) across all time points. Finally, by calculating the Euclidean distance from each hourly data point to both the positive and negative ideal solutions, we compute a relative closeness index. This index serves as a comprehensive risk score, where a value closer to 1 signifies a higher risk level for that specific hour.

The hourly comprehensive evaluation indices are accumulated, and then the k-means algorithm is applied to cluster the resulting cumulative risk Indices [15]. Specifying the number of clusters as 4, the center points (centroids) for each cluster were obtained as 0.4839, 1.504, 2.997, and 5.395, respectively.This consequently allows for the division of online public opinion risk into four levels, as shown in Table 8.

Table 7. Index Information Entropy and Weights

	S1	S2	S3	S4	S5	S6	W1	W2	W3
Information entropy	0.9712	0.7311	0.7591	0.9896	0.9896	0.9906	0.7487	0.7503	0.7667
Weight	0.0077	0.0716	0.0642	0.0028	0.0028	0.0025	0.0669	0.0665	0.0621
	W4	W5	W6	W7	W8	W9	W10	W11	W12
Information entropy	0.7543	0.7817	0.2545	0.8512	0.7923	0.8681	0.8226	0.8223	0.9997

Weight	0.0654	0.0581	0.0802	0.0396	0.0553	0.0351	0.0473	0.0314	0.0001
	W13	W14	W15	W16	W17	M1	M2	M3	P1
Information entropy	0.7838	0.9997	0.9937	0.929	0.9949	0.7278	0.7448	0.7806	0.9434
Weight	0.0576	0.0001	0.0017	0.0189	0.0014	0.0725	0.068	0.0584	0.0151

Table 8. Risk Level Table

Risk Level	Mild	Moderate	Severe	Dangerous
Cluster center point	0.4839	1.504	2.997	5.395
Comprehensive evaluation index	[0,0.994)	[0.994, 2.251)	[2.251, 4.196)	[4.196,+]

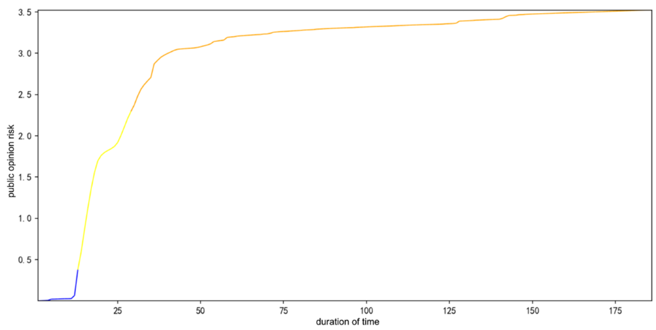

Figure 9. Public Opinion Risk Line Chart

The thresholds established in Table 8 form the core of our risk level determination mechanism. In a practical early warning scenario, the continuous cumulative risk index predicted by our SSA-CNN-LSTM-Attention model is directly compared against these intervals to assign a discrete risk level. For instance, if the model outputs a predicted risk index of 3.5 at a specific hour, this value falls within the [2.251, 4.196) range. Consequently, the system would classify the current risk as 'Severe' and could trigger a corresponding alert. This mapping process translates the quantitative model output into an actionable qualitative warning for decision-makers.

Represent the four risk levels—Mild, Moderate, Severe, and Dangerous—with blue, yellow, orange, and red colors, respectively. The trend of the risk level variation over time for the public opinion incident "Former Police Chief in Tangshan assault case sentenced to 12 years" is illustrated in Figure 9. Among these phases, the mild risk phase exhibits the shortest duration. Towards the latter part of the rapid escalation phase of the public opinion risk, the risk begins to transition into the severe risk phase.

3.4. Establishment and Analysis of Risk Early Warning Models

Research Based on the CNN-LSTM Model

This paper implements two baseline models for risk early warning. The first is an LSTM model, which consists of an input layer, a hidden LSTM layer with 50 units, and a fully connected output layer. The second is a CNN-LSTM model, whose structure is depicted in Figure 10.Specifically, this model processes the 23-dimensional time-series data. The input is first fed into a one-dimensional convolutional layer, which is configured with filters and a kernel size of and uses the ReLU activation function. This layer is designed to extract local temporal features from the input sequence. The output feature map is then passed to a max-pooling layer with a pool size of 2. Subsequently, an LSTM layer with hidden units processes the pooled features to capture long-term temporal dependencies. Finally, a fully connected output layer with a single neuron predicts the continuous cumulative risk index. Both models utilize the Mean Squared Error as the loss function and the Adam optimizer. During training, which for the CNN-LSTM model ran for 100 epochs with a learning rate of 0.001, data batches are processed through forward propagation to compute loss, followed by backpropagation to update model parameters.

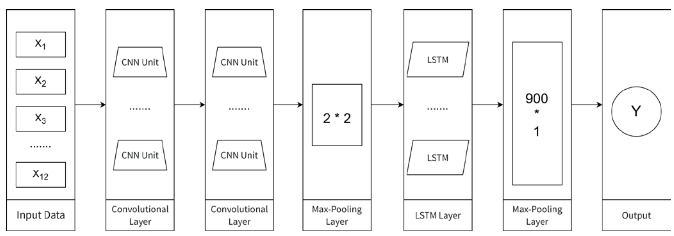

Figure 10. Structure Diagram of CNN-LSTM Model

Construction of the Risk Early Warning Model Based on SSA-CNN-LSTM-Attention

An attention mechanism, consisting of linear layers and a Softmax function, was incorporated into the CNN-LSTM model to calculate and apply attention weights to the CNN component's output. To improve model performance and balance complexity, three algorithms—Bayesian Optimization, Particle Swarm Optimization, and the Sparrow Search Algorithm (SSA)—were used to optimize the neuron count in the attention layer. These algorithms search for the optimal parameter configuration by evaluating the model's loss function. Figure 11 illustrates the final structure of the SSA-optimized CNN-LSTM-Attention model.

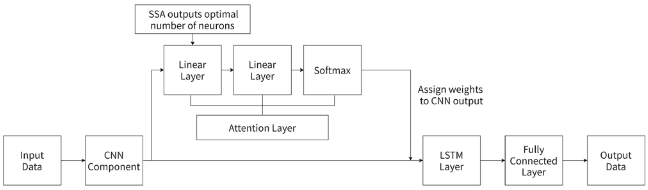

Figure 11. Structural Diagram of SSA-CNN-LSTM-Attention Model

Model Comparison and Analysis

The empirical research utilized data from public opinion incidents, including the "Former Police Chief in Tangshan assault case sentenced to 12 years" and the "Quzhou traffic police post accused of suspected gender discrimination." This data, comprising 23 quantifiable hourly indicators, was processed into standardized 12-step time-series subsequences to predict the subsequent cumulative risk index. For the Sparrow Search Algorithm optimization, key parameters included an attention layer neuron count range of 1 to 500, a population of 20, and 50 iterations. These values were selected based on preliminary tuning tests, which indicated that further increasing the population size or iterations yielded only marginal performance gains at a significantly higher computational cost.

Models were evaluated using three standard metrics. Mean Absolute Percentage Error (MAPE) expresses the average relative error as a percentage; smaller values are better. The formula follows:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| * 100\% \tag{1}$$

Root Mean Square Error (RMSE) reflects the deviation between predicted and actual values; smaller values indicate higher accuracy.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2}$$

Mean Absolute Error (MAE) calculates the average absolute deviation between predicted and actual values; a smaller value indicates a better model. The formula is as follows:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{3}$$

The comparison chart of predicted values for the Tangshan case is shown in Figure 12, and the comprehensive evaluation indicators for each model across all incidents are presented in Table 9.

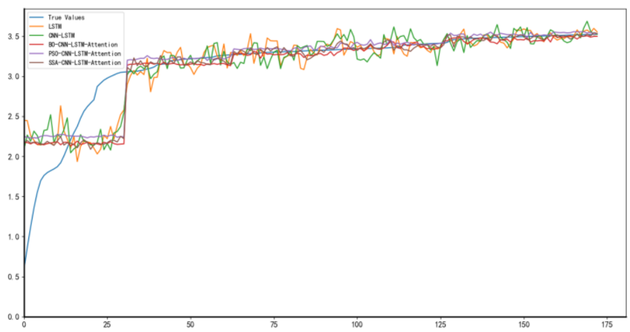

Figure 12. Structural Diagram of SSA

Table 9. Index Information Entropy and Weights

Public opinion incident	Evaluation Indicator	LSTM	CNN-LSTM	BO-CNN-LSTM-Attention	PSO-CNN-LSTM-Attention	SSA-CNN-LSTM-Attention
Elderly woman fatally shot by police after knife attack	RMSE	0.186	0.262	0.243	0.146	0.237
	MAPE(%)	8.832	12.965	9.305	7.425	10.304
	MAE	0.153	0.209	0.176	0.123	0.181
Ningxia officer accused of protecting obscene venue and gang ties	RMSE	0.03	0.031	0.026	0.025	0.023
	MAPE(%)	6.445	6.405	5.05	4.917	4.629
	MAE	0.024	0.024	0.019	0.018	0.016
Shaanxi officer reported beating disabled person at foot spa after drinking	RMSE	0.054	0.057	0.06	0.052	0.048
	MAPE(%)	8.553	9.628	11.156	7.863	6.875
	MAE	0.04	0.043	0.048	0.039	0.035
Auxiliary officer hits elderly person with stool	RMSE	0.147	0.155	0.151	0.149	0.148
	MAPE(%)	36.631	39.998	36.211	35.86	35.532
	MAE	0.095	0.1	0.093	0.091	0.091
Shandong auxiliary officer threatens citizen's life	RMSE	0.039	0.034	0.031	0.028	0.028
	MAPE(%)	44.208	35.922	28.959	26.992	30.184
	MAE	0.019	0.02	0.015	0.014	0.014
Netizen alleges rape by police officer in online post	RMSE	0.205	0.202	0.193	0.211	0.199
	MAPE(%)	9.869	9.771	8.935	9.936	9.408
	MAE	0.137	0.14	0.119	0.136	0.116
Quzhou traffic police post accused of suspected gender discrimination	RMSE	0.246	0.266	0.214	0.214	0.215
	MAPE(%)	139.117	154.077	110.662	110.059	114.47
	MAE	0.143	0.162	0.118	0.122	0.114
Anti-fraud Officer Chen resigns	RMSE	0.115	0.108	0.079	0.08	0.078

	MAPE(%)	12.554	12.126	9.081	9.118	8.816
	MAE	0.085	0.083	0.06	0.062	0.061
Starbucks expels on-duty officers	RMSE	0.228	0.248	0.223	0.1992	0.191
	MAPE(%)	15.996	17.532	14.397	13.533	11.226
	MAE	0.153	0.168	0.179	0.129	0.142
Anhui wife reports police husband for alleged violations	RMSE	0.35	0.282	0.221	0.254	0.24
	MAPE(%)	155.482	103.174	59.466	78.9	68.558
	MAE	0.258	0.207	0.134	0.151	0.163
Binzhou traffic police report flaws in DUI check	RMSE	0.128	0.131	0.116	0.116	0.123
	MAPE(%)	12.705	13.385	10.37	10.452	13.452
	MAE	0.078	0.084	0.064	0.065	0.072
Anhui Wangjiang woman drowning incident	RMSE	0.221	0.278	0.222	0.196	0.192
	MAPE(%)	9.256	12.883	9.068	8.436	8.3
	MAE	0.17	0.228	0.184	0.165	0.156
Shenzhen Bao'an officer domestic violence case	RMSE	0.214	0.209	0.195	0.198	0.218
	MAPE(%)	118.322	116.81	104.5	105.565	116.8
	MAE	0.121	0.115	0.097	0.102	0.121
Fourth Chinese People's Police Day	RMSE	0.359	0.407	0.33	0.328	0.299
	MAPE(%)	5.902	6.842	4.836	4.885	4.166
	MAE	0.238	0.277	0.19	0.194	0.153
Former Police Chief in Tangshan assault case sentenced to 12 years	RMSE	0.311	0.305	0.293	0.291	0.29
	MAPE(%)	8.808	8.74	7.406	7.639	7.345
	MAE	0.169	0.175	0.138	0.138	0.134

The empirical research shows that all optimized models significantly improve prediction accuracy compared to the baseline LSTM model. However, for incidents like the “Former Police Chief in Tangshan assault case,” which are sudden and short-lived, all models exhibited larger errors during the early risk phases, likely due to insufficient early-stage training data.

Compared to the baseline LSTM, the optimization algorithm models markedly reduced errors. While the performance of models optimized by Particle Swarm Optimization and Sparrow Search Algorithm (SSA) was similar, the SSA-optimized model demonstrated superior and more consistent performance across the majority of the 15 analyzed incidents. It achieved the minimum Root Mean Square Error, Mean Absolute Percentage Error, and Mean Absolute Error in 9, 8, and 10 incidents respectively, indicating the most outstanding overall capability.

From a practical application perspective, the proposed model can serve as the core engine for an intelligent early warning system designed for public security organs. In a real-world scenario, this system would manifest as a real-time monitoring dashboard for police-related public opinion incidents. For each ongoing event, the system would automatically process social media data and output a dynamic risk index, which could be visualized using a color-coded alert system. When the predicted risk index surpasses a predefined threshold and enters a high-risk level such as 'Severe', the system would automatically trigger an alert—via email, text message, or an internal platform—to the relevant public relations officers. This would enable law enforcement agencies to move from passive monitoring to proactive management, facilitating timely and targeted intervention to guide public opinion effectively.

4. Conclusion

This paper constructed and empirically validated an optimized fusion model for the risk early warning of police-related online public opinion, proposing targeted response strategies. The SSA-CNN-LSTM-Attention model demonstrated the most effective performance. However, this study has several limitations that should be acknowledged. First and most notably, our empirical results indicate that all models, including our proposed SSA-CNN-LSTM-Attention model, exhibit larger prediction errors during the early stages of sudden and short-lived incidents. This limitation is primarily due to the 'cold-start' problem, where the model has insufficient initial data to effectively capture the rapidly escalating risk dynamics.

To address this, future research could explore advanced methodologies such as transfer learning, which leverages knowledge from well-documented historical events, or few-shot learning techniques that are specifically designed for high performance in data-sparse scenarios. Second, the current indicator system could be further expanded to incorporate more nuanced dimensions based on actual public security operational needs. By addressing these aspects, future work can further enhance the accuracy and robustness of risk early warning systems.

Acknowledgements.s

This work was supported in part by the Hebei Provincial Social Science Fund Project “Big Data Driven Social Security Risk Computing and Intelligent Decision-making for the Beijing-Tianjin-Hebei Region” (Project Number: HB22SH011). We are grateful to the anonymous reviewers for their constructive comments that improved the quality of this work.

References

[1] Liang R, Xu Y. Research on Network Public Opinion Communication Based on Improved SERI Model. Computer Simulation 2023; 40: 333–340.

[2] Sun Y, Liu F, Gao Y, et al. Research on Differentiated Guidance of Network Public Opinion Based on Echo Chamber Super-network. Information Studies: Theory & Application 2022; 45: 138–145.

[3] Xing P, Li X. Research on Formation Mechanism of Online Public Opinions and Guiding Strategy in Major Epidemic Prevention and Control——Qualitative Analysis Based on Online Public Opinion Texts During the Outbreak of COVID-19. Journal of Intelligence 2020; 39: 67–74, 158.

[4] Yang X, Zheng S, Jin Z, et al. Research on Evaluation Index System of Network Public Opinion Hotness Based on Information Ecology Theory. Information Studies: Theory & Application 2021; 44: 143–148.

[5] Tian S, Wang C, Yang W. Research on the Early Warning Mechanism of Network Public Opinion Crisis for Emergencies Based on ANP and Stochastic Petri Net. Chinese Journal of Management Science 2023; 31: 215–224.

[6] Yang L, Luo W, Deng C, et al. Classification and Early Warning Model of Public Opinion Based on Grey Correlation Analysis. Information Science 2020; 38: 28–34.

[7] Luo W, Ma X. An Empirical Study of Multi-Stage Network Public Opinion Early Warning Based on Bayes-ian Network. Information Science 2021; 39: 68–74.

[8] Huang X, Jin H, Zhang Y. Risk assessment of earthquake network public opinion based on global search BP neural network. PLoS ONE 2019; 14: e0212839.

[9] Shen G. The Early-warning of Police-related Network Public Opinions based on TSFNN and Ensemble Learning. Journal of Heze University 2022; 44: 10–14, 20.

[10] Zander KK, Garnett ST, Ogie R, et al. Trends in bushfire related tweets during the Australian ‘Black Summer’ of 2019/20. Forest Ecology and Management 2023; 545: 121274.

[11] Esparza D, Ugues A. The Impact of Law Enforcement Centralisation and Professionalisation on Public Opinion of the Mexican Police. Journal of Politics in Latin America 2020; 12: 104–120.

[12] Deng J, Wu L, Qi J, et al. Risk Classification and Prediction of Negative Network Public Opinion Based on Information Correlation. Information Science 2022; 40: 38–43.

[13] Cao H, Hou T. Research on the Structural Relationship of Influencing Factors of Political Public Opinion Risk Based on ISM. Hubei Social Sciences 2021; 43–51.

[14] Zhang J, Zhu L, Ding W. Research on the Impact Factors, Risk Assessment and Coping Strategy of Police-involved Network Public Opinion. Journal of Zhejiang Police College 2021; 59–64.

[15] Yang L, Xu Y, Deng C. Risk Assessment and Early Warning of University Network Public Opinion. Information Science 2022; 40: 65–72, 83.