# Multi-KGS: Generating Social Network-based Meteorological Decision Reports Fusing with Multiple Knowledge

Kaize Shi, *Member, IEEE,* Xueping Peng, *Senior Member, IEEE,* Yifan Zhu, *Member, IEEE,*
Hui He, Kun Yi, Zhendong Niu

***Abstract*—The increasing prevalence of meteorological disasters necessitates advanced spatiotemporal data analytics to enhance emergency response in smart cities. Social networks, as real-time crowdsourcing sensors, provide critical data streams that Generative AI (GenAI) can fuse and summarize to generate comprehensive meteorological decision reports for enhanced emergency response during abrupt weather crises. This paper introduces a Multiple Knowledge Guided Summarization (Multi-KGS) model designed to generate meteorological decision reports by fusing posts from Sina Weibo. Specifically, the Multi-KGS model comprises a summary generation module and a multiple knowledge guidance module. The summary generation module synthesizes the content of the decision report, while the multiple knowledge guidance module steers and constrains the summarization process using knowledge of meteorological events and geographical locations, ensuring that the generated report highlights the core knowledge expressed in the source posts. Compared to baseline models, Multi-KGS achieves superior performance in content evaluation, as measured by $ROUGE-1$, $ROUGE-2$, and $ROUGE-L$, as well as in sentiment evaluation, with the best $F_1$ score. This study provide a generative decision support paradigm for servicing urban computing.**

***Index Terms*—Smart Cities, Social Sensors, Generative Decision Support, GenAI-based Emergency Management, Urban Computing.**

## I. Introduction

The increasing frequency of meteorological disasters due to global warming has led to substantial societal and economic losses worldwide [1]. Integrating Generative Artificial Intelligence (GenAI) into meteorological emergency management is essential for enhancing decision support services in smart cities [2], [3]. Decision reports serve as a crucial mechanism in this domain, facilitating the fusion of critical insights from vast amounts of unstructured data [4], [5]. Social networks, characterized by user-generated content (UGC), function as real-time social sensors that provide timely and intuitive feedback during extreme meteorological events [6]. The automatic generation of meteorological decision reports from social network data is thus pivotal for improving decision-making efficiency and mitigating disaster impacts.

K. Shi is with the School of Mathematics, Physics and Computing, University of Southern Queensland, QLD 4350, Australia. (E-mail: Kaize.Shi@unisq.edu.au)

X. Peng is with the Australian Artificial Intelligence Institute, University of Technology Sydney, NSW 2007, Australia.

Y. Zhu is with the School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China.

H. He is with the School of Medical Technology, Beijing Institute of Technology, Beijing 100081, China.

K. Yi is with the State Information Center of China, Beijing 100045, China.

Z. Niu is with the School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China.

The distributed nature of social networks makes them indispensable sensors for fusing information from the socialized perspectives [7]. Users actively contribute social signals reflecting their real-time perceptions and experiences of meteorological events, offering a valuable data source for decision support [8]. For instance, during the catastrophic Henan floods in July 2021 [9], Sina Weibo emerged as a primary communication channel for affected individuals, enabling them to report flood conditions, request assistance, and express their sentiments. These social signals played a crucial role in mobilizing official and civilian rescue efforts, underscoring the potential of social media in meteorological disaster management. Effectively integrating such signals into urban computing is imperative for enhancing government decision-making and improving public services.

Conventional meteorological decision report generation primarily relies on single-document summarization models [4]. These models exhibit clear limitations in integrating diverse perspectives and sentiment expressions from multiple sources. Recent advancements in large language models (LLMs) have revolutionized the generation paradigm [10], [11], their practical deployment still presents challenges. Specifically, the high computational costs of LLMs training and inference require users to rely on third-party API services, which can introduce risks related to data security in public services [12]. Furthermore, LLMs may generate inaccurate or incomplete descriptions of key meteorological events due to a lack of domain-specific knowledge constraints, potentially affecting the effectiveness of decision reports [13], [14].

To address the aforementioned limitation, this paper proposes the **Multi**ple **K**nowledge **G**uided **S**ummarization (Multi-KGS) model. The Multi-KGS model consists of a summary generation module and a multiple knowledge guidance module. The summary generation module is in an unsupervised structure to summarize the core knowledge in the source text. The multiple knowledge guidance module constrains the generation process by incorporating critical knowledge, including 14 meteorological event knowledge[1] and 367 geographical location knowledge[2], to produce comprehensive and localized decision reports. Considering the ubiquitous independent and co-occurrence characteristics of meteorological events and the wide-area impact of the same meteorological event on similar locations, the multiple knowledge guidance module is based on the multi-label classification structure to realize the correlation mining by taking into account the independent and co-occurrence characteristics of the knowledge carrier words. To introduce prior meteorological knowledge, we propose the MET-BERT model to initialize the word embedding of the Multi-KGS model. For the automated pipeline, this paper presents a framework for automatically generating meteorological decision reports, as shown in Fig. 1. The framework includes a data collection module that gathers meteorological posts from Sina Weibo, a data processing module that constructs the corpus and extracts relevant knowledge, a Multi-KGS model for generating the report content, and a report overview module that displays and delivers the report to decision-makers. The framework supports rapid meteorological decision-making in disaster scenarios.

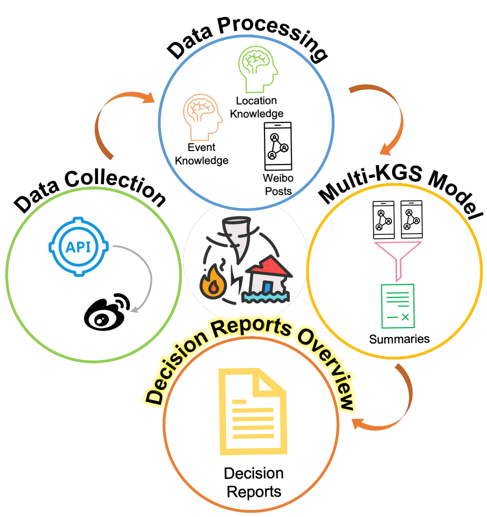


Fig. 1: The framework for fusing multiple information and generating the meteorological decision reports.

The key contributions of this paper are as follows:

- We propose a Multiple Knowledge Guided Summarization (Multi-KGS) model to generate meteorological decision report content fused with meteorological event and geographical location knowledge.
- We introduce MET-BERT, a domain-specific language model fine-tuned on meteorological corpora to enrich the Multi-KGS model with specialized meteorological knowledge.
- We validate our approach through extensive experiments, demonstrating that the proposed Multi-KGS model outperforms all baseline models in both quantitative metrics and qualitative evaluations.

[1]The extreme meteorological events include Hail, Sandstorm, Blizzard, Fog, Drought, Cold Wave, Typhoon, Thunder, Icing, High Temperature, Gale, Frost, Rainstorm, and Haze. These events are defined by the China Meteorological Administration (CMA) and are also recognized as a national standard by the General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China and the Standardisation Administration of the People's Republic of China (2011).

[2]Geographical location knowledge is defined as the 367 administrative regions in China at the prefecture level and above, including 333 prefecture-level administrative regions and 34 provincial-level administrative regions. These classification criteria are defined by the Administrative Divisions of the People's Republic of China.

## II. Related Work

The decision report is an effective way for urban computing [15]. Most of the current decision report construction methods are based on fixed templates [16], [17]. Reiter et al. [17] proposed SUMTIME-MOUSAM, which groups the words in weather forecasts into fixed categories and generates grammatical weather forecasts based on numerical data. The result of manual verification shows that users prefer the weather forecast text generated by the model instead of manually written since the model has better word choices. With the development of GenAI, the generative decision support services have gradually become mainstream [4], [5].

To improve the factual accuracy and adaptability of such generative services, recent studies have introduced controllable generation mechanisms [18], where external prior knowledge plays an essential role. Narayan et al. [19] proposed a summary generation method based on the ordered sequence of entities to plan the summary generation process by taking entity sequences and source texts as inputs. Chen et al. [20] proposed a relation-aware multi-document summarization model, which uses the dependencies in the relationship graph to establish associations between documents and generates multi-documents-based literature review in a specific domain. The prior knowledge introduced by the keywords always contains specific attributes, such as topics and emotions [21], [22]. To this end, Zhu et al. [21] classified the topic from Wikipedia and generated the topic-centered summary from the sentence's topic-aware representations. Focusing on the explicit document semantics, Wang et al. [22] integrated the topic models into the summary generation process by a topic assistant, which improves the performance of the transformer-based summarization models. The prior knowledge has also been introduced by the pre-trained language model, where the BERT [23] model is commonly used. Ma et al. [24] proposed the T-BERTSum, a summarization model sensitive to a specific topic. Du et al. [25] proposed the BioBERTSum model to introduce the biomedical domain's prior knowledge for extractive summarization.

## III. Methods

### A. Method Overview

The proposed multiple knowledge guided summarization (Multi-KGS) model is illustrated in Fig. 2. The Multi-KGS model consists of two components: a summary generation module and a multiple knowledge guidance module. The former produces decision report content by processing a collection of Weibo posts $X^{1:K} \in \mathbb{R}^K$, where $K$ represents the total number of Weibo entries to be synthesized. Meanwhile, the latter component enhances and regulates the generation process by incorporating meteorological event knowledge $E_g$ and geographical location knowledge $L_g$ as guiding constraints.

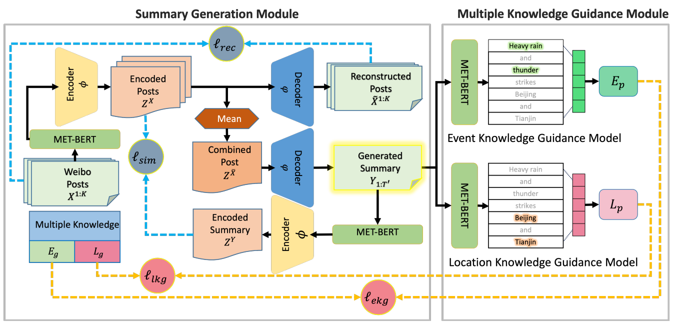


Fig. 2: The structure of the multiple knowledge guided summarization model.

To make the Multi-KGS model more sensitive to the semantic features of the meteorological domain, we introduce prior meteorological knowledge through the MET-BERT model. The MET-BERT model is fine-tuned by the meteorological event classification task with the meteorological Weibo posts based on the general "BERT-Base, Chinese" [23][3] model. The word embedding of the Multi-KGS model is initialized by the MET-BERT model and is fixed during the training process.

The summary generation module is based on an unsupervised auto-encoder structure. The source Weibo posts $X^{1:K}$ are embedded into word vectors and encoded into $\mathbf{Z}^X \in \mathbb{R}^{2\times K\times dim^\phi}$ by the encoder $\phi$. The mean function combines $\mathbf{Z}^X$ to $\mathbf{Z}^{\bar{X}} \in \mathbb{R}^{2\times dim^\phi}$ that contains the mean semantic features of $X^{1:K}$. The decoder $\varphi$ reconstructs the $\mathbf{Z}^X$ to $\widetilde{X}^{1:K}$, and calculates the reconstruction loss $\ell_{rec}$ between $X^{1:K}$ and $\widetilde{X}^{1:K}$, while $\varphi$ also generates the summary $Y_{1:T'} \in \mathbb{R}^{T'}$ that contains the comprehensive semantic features of $X^{1:K}$, where $T'$ is the length of the generated sequence. The $Y_{1:T'}$ is encoded into $\mathbf{Z}^Y \in \mathbb{R}^{2\times T' \times dim^\phi}$ by $\phi$, and the semantic similarity is constrained by optimizing the similarity loss $\ell_{sim}$ between each $\mathbf{Z}^k \in \mathbf{Z}^X$ and $\mathbf{Z}^Y$.

The multiple knowledge guidance module is a convolutional neural network (CNN)-based structure, which predicts both the meteorological event knowledge $E_p$ and the geographical location knowledge $L_p$ contained in the generated summary $Y_{1:T'}$. The knowledge guidance process is achieved by optimizing the loss function of $\ell_{ekg}$ and $\ell_{lkg}$, which minimize the variance of the $E_p$ and $L_p$ with the corresponding ground-truth knowledge $E_g$ and $L_g$. In the rest of this section, we describe each module in detail.

### B. Summary Generation Module

The summary generation module incorporates two interconnected sub-models: a reconstruction model and a similarity alignment model. Both employ an auto-encoder architecture wherein the encoder and decoder components are implemented as two-layer unidirectional LSTM networks. To enhance semantic coherence, we establish parameter sharing between the encoder-decoder weights across both sub-models. This weight-tying mechanism ensures that the generated summary $Y_{1:T'}$ maximizes semantic alignment with each individual Weibo post $X^k_{1:T} = \{x^k_1, x^k_2, ..., x^k_T\} \in \mathbb{R}^T$, where $T$ denotes the sequence length of the $k^{th}$ Weibo post.

*1) Reconstruction Model:* The reconstruction model's encoder, denoted as $\phi$, is a two-layer unidirectional LSTM that processes $X^k_{1:T}$ into $\mathbf{Z}^k = \phi\left(X^k_{1:T}\right) = \left[\mathbf{h}^k_T, \mathbf{c}^k_T\right] \in \mathbb{R}^{2\times dim^\phi}$. Here $\mathbf{h}^k_T$ and $\mathbf{c}^k_T$ correspond to the hidden and cell states at time step $T$, respectively, while $dim^\phi$ denotes the dimensionality of the LSTM states.

The decoder $\varphi$ is constructed with a two-layer unidirectional LSTM, which reconstructs $\mathbf{Z}^k$ into $\widetilde{X}^k_{1:T''} = \left\{\widetilde{x}^k_1, \widetilde{x}^k_2, ..., \widetilde{x}^k_{T''}\right\} = \varphi\left(\mathbf{Z}^k\right) \in \mathbb{R}^{T''}$. Following Chu et al [26], we use cross-entropy loss with teacher-forcing mechanism [27] to train the reconstruction model as Eq. 1:

$$\ell_{rec} = -\sum_{k=1}^{K} log\left(p\left(X^k_{1:T} \mid \varphi\left(\phi\left(X^k_{1:T}\right)\right)\right)\right). \tag{1}$$

We calculate the mean semantic features $\mathbf{Z}^{\bar{X}}$ as Eq. 2, where $\mathbf{Z}^{\bar{X}}$ contains each of the semantic features of $X^k_{1:T}$ and tends to retain the common features in $X^{1:K}$.

$$\mathbf{Z}^{\bar{X}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{Z}^k. \tag{2}$$

*2) Similarity Alignment Model:* The summary $Y_{1:T'} = \{y_1, y_2, ..., y_{T'}\} \in \mathbb{R}^{T'}$ is generated by the similarity alignment model's decoder $\varphi$ with the input of $\mathbf{Z}^{\bar{X}}$, $Y_{1:T'} = \varphi\left(\mathbf{Z}^{\bar{X}}\right)$.

To constrain the semantic similarity between $Y_{1:T'}$ and $X^k_{1:T}$, we re-encode the $Y_{1:T'}$ into $\mathbf{Z}^Y$, $\mathbf{Z}^Y = \phi(Y_{1:T'}) = \left[\mathbf{h}^Y_{T'}, \mathbf{c}^Y_{T'}\right] \in \mathbb{R}^{2\times dim^\phi}$, and calculate the cosine similarity $cos\ (\cdot)$ between $\mathbf{h}^Y_{T'}$ of $\mathbf{Z}^Y$ and $\mathbf{h}^k_T$ of $\mathbf{Z}^k$ as the loss function, which is shown in Eq. 3. Since each $y_t \in Y_{1:T'}$ are discrete and the model without the ground-truth summary corresponding to $Y_{1:T'}$, here we use the straight through gumbel-softmax [28] to optimize the $\ell_{sim}$.

$$\ell_{sim} = \frac{1}{K}\sum_{k=1}^{K} cos(\mathbf{h}^k_T, \mathbf{h}^Y_{T'}). \tag{3}$$

Eq. 4 specifies the loss function for the summary generation module. There is no hyper-parameter to balance the $\mathcal{L}_{SUMM}$ because each sub-model's encoder and decoder are weight-tied.

$$\mathcal{L}_{SUMM} = \ell_{rec} + \ell_{sim}. \tag{4}$$

[3] https://storage.googleapis.com/bert_models/2018_11_03/chinese_L-12_H-768_A-12.zip

### C. Multiple Knowledge Guidance Module

The multiple knowledge guidance module consists of an event knowledge guidance model and a location knowledge guidance model based on the CNN structure. To capture the distinct features of individual meteorological events or geographical location knowledge, we employ a convolution kernel with a size of 1. This design allows the model to effectively extract co-occurrence relationships between meteorological events and geographical knowledge while highlighting the unique contributions of these independent features in guiding the summary generation.

Let $\mathbf{V}_{1:T'} \in \mathbb{R}^{T' \times dim^{KG}}$ represent the vectorized form of $Y_{1:T'}$, embedded with the MET-BERT model that is shared within the summary generation module. The $dim^{KG}$ denotes the dimensionality of word embeddings in multiple knowledge guidance module. Additionally, let $\mathbf{m} \in \mathbb{R}^{T'-f+1}$ be the feature map obtained by convolution operation $*$, where $f \in [1, 3, 4, 5]$ defines the receptive field size. Each feature $\mathbf{m}_i \in \mathbf{m}$ can be calculated as:

$$\begin{aligned} \mathbf{m}_i &= R(\mathbf{w}^f * \mathbf{V}_{1:T'})_i \\ &= R(\mathbf{w}^f \cdot \mathbf{V}_{1:T'[i-f+1:i]}) \\ &= R(\sum_{n=i}^{i+f-1} \mathbf{w}_n^f \mathbf{V}_{1:T'n}), \end{aligned} \quad (5)$$

where $R(\cdot)$ represents the non-linear activation function ReLU [29], and $\mathbf{w}^f$ is the convolution kernel with size of $f$. $\hat{\mathbf{m}}$ represents the max-overtime pooling of $\mathbf{m}$, which captures the most significant features. Let $K_p$ be the predicted knowledge, $K_p \in \{E_p, L_p\}$, where $E_p \in T^e$ and $L_p \in T^l$. The loss function of each multiple knowledge guidance module can be calculated as follows:

$$\ell_{EKG/LKG} = -\frac{1}{N}\sum_{n=1}^{N} log(p({K_g}^{(n)} \mid {Y_{1:T'}}^{(n)}, \theta^{KG})), \quad (6)$$

where $K_g \in \{E_g, L_g\}$ represents the ground-truth knowledge corresponding to $K_p$, and $\theta^{KG}$ represents the parameters in the multiple knowledge guidance module. The learning process encourages the multiple knowledge guidance module to maximize the probability of the $K_p$ close to the $K_g$ in $Y_{1:T'}$, thus encouraging the summary generation module to generate the summary sequence containing the corresponding knowledgeable words.

The loss function of the entire multiple knowledge guidance module is as follows:

$$\mathcal{L}_{KG} = \lambda_{EKG} \cdot \ell_{EKG} + \lambda_{LKG} \cdot \ell_{LKG}. \quad (7)$$

Since the event and location knowledge guidance models are independent of one another, with each focusing on distinct features, here we introduce the hyper-parameters $\lambda_{EKG}$ and $\lambda_{LKG}$ to balance the overall loss function of the multiple knowledge guidance module to prevent $\mathcal{L}_{KG}$ dominated by a single task due to the different model scale. The hyper-parameters also encourage the multiple knowledge guidance module to produce the most positive impact on the summary generation module.

### D. Training

The overall loss function of the Multi-KGS model ($\mathcal{L}_{MKGS}$) is as Eq.8, consisting the loss function of summary generation module ($\mathcal{L}_{SUMM}$) and multiple knowledge guidance module ($\mathcal{L}_{KG}$).

$$\begin{aligned} \mathcal{L}_{MKGS} &= \mathcal{L}_{SUMM} + \mathcal{L}_{KG} \\ &= \ell_{rec} + \ell_{sim} + \lambda_{EKG} \cdot \ell_{EKG} + \lambda_{LKG} \cdot \ell_{LKG}. \end{aligned} \quad (8)$$

By continuously minimizing $\mathcal{L}_{MKGS}$, the training process encourages the summary generation module to produce $y_t \in Y_{1:T'}$, which is strongly correlated with $E_g$ and $L_g$. This structure ensures that $Y_{1:T'}$ retains as much meteorological event knowledge and geographical location knowledge from $X^k_{1:T}$ as possible, learning $Y_{1:T'}$ by using $E_g$ and $L_g$ as the core descriptive elements.

## IV. Experiments

### A. Dataset

The dataset is derived from posts on Sina Weibo[4], a Chinese leading social media platform. The Jieba[5] is utilized to segment the posts in natural language into word sequences. In data preprocessing, high-quality Weibo posts are selected using regular expression-based rules. Posts containing empty content or characters such as "Emoji" and "\u3000" are filtered out. Since post length is a key factor influencing model performance, we employ boxplots and histograms to filter posts of appropriate length. As shown in Fig. 3, posts with lengths ranging from 2 to 474 ($T \in [2, 474]$) are retained. We also exclude posts that lack event or location knowledge. For posts' location knowledge outside of the prefecture's administrative regions, the corresponding provincial administrative regions are used to represent the location. After data preprocessing, a total of 722,349 Weibo posts remained in the dataset. Examples of the dataset are in Table I.

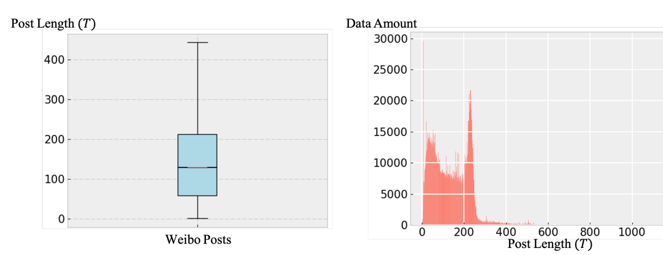


Fig. 3: The box plot and histogram illustrate the distribution of Weibo post lengths in the dataset. Data points outside the whiskers of the box plot, which represent outliers, are not displayed.

[4]https://www.weibo.com/
[5]https://github.com/fxsjy/jieba

TABLE I: Example of a dataset with event and location knowledge in bold.

| Weibo Posts | Event Knowledge | Location Knowledge |
|---|---|---|
| 初秋偶遇，不巧京城又是两天**雾霾**天气，**北京**。We met by chance in early autumn, but unfortunately it was another two days of **haze** weather in **Beijing**. | 霾 Haze | 北京 Beijing |
| **雷电**预警。**邯郸**市气象台 23 时 55 分发布**雷电**黄色预警信号，预计今天夜间我市将受**雷电**天气影响，请注意防范。**Thunder** warning. The **Handan** Meteorological Observatory issued a yellow **thunder** warning signal at 23:55. It is expected that our city will be affected by **thunder** tonight, please pay attention to precautions. | 雷电 Thunder | 邯郸 Handan |
| **上海**突降**暴雨**，半分钟积水成河。**Heavy rain** fell suddenly in **Shanghai**, and the water became a river in half a minute. | 暴雨 Heavy rain | 上海 Shanghai |

### B. Experiments Settings

The experiments are conducted on an NVIDIA RTX 3090 GPU. Adam [30] is employed as the optimizer for the Multi-KGS model. The Weibo post set consists of 8 posts ($K = 8$), and the learning rate is set to 0.0001. The word embedding dimensions in both the summary generation module and the knowledge guidance module are 768, while the hidden state dimensions in the encoder ($\phi$) and decoder ($\varphi$) are also 768 ($dim^{\phi} = 768$, $dim^{\varphi} = 768$). The number of meteorological event knowledge instances is 14 ($T^e = 14$), and the number of geographical location knowledge instances is 367 ($T^l = 367$).

The hyper-parameters, $\lambda_{EKG}$ and $\lambda_{LKG}$, which are used to balance the overall loss function $\mathcal{L}_{MKGS}$ are 0.1 and 0.9 ($\lambda_{EKG} = 0.1$, $\lambda_{LKG} = 0.9$). In the following sections, we will discuss the influence of the different hyper-parameters on the overall model performance.

## V. Results and Analysis

### A. Evaluation Metric

*1) Content Evaluation:* We use *ROUGE* as the quantitative evaluation metric, which is a set of metrics to evaluate the similarity between the source Weibo posts ($X^{1:K}$) and the generated summary ($Y_{1:T'}$).

*2) Sentiment Evaluation:* Most social network posts contain significant sentiment features carried by sentiment words, crucial for assessing public opinion and attitudes toward meteorological events. Sentiment consistency proves that the generated summary and the source posts maintain similar sentiment words, further proving the semantic consistency between these two texts.

Specifically, we label the sentiment attributes of each $X^k_{1:T} \in X^{1:K}$ by the Senta [31], where the sentiment value $\theta^{X^k_{1:T}}_{sent} = 1$ represents $X^k_{1:T}$ with positive sentiment attribute and $\theta^{X^k_{1:T}}_{sent} = 0$ represents $X^k_{1:T}$ with negative sentiment attribute. After that, we train a Text-CNN-based sentiment classification model by using $X^k_{1:T}$ as input, which achieves fine-grained sentiment classification fitting the semantic features of the meteorological domain. We calculate the *Accuracy* (*Acc*), *Precision* ($P$), *Recall* ($R$), and $F_1$ between the predicted sentiment attribute $\theta^{X^k_{1:T}}_{sent_p}$ and the ground-truth sentiment attribute $\theta^{X^k_{1:T}}_{sent_g}$. Table II presents the evaluation results of the sentiment classification model, showing its ability in evaluating the sentiment polarity preservation between the generated summary and the original posts.

TABLE II: The testing results of the model for the quantitative sentiment evaluations.

| Metric | Value |
|---|---|
| *Accuracy* | 0.983 |
| *Precision* | 0.958 |
| *Recall* | 0.976 |
| $F_1$ | 0.966 |

We label the ground-truth sentiment attribute ($\theta^{Y_{1:T'}}_{sent_g}$) of the generated summary as Eq. 9, and we use the trained sentiment classification model to evaluate whether the sentiment polarity of the generated summary remains the same probability distribution with $X^{1:K}$.

$$\theta^{Y_{1:T'}}_{sent_g} = \begin{cases} 1, \frac{1}{K}\sum_{k=1}^{K}\theta^{X^k_{1:T}}_{sent} \geq 0.5 \\ 0, \frac{1}{K}\sum_{k=1}^{K}\theta^{X^k_{1:T}}_{sent} < 0.5 \end{cases}. \quad (9)$$

### B. Baseline Methods

We take the unsupervised text summarization models as the baseline methods: the Extractive summarization model [32], MeanSum [26], and Copycat [33]. The Extractive model [32] is a centroid-based summarization model that exploits word embeddings' compositional capabilities. The MeanSum [26] is the model that generates the multi-document summary by considering the mean semantic features among input texts. The Copycat [33] focuses on preserving the specific information of each text in the generated summary.

### C. Quantitative Evaluation

*1) Content Evaluation:* The quantitative content evaluation results for the proposed Multi-KGS model, along with other baseline models, are presented in Table III. The Multi-KGS model outperforms the baseline models in terms of $ROUGE-1$, $ROUGE-2$, and $ROUGE-L$, demonstrating its effectiveness in summarizing the essential knowledge conveyed by the source Weibo posts.

TABLE III: Quantitative content evaluation results of the Multi-KGS model and other baseline models.

| | ROUGE-1 | ROUGE-2 | ROUGE-L |
|---|---|---|---|
| Extractive [32] | 0.1661 | 0.0564 | 0.1472 |
| MeanSum [26] | 0.1869 | 0.0589 | 0.1640 |
| Copycat [33] | 0.1874 | 0.0632 | 0.1587 |
| **Multi-KGS** | **0.2068** | **0.0713** | **0.1713** |

*2) Sentiment Evaluation:* We use quantitative sentiment evaluation to assess the similarity in sentiment polarity between the generated summaries and the source Weibo posts. The results, shown in Table IV, indicate that the proposed Multi-KGS model achieves the best performance in terms of $Acc$, $P$, $R$, and $F_1$. This demonstrates that the Multi-KGS model is more effective at preserving the sentiment polarity's probability distribution.

TABLE IV: Quantitative sentiment evaluation results of the Multi-KGS model and other baseline models.

| | $Acc$ | $P$ | $R$ | $F_1$ |
|---|---|---|---|---|
| Extractive [32] | 0.779 | 0.438 | 0.500 | 0.467 |
| MeanSum [26] | 0.862 | 0.633 | 0.625 | 0.619 |
| Copycat [33] | 0.824 | 0.602 | 0.617 | 0.610 |
| **Multi-KGS** | **0.868** | **0.649** | **0.655** | **0.652** |

## D. Qualitative Evaluation

This section presents a qualitative evaluation using the summaries generated by the Multi-KGS model and other baseline models. The results are shown in Table V.

The core content expressed in the source posts can be divided into five parts: (1) the heavy rain in Chongqing; (2) the lightning in Longchuan; (3) Fujian issued a yellow warning signal for lightning; (4) the gale in Beijing; (5) the gale in Tianjin. The source posts have three meteorological event knowledge (heavy rain, lighting, and gale) and five location knowledge (Chongqing, Longchuan, Fujian, Beijing, and Tianjin). The Extractive [32] generated summary described the one of core content ($\frac{1}{5}$), meteorological event knowledge ($\frac{1}{3}$), and geographical location knowledge ($\frac{1}{5}$) in the source posts. Besides, the model mistakenly generated the time describing words "this morning", which would mislead the decision-making process. The summary generated by the MeanSum [26] model described two of core content ($\frac{2}{5}$), meteorological event knowledge ($\frac{2}{3}$), and geographical location knowledge ($\frac{2}{5}$) in the source posts, but it mistakenly generated another meteorological event knowledge: high temperature, which interference the decision process. The Copycat [33] model generated three of core content ($\frac{3}{5}$), two of meteorological event knowledge ($\frac{2}{3}$), and geographical location knowledge ($\frac{2}{5}$) expressed in the source posts. However, it mistakenly maps the lighting event to the location of Chongqing. We speculate that this is due to the frequent co-occurrence of heavy rain and lightning events, causing the model to incorrectly associate the events based on inter-word probability. The summary generated by the Multi-KGS model described fore of core content ($\frac{4}{5}$), three of meteorological event knowledge ($\frac{3}{3}$), and fore of geographical location knowledge ($\frac{4}{5}$) expressed in the source posts. In contrast, the Multi-KGS model is more sensitive to the specific meteorological event and geographical location knowledge, giving the generated summaries more efficient decision support service capabilities. More importantly, the summary generated by the Multi-KGS model accurately retains the corresponding relationship between meteorological event knowledge and geographical location knowledge, which is beneficial for decision-makers to better understand the content of decision reports and formulate more fine-grained decision-support strategies.

## E. Receptive Field Configuration Analysis

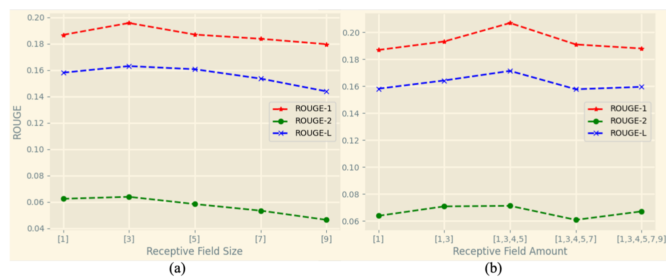


Fig. 4: Effect of (a) receptive field size and (b) receptive field amount in the multiple knowledge guidance module.

We analyze two key parameters in the multiple knowledge guidance module through Fig. 4. First, subfigure (a) reveals that a receptive field size of 3 achieves optimal $ROUGE$ scores, as single-grained fields struggle to capture co-occurrence features across meteorological events and geographical locations. Notably, combined sizes of 1, 3, and 5 show synergistic benefits. Second, (b) demonstrates that the configuration yields peak performance by capturing critical spatiotemporal patterns. These findings collectively validate our multi-scale design choice for feature extraction in meteorological decision reports.

# VI. Ablation Study

## A. Multiple Knowledge Guidance Module

The ablation study results for knowledge guidance models are in Table VI and Table VII, which are the content and sentiment evaluation. The results reveal that the Multi-KGS model, enhanced by the event and location knowledge guidance modules, achieves superior content and sentiment evaluation scores. These findings further highlight their positive impact on overall performance.

## B. $\lambda$ Selection

We show the influence of different $\lambda_{EKG}$ and $\lambda_{LKG}$ on $ROUGE-1$, $ROUGE-2$, and $ROUGE-L$ in Fig. 5. All

TABLE V: Qualitative evaluation results of the Multi-KGS model and other baseline models.

| | |
|---|---|
| Source | $X^1$: **重庆**在下**暴雨**，我服了。这裙子真好看，但是拍不出来。It is **heavy rain** in **Chongqing**, and I took it. This dress is so pretty, but I can't photograph it.<br>$X^2$: **龙川**天气。提醒目前有雷雨云团逐渐东移，影响我县。预计未来我县局部地区将出现雷雨天气并伴有短时强降水、短时大风、**雷电**等强对流天气。请注意防御。龙川县气象局 19 日 17 时 55 分发布。**Longchuan** weather. Remind that there are thunderstorm clouds gradually moving eastward, affecting our county. It is expected that there will be thunderstorms in some parts of the county in the future, accompanied by strong convective weather such as short-term heavy precipitation, short-term gale, and **thunder**. Beware of defense. The Longchuan County Meteorological Bureau released it at 17:55 on the 19th.<br>$X^3$: **福建**省气象台 5 月 19 日 17 时 15 分继续发布**雷电**黄色预警信号。The **Fujian** Meteorological Observatory continued to issue a yellow **thunder** warning signal at 17:15 on May 19.<br>$X^4$: **北京大风**。吹到头痛。**Beijing gale**, blowing headache.<br>$X^5$: 今天**北京**迎来了一场**大风**，明天会是一个好天气。社会人第一天希望一切顺顺利利。There is a **gale** in **Beijing** today, and tomorrow will be a good weather. The first day of society people hope everything goes well.<br>$X^6$: **北京大风**。我早上睡醒就在挣扎要不要出门买粮（家里没有吃的只有点小米了），然后一天了到现在都没敢出去，今天就吃了点小米粥啊。**Beijing gale**. When I woke up in the morning, I was struggling whether to go out to buy food (the only thing I didn't eat at home was some millet), and then I didn't dare to go out for a day, so I ate some millet porridge today.<br>$X^7$: **北京大风**。这么大风适合放风筝。**Beijing gale**. Such gale are suitable for kite flying.<br>$X^8$: 今天在**大风**中凌乱的你可以松口气了，高温才是下周天气的关键词啊，更多天气信息请关注**天津**气象。Today in the **gale** messy you can breathe a sigh of relief, high temperature is the key word of the weather next week, more weather information please pay attention to **Tianjin** weather. |
| Extractive [32] | **重庆**身边事。今天早上有**暴雨**袭击。Things around **Chongqing**. **Heavy rain** hit this morning. |
| MeanSum [26] | **重庆**今天白天有大到**暴雨**。**北京大风**，今天白天最高温度:39 度。**Chongqing** has **heavy rain** during the day today. **Beijing** is **gale**, and the highest temperature today is 39 degrees. |
| Copycat [33] | **重庆**下**暴雨**伴随**雷电**预警，**北京大风**。**Heavy rain** in **Chongqing** accompanied by **thunder** warning, **gale** in **Beijing**. |
| **Multi-KGS** | **重庆**今天又是**暴雨**。**福建**、**龙川**发布**雷电**黄色预警信号。今天**北京大风**。**Chongqing** has another **heavy rain** today. **Fujian** and **Longchuan** issued yellow **thunder** warning signals. **Beijing** is **gale** today. |

TABLE VI: Ablation study results in content evaluation.

| | ROUGE-1 | ROUGE-2 | ROUGE-L |
|---|---|---|---|
| w/o EKG | 0.1937 | 0.0661 | 0.1646 |
| w/o LKG | 0.1877 | 0.0667 | 0.1594 |
| **Multi-KGS** | **0.2068** | **0.0713** | **0.1713** |

TABLE VII: Ablation study results in sentiment evaluation.

| | *Acc* | *P* | *R* | $F_1$ |
|---|---|---|---|---|
| w/o EKG | 0.813 | 0.643 | 0.627 | 0.635 |
| w/o LKG | 0.819 | 0.651 | 0.636 | 0.643 |
| **Multi-KGS** | **0.867** | **0.671** | **0.663** | **0.668** |

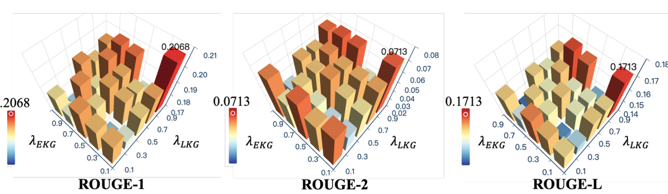


Fig. 5: Verification histogram showing the effect of $\lambda_{EKG}$ and $\lambda_{LKG}$ on model performance.

the metrics achieve the best results when $\lambda_{EKG} = 0.1$ and $\lambda_{LKG} = 0.9$. This result proves that when $\lambda_{EKG} = 0.1$ and $\lambda_{LKG} = 0.9$, the multiple knowledge guidance module positively impacts the summary generation module. This conclusion also supports the setting of these hyperparameters during the experiment.

## VII. Conclusion and Future Work

This paper proposes a multiple knowledge-guided summarization model, Multi-KGS, to enhance meteorological decision-making by fusing information from social networks. By integrating a summary generation module and a multiple knowledge guidance module, the model ensures that generated decision reports are both comprehensive and contextually relevant. The effectiveness of Multi-KGS in content and sentiment evaluation demonstrates the potential of GenAI-driven information fusion in real-world emergency management scenarios.

Future work will extend the Multi-KGS framework to improve the granularity and contextual sensitivity of the generated summaries. In particular, we aim to increase the model's responsiveness to low-frequency but critical social media posts, ensuring that minority yet essential information is incorporated into the decision reports. We will also refine the model's capacity to preserve numerical information such as timestamps and severity levels, thereby improving summary precision. To further strengthen information fusion, we plan to integrate a Retrieval-Augmented Generation mechanism, allowing dynamically retrieved meteorological knowledge from multiple sources to inform the generation process. This will support a more robust decision support system that combines real-time social signals with authoritative meteorological data, advancing the application of generative AI in emergency management.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant 62406036, the National Key Research and Development Program of China under Grant 2024YFC3308500, and also sponsored by SMP-Zhipu.AI Large Model Cross-Disciplinary Fund under Grant ZPCG20241029322.


## References


[1] D. Degroot, K. Anchukaitis, M. Bauch, J. Burnham, F. Carnegy, J. Cui, K. de Luna, P. Guzowski, G. Hambrecht, H. Huhtamaa, *et al.*, "Towards a rigorous understanding of societal responses to climate change," *Nature*, vol. 591, no. 7851, pp. 539–550, 2021.

[2] J. Palmer, "The new science of volcanoes harnesses ai, satellites and gas sensors to forecast eruptions.," *Nature*, vol. 581, no. 7808, pp. 256–260, 2020.

[3] A. Ajay, Y. Du, A. Gupta, J. B. Tenenbaum, T. S. Jaakkola, and P. Agrawal, "Is conditional generative modeling all you need for decision making?," in *The Eleventh International Conference on Learning Representations*, 2023.

[4] K. Shi, H. Lu, Y. Zhu, and Z. Niu, "Automatic generation of meteorological briefing by event knowledge guided summarization model," *Knowledge-Based Systems*, vol. 192, p. 105379, 2020.

[5] K. Shi, Y. Wang, H. Lu, Y. Zhu, and Z. Niu, "Ekgtf: A knowledge-enhanced model for optimizing social network-based meteorological briefings," *Information Processing & Management*, vol. 58, no. 4, p. 102564, 2021.

[6] M. Cao, T. Gai, J. Wu, F. Chiclana, Z. Zhang, Y. Dong, E. Herrera-Viedma, and F. Herrera, "Social network group decision making: Characterization, taxonomy, challenges and future directions from an ai and llms perspective," *Information Fusion*, p. 103107, 2025.

[7] Y. Shen, X. Ma, H. Zhang, and J. Zhan, "Fusion social network and regret theory for a consensus model with minority opinions in large-scale group decision making," *Information Fusion*, vol. 112, p. 102548, 2024.

[8] K. Shi, X. Peng, H. Lu, Y. Zhu, and Z. Niu, "Application of social sensors in natural disasters emergency management: A review," *IEEE Transactions on Computational Social Systems*, vol. 10, no. 6, pp. 3143–3158, 2023.

[9] L. Yin, F. Ping, J. Mao, and S. Jin, "Analysis on precipitation efficiency of the "21.7" henan extremely heavy rainfall event," *Advances in Atmospheric Sciences*, vol. 40, no. 3, pp. 374–392, 2023.

[10] K. Shi, X. Sun, D. Wang, Y. Fu, G. Xu, and Q. Li, "LLaMA-E: Empowering E-commerce authoring with object-interleaved instruction following," in *Proceedings of the 31st International Conference on Computational Linguistics* (O. Rambow, L. Wanner, M. Apidianaki, H. Al-Khalifa, B. D. Eugenio, and S. Schockaert, eds.), (Abu Dhabi, UAE), pp. 870–885, Association for Computational Linguistics, Jan. 2025.

[11] K. Shi, X. Sun, L. He, D. Wang, Q. Li, and G. Xu, "AMR-TST: Abstract Meaning Representation-based text style transfer," in *Findings of the Association for Computational Linguistics: ACL 2023* (A. Rogers, J. Boyd-Graber, and N. Okazaki, eds.), (Toronto, Canada), pp. 4231–4243, Association for Computational Linguistics, July 2023.

[12] M. Xu, D. Cai, W. Yin, S. Wang, X. Jin, and X. Liu, "Resource-efficient algorithms and systems of foundation models: A survey," *ACM Comput. Surv.*, vol. 57, Jan. 2025.

[13] L. Huang, W. Yu, W. Ma, W. Zhong, Z. Feng, H. Wang, Q. Chen, W. Peng, X. Feng, B. Qin, *et al.*, "A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions," *ACM Transactions on Information Systems*, vol. 43, no. 2, pp. 1–55, 2025.

[14] K. Shi, X. Sun, Q. Li, and G. Xu, "Compressing long context for enhancing rag with amr-based concept distillation," *arXiv preprint arXiv:2405.03085*, 2024.

[15] X. Zou, Y. Yan, X. Hao, Y. Hu, H. Wen, E. Liu, J. Zhang, Y. Li, T. Li, Y. Zheng, *et al.*, "Deep learning for cross-domain data fusion in urban computing: Taxonomy, advances, and outlook," *Information Fusion*, vol. 113, p. 102606, 2025.

[16] D. Das, M. Kumar, and A. Rudnicky, "Automatic extraction of briefing templates," in *Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-I*, 2008.

[17] E. Reiter, S. Sripada, J. Hunter, J. Yu, and I. Davy, "Choosing words in computer-generated weather forecasts," *Artificial Intelligence*, vol. 167, no. 1, pp. 137–169, 2005. Connecting Language to the World.

[18] S. Wang, Y. Du, X. Guo, B. Pan, Z. Qin, and L. Zhao, "Controllable data generation by deep learning: A review," *ACM Comput. Surv.*, vol. 56, Apr. 2024.

[19] S. Narayan, Y. Zhao, J. Maynez, G. Simões, V. Nikolaev, and R. McDonald, "Planning with learned entity prompts for abstractive summarization," *Transactions of the Association for Computational Linguistics*, vol. 9, pp. 1475–1492, 2021.

[20] X. Chen, H. Alamro, M. Li, S. Gao, X. Zhang, D. Zhao, and R. Yan, "Capturing relations between scientific papers: An abstractive model for related work section generation," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, (Online), pp. 6068–6077, Association for Computational Linguistics, Aug. 2021.

[21] F. Zhu, S. Tu, J. Shi, J. Li, L. Hou, and T. Cui, "Twag: A topic-guided wikipedia abstract generator," *arXiv preprint arXiv:2106.15135*, 2021.

[22] Z. Wang, Z. Duan, H. Zhang, C. Wang, L. Tian, B. Chen, and M. Zhou, "Friendly topic assistant for transformer based abstractive summarization," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 485–497, 2020.

[23] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, (Minneapolis, Minnesota), pp. 4171–4186, Association for Computational Linguistics, June 2019.

[24] T. Ma, Q. Pan, H. Rong, Y. Qian, Y. Tian, and N. Al-Nabhan, "T-bertsum: Topic-aware text summarization based on bert," *IEEE Transactions on Computational Social Systems*, pp. 1–12, 2021.

[25] Y. Du, Q. Li, L. Wang, and Y. He, "Biomedical-domain pre-trained language model for extractive summarization," *Knowledge-Based Systems*, vol. 199, p. 105964, 2020.

[26] E. Chu and P. Liu, "Meansum: a neural model for unsupervised multi-document abstractive summarization," in *International Conference on Machine Learning*, pp. 1223–1232, PMLR, 2019.

[27] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," *Neural Comput.*, vol. 1, p. 270–280, June 1989.

[28] E. Jang, S. Gu, and B. Poole, "Categorical reparameterization with gumbel-softmax," *arXiv preprint arXiv:1611.01144*, 2016.

[29] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pp. 315–323, JMLR Workshop and Conference Proceedings, 2011.

[30] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[31] H. Tian, C. Gao, X. Xiao, H. Liu, B. He, H. Wu, H. Wang, and F. Wu, "Skep: Sentiment knowledge enhanced pre-training for sentiment analysis," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 4067–4076, 2020.

[32] G. Rossiello, P. Basile, and G. Semeraro, "Centroid-based text summarization through compositionality of word embeddings," in *Proceedings of the MultiLing 2017 Workshop on Summarization and Summary Evaluation Across Source Types and Genres*, pp. 12–21, 2017.

[33] A. Bražinskas, M. Lapata, and I. Titov, "Unsupervised opinion summarization as copycat-review generation," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault, eds.), (Online), pp. 5151–5169, Association for Computational Linguistics, July 2020.