

Research Article

RAPOT: An Adaptive Multifactor Risk Assessment Framework on Public Opinion for Trial Management

Weina Jiang ¹, Qi Yong,¹ Ning Liu,¹ and Yuze Luo ²

¹School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China

²National Engineering Research Center of Digital Life, Sun Yat-sen University, Guangzhou 510006, China

Correspondence should be addressed to Yuze Luo; yzluo26@126.com

Received 8 January 2021; Revised 8 April 2021; Accepted 17 April 2021; Published 17 May 2021

Academic Editor: Amr Tolba

Copyright © 2021 Weina Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Since public opinion from social media has a growing impact and supervision on trial, risk assessment on public opinion is increasingly important in refined trial management. However, the tremendous amount of public opinion and the insufficient historical logs of trial procedures bring challenges to risk assessment on public opinion. To address this, we propose an adaptive multifactor risk assessment framework on public opinion with fuzzy numbers. Initially, we establish a multilayer indicator model for assessing the risk of public opinion (POR) with multilayer analysis and decision methods. Then, we explore the association rules hidden in the process logs to update the indicator model periodically. Moreover, we design a public opinion analysis module for indicator evaluation, including analysis in public opinion sentiment, hot search, and social media coverage to deal with big data on social media. Especially, the public opinion sentiment is classified by topic-based BiLSTM (T-BiLSTM), which is more accurate. Finally, the fuzzy number similarity is employed to determine POR's level in the nine-level risk system. Experimental results validate the efficiency of our framework when assessing the POR.

1. Introduction

Serious and complicated cases bring severe challenges to trial management nowadays. Some of them have raised much attention due to their case type, related parties, and well-known judges. Simultaneously, people are used to expressing their opinions for the concerned cases on platforms such as Facebook, WeChat, Weibo, and Twitter. A mass of public opinion has both positive and negative impacts on trial procedures. Hence, public opinion assessment and supervision are crucial for credible trials. Actually, public opinion in social media has its characteristics, such as mass amount, fast propagation, and chaotic content. Furthermore, the mass data in social media reveals the inherent information we are concerned about. After analyzing multisource public opinion comprehensively, we could figure out its propagation mode to make POR's warning come earlier. Therefore, POR assessment is beneficial for early responding to negative public opinions and improving the court's initiative ability. There are two main tasks while accomplishing the task. One is to

handle public opinion with big data theory, and the other is to conduct the risk assessment with insufficient historical data.

For the explosive comments that emerge on social media, sentiment analysis has become a research hotspot. Besides, sentiment analysis on comments about hot cases plays a vital role in promoting trial management. Thus, it is crucial to carry out an efficient analysis and supervision method for comments about cases. So far, research on machine learning-based sentiment analysis has a lot of achievements, such as KNN [1], maximum entropy [2], SVM [3], and Bayes [4]. Nowadays, with the rapid development and outstanding performance of deep learning, many researchers concentrate on methods with CNN [5], RNN [6], and LSTM [7] to improve the classification accuracy and have significant progress.

For risk assessment, due to insufficient historical data, together with the fuzziness and uncertainty of risks, researchers adopt a fuzzy set theory to analyze the risk [8]. Singh et al. [9] propose an assessment framework for risk

analysis of food disaster based on fuzzy similarity, and they quantitatively calculate the risk level of targets separately. The fuzzy similarity-based method performs well with a quantitative risk assessment for trial cases.

However, there still exists some challenges to achieve the assessment of POR. Firstly, there is no suitable indicator model for this task. An efficient assessment relies on fine-grained indicators and objective weights for indicators, and it remains unsolved. Secondly, comments about cases in the trial on social media have many characteristics that are hard to analyze. Hence, it remains much work to ensure the accuracy of sentiment classification for the specific use. Thirdly, how to evaluate risks quantitatively is not easy but crucial.

To address these issues, this paper implements a Risk Assessment framework on Public Opinion for Trial management (RAPOT). The framework provides a fine-grained risk assessment based on fuzzy numbers. By computing fuzzy number similarities, the framework decides its risk level in the nine-level assessment system. Our main contributions in this paper are as follows:

- (i) Fine-Grained Risk Rating System. We employ fuzzy number similarities to achieve risk assessment with little historical data in trial procedure management. At first, a multilayer risk indicator model is established based on the analytic hierarchy process (AHP) method and extended technique for order preference by similarity to an ideal solution (extended TOPSIS) method. The model contains a fine-grained indicator layer, and each one contains a risk indicator and its impact factor. When assessing the risks, we transform both impact factors and indicator values into fuzzy numbers. Then, we aggregate the fuzzy numbers into one and rank the integrated one in the nine-level assessment system
- (ii) Adaptive Indicator Model. Considering that the system logs accumulated during trial processing contain many latent association rules of the procedures, we propose the RAPriori algorithm to explore the association rules. These latent rules are updated to the indicator model for improving the applicability and robustness of the model
- (iii) Efficient Comment Sentiment Analysis. We define three kinds of input sources and submodules for indicator evaluation. Significantly, the sentiment of public opinion is classified based on topics. The sentiment analysis that we propose consists of single-pass-based topic clustering and T-BiLSTM-based sentiment analysis. Sentiment analysis for topics is precise and more comprehensive. Besides, our framework has extensive indicators such as the topic's heat and coverage of media
- (iv) Experimental Evaluations. To demonstrate the performance of RAPOT, we conduct a case study with three cases that are paid much attention recently. The results illustrate that our framework is applica-

ble and efficient in practical cases with a reasonable assessment level

The rest of this paper is structured as follows. We talk about the related work in Section 2. The RAPOT framework is described in Section 3. In Section 4, we illustrate the experimental results, and we conclude the paper in Section 5.

2. Related Work

Due to the fuzziness and uncertainty of risks, researchers adopt a fuzzy set theory to analyze the risk. The theory of fuzzy numbers has been widely applied in risk analysis [10], approximate reasoning [11], and risk pattern recognition [12]. For risk analysis, the existing methods can be divided into the fuzzy ranking-based [13], fuzzy inference-based [14, 15], fuzzy matrix-based [16, 17], and fuzzy number similarity-based [9] risk assessment models. Zhang et al. [13] figure out the risky area based on a water security evaluation framework by comparing the risk of related areas. Hence, the qualitative analysis measures the risk level comparatively. Nevertheless, in the trial, the POR of the two cases cannot be compared at the same pace. Karasan et al. [14] propose the safety and critical effect analysis (SCEA). Furthermore, it adopts Pythagorean fuzzy sets [18] to provide a comprehensive risk assessment. However, fuzzy inference-based methods are usually used in industry and are not suitable for trial applications. Can et al. [16] present a three-stage fuzzy risk matrix-based risk assessment and dynamically combine multicriteria decision-making with fuzzy logic. Though fuzzy matrix-based methods can reduce risk ties [19] efficiently, they still provide a qualitative assessment that is not precise enough. As for similarity-based method, Khorshidi and Nikfalazar [20] present an improved method to compute the degree of similarity between generalized fuzzy numbers. The proposed method has been used for fuzzy risk analysis, and it could determine each manufacturer's risk level. In summary, the similarity-based model is suitable for the quantitative risk assessment for an individual object. At the same time, risks in the trial process management system (TPMS) are quite fuzzy and uncertain in fact. Besides, the historical data have not been digitalized well. So we adopt a fuzzy number similarity-based model to achieve risk assessment.

The existing fuzzy number similarity-based methods always have three main modules. They are the risk indicator model, risk aggregation, and risk level determination. Among them, fuzzy number similarity calculation is important for risk level determination precisely. Referring to fuzzy number similarities, researchers have defined various features of generalized fuzzy number (GFN) to distinguish the numbers, such as the center of gravity (COG) [21], the area [20], and the radius of gyration [22]. Then, researchers adopt geometric distance, Hausdorff distance [23], and so on to measure the similarity of the feature values. Xu et al. [24] present a COG-based method while with a limitation that two different fuzzy numbers may have the same COG. To address the limitation, Yong et al. [25] employ ROG of the area to measure the similarities. Moreover, Chutia and Gogoi [10] expand

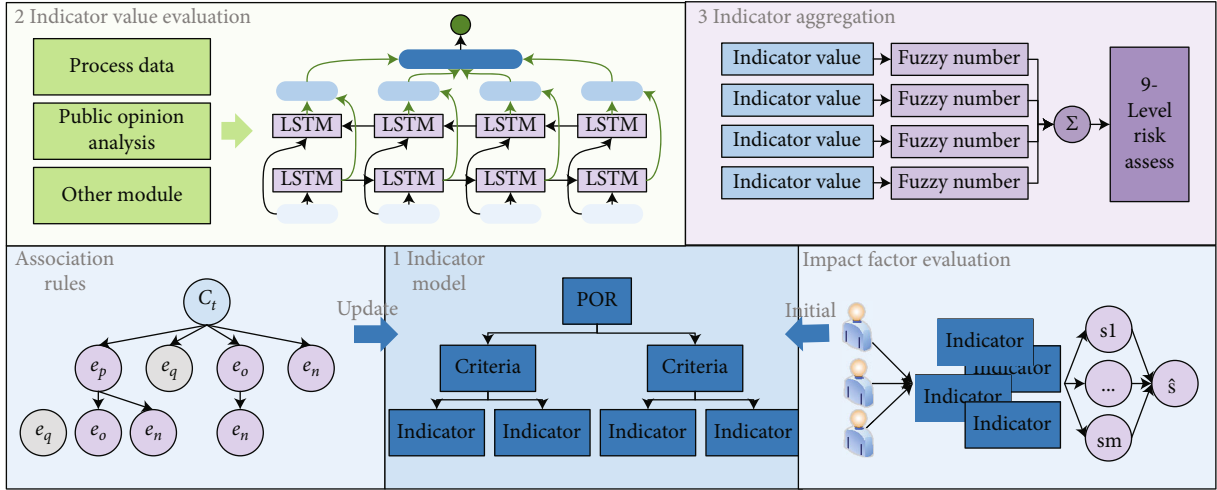


FIGURE 1: The framework of RAPOT.

GFN with left height and the right height to further distinguish traditional GFNs with the same COG. However, these two methods still suffer from invalid results. Therefore, we select a similarity measurement on generalized fuzzy numbers to map the integrated fuzzy number into a linguistic term in the nine-level risk system [26]. The similarity measure algorithm we employ constrains the similarity of two fuzzy numbers in the range of $[0, 1]$, with fewer invalid results and at the same time has a high distinguishability.

3. Framework of Risk Assessment of Public Opinion

In this section, we discuss the critical issues while assessing the POR. Firstly, we present the risk indicator model in Sections 3.1 and 3.2. Secondly, we talk about evaluating risk indicators and the public opinion sentiment analysis in Section 3.3. Then, we explain the fuzzy number similarity-based risk assessment method in Section 3.4.

Figure 1 is the framework of our RAPOT. It includes a risk indicator model, indicator evaluation module, and risk aggregation module. In the beginning, a multilayer indicator model is built to define the fine-grained risk indicators with corresponding impact factors, and the model is dynamically updated by exploring new association rules on system logs. Then, the indicator evaluation module figures out the indicator values based on process data, data from the public opinion analysis module, and the other systems. The indicator aggregation module is aimed at deciding the risk level with the impact factors and the risk probabilities.

3.1. Risk Indicator Model Initialization. To overcome the difficulty of lacking historical data, we employ AHP and extended TOPSIS to construct an initialized risk indicator model. The hierarchy model defines amounts of risk indicators along with their impact factors. Figure 2 describes the procedures for building our indicator model. First, a hierarchy structure is built, and an evaluation dataset for risk indicators is collected based on AHP. Then, an evaluation matrix for the risk indicators is constructed based on the collected

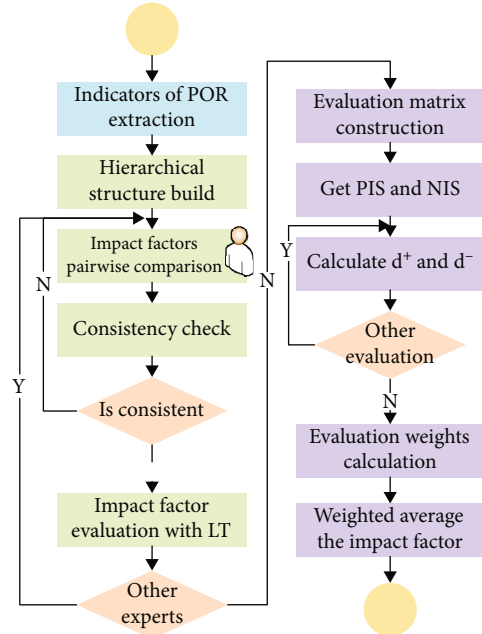


FIGURE 2: The flowchart of risk indicator model initialization.

dataset. We adopt extended TOPSIS to analyze the evaluation dataset to calculate the impact factors of the indicators. Our model construction method combines AHP and extended TOPSIS to work out a group of accurate impact factors with limited historical data.

3.1.1. Hierarchical Structure Determination. AHP is an efficient multilayer analysis and decision method [27, 28]. It first composes the decision problem into a hierarchy of subproblems that each one can be treated independently. Once the hierarchy is built, the expert group evaluates the elements in the same layer by comparing them to each other according to their impact on the father element. Table 1 shows the 1-9 scales used to evaluate each element’s impact factor. The AHP transforms the evaluations to numerical values that can be calculated over the decision problem’s entire range.

TABLE 1: 1-9 scales of relative importance [29].

Intensity of importance	Definition
1	Equal importance
2	Weak
3	Moderate importance
4	Moderate plus
5	Strong importance
6	Strong plus
7	Very strong or demonstrated importance
8	Very, very strong
9	Extreme importance

Finally, a priority is derived for each element in the hierarchy by iteratively verifying the comparison matrix's consistency after adjusting the priorities each time.

At first, we refer to expertise, existing laws, regulations, and the classical hot cases and form the set of risks as $R = \{r_1, r_2, \dots, r_N\}$, where N is the number of risks. Then, the hierarchical structure is established based on AHP. As shown in Figure 3, our risk indicator model consists of three layers:

- (i) Objective Layer (OL). Risk assessment of public opinion for trial management is the objective of our work. We need to figure out the impacts of public opinion on the trial procedure
- (ii) Criteria Layer (CL). The elements in this layer are the judge, the parties involved, the case, and the public opinion. The expert group defines the elements referring to the existing documents
- (iii) Indicator Layer (IL). This layer contains the indicators which would impact the trial procedure by public opinion. Each indicator belongs to their father elements in the criteria layer

After that, an evaluation dataset is collected to gain the indicators' impact factors, and the impact factor represents the indicator's weight when integrating the POR. To evaluate the impact factor accurately, the expert compares the risk indicators with pairs to complete a comparison matrix as

$$\Delta = \begin{bmatrix} \delta_{11} & \cdots & \delta_{1N} \\ \vdots & \ddots & \vdots \\ \delta_{N1} & \cdots & \delta_{NN} \end{bmatrix}, \quad (1)$$

where δ_{ij} is the comparison value of r_i and r_j . The expert assigns the value δ according to Table 1. Then, the consistency of Δ has to be verified by

$$CI = \frac{\lambda - n}{n - 1}, \quad (2)$$

where λ is the maximum eigenvalue of Δ and n is the dimension of the matrix. The consistency is complete when $CI = 0$ and decreases with CI increasing. Then, AHP uses a random

consistency indicator RI to define a refined CR which is

$$CR = \frac{CI}{RI}. \quad (3)$$

When $CR < 0.1$, the matrix Δ is consistent and RI is a predefined dictionary [29]. If the validation fails, the expert has to adjust the comparison matrix until the validation comes to success.

The eigenvector of the approved evaluation matrix gives a sort of risk indicators by their impact factors. For risk assessment with fuzzy numbers, the expert assigns a linguistic term in $LT = \{\text{"AbsolutelyLow (AL)", "VeryLow (VL)", "Low (L)", "FairlyLow (FL)", "Medium (M)", "FairlyHigh (FH)", "High (H)", "VeryHigh (VH)", "AbsoluteHigh (AH)"}\}$ to each risk indicator based on the order.

3.1.2. Impact Factor Calculation. Several law experts evaluate the impact factors according to our hierarchical structure and construct an evaluation dataset. The dataset contains several evaluation items $\{l_1, l_2, \dots, l_M\}$ for M experts and l_m consists of $l_{m1}, l_{m2}, \dots, l_{mN}$, each item comes from a law expert for r_n in set R . Then, we employ TOPSIS to aggregate the evaluations of different experts. TOPSIS is a multicriteria decision analysis method, which identifies weights for each criterion by calculating the geometric distances from each alternative to the positive ideal solution and the negative ideal solution, respectively [30]. When evaluating the risk indicator's impact factor, the positive ideal solution is defined as the lowest impact on cost optimization. Namely, the lower impact of the risk indicator brings less cost in risk prevention and control. Hence, we adopt the extended TOPSIS [31] to calculate the impact factors for the POR assessment designed for the trial scene.

First, an evaluation matrix with linguistic terms is established based on the dataset as

$$L = \begin{bmatrix} l_{11} & \cdots & l_{1N} \\ \vdots & \ddots & \vdots \\ l_{M1} & \cdots & l_{MN} \end{bmatrix}, \quad (4)$$

where M, N are the number of experts and risk indicators. In the matrix, l_{mn} is given by expert m for the indicator n to measure the importance of the indicator. And then, l_{mn} is transformed into a fuzzy number according to Table 2 for weight fusion of impact. After that, we get an evaluation matrix with fuzzy numbers.

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1N} \\ \vdots & \ddots & \vdots \\ s_{M1} & \cdots & s_{MN} \end{bmatrix}, \quad (5)$$

here s_{mn} is a generalized fuzzy number represented as $(a, b, c, d; w)$ and $a, b, c, d, w \in R$.

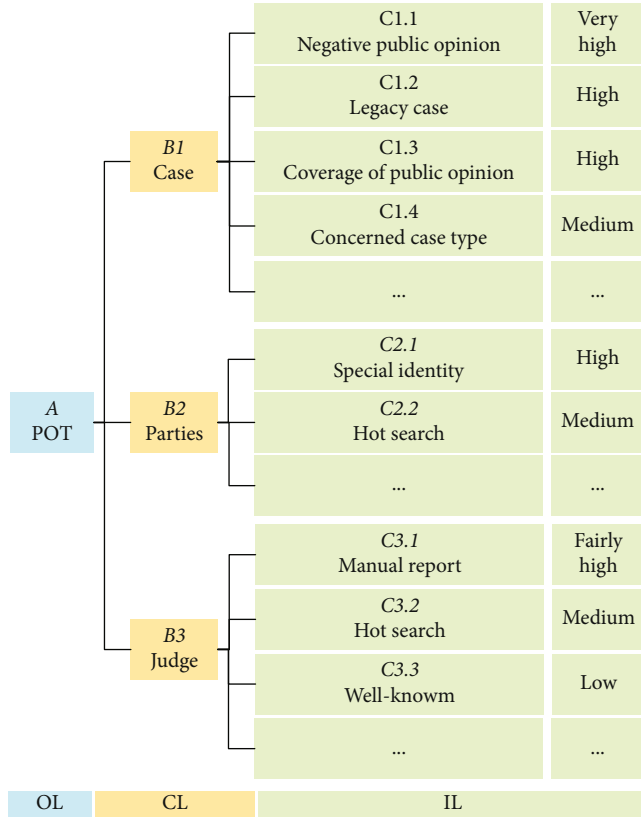


FIGURE 3: The indicator model for the POR.

TABLE 2: The transform from linguistic terms to the fuzzy numbers [26].

Linguistic terms	Generalized fuzzy numbers
AbsolutelyLow	(0.0, 0.0, 0.0, 0.0; 1.0)
VeryLow	(0.0, 0.0, 0.02, 0.07; 1.0)
Low	(0.04, 0.1, 0.18, 0.23; 1.0)
FairlyLow	(0.17, 0.22, 0.36, 0.42; 1.0)
Medium	(0.32, 0.41, 0.58, 0.65; 1.0)
FairlyHigh	(0.58, 0.63, 0.80, 0.86; 1.0)
High	(0.72, 0.78, 0.92, 0.97; 1.0)
VeryHigh	(0.93, 0.98, 1.0, 1.0; 1.0)
AbsolutelyHigh	(1.0, 1.0, 1.0, 1.0; 1.0)

In the extended TOPSIS, the positive and negative ideal solutions are

$$\begin{aligned} \text{PIS} &= [s_1^+, s_2^+, \dots, s_N^+], \\ \text{NIS} &= [s_1^-, s_2^-, \dots, s_N^-], \end{aligned} \quad (6)$$

here, s_n^+ and s_n^- are defined as

$$\begin{cases} s_n^+ = \min(s_{mn}) \\ s_n^- = \max(s_{mn}) \end{cases}, \quad (7)$$

Then, the distance between $s_m = [s_{m1}, s_{m2}, \dots, s_{mN}]$ and the positive ideal solution PIS is calculated as

$$d^+ = d(s_m, \text{PIS}) = \sqrt{\sum_{n=1}^N \boxtimes \sum_{j=1}^4 \boxtimes (s_{mnj} - s_{nj}^+)^2}. \quad (8)$$

Similarly, the geometric distance between s_m and the negative ideal solution NIS is

$$d^- = d(s_m, \text{NIS}) = \sqrt{\sum_{n=1}^N \boxtimes \sum_{j=1}^4 \boxtimes (s_{mnj} - s_{nj}^-)^2}, \quad (9)$$

here $s_{mn} = (s_{mn1}, s_{mn2}, s_{mn3}, s_{mn4}; w)$, $s_n^+ = (s_{n1}^+, s_{n2}^+, s_{n3}^+, s_{n4}^+; w)$, and $s_n^- = (s_{n1}^-, s_{n2}^-, s_{n3}^-, s_{n4}^-; w)$ are generalized fuzzy numbers. After that, we obtain the weight of each alternative s_m by normalizing the distance ratios as

$$\lambda_m = \text{II} \left(\frac{2}{|d_m^+ - d_m^-|} \right). \quad (10)$$

Finally, the impact factor of indicator n is calculated by weighted summing the alternatives as

$$\hat{s}_n = \frac{\sum_{m=1}^M \lambda_m \boxtimes s_{mn}}{\sum_{m=1}^M \lambda_m}. \quad (11)$$

3.2. Risk Indicator Model Update. Considering that the trial process is strict and complicated, POR's initial indicator model can be hardly applicable to the POR assessment continuously. Also, the system logs accumulated during trial processing contain many latent association rules of the procedures. Figure 4 shows a fragment of the trial process, each block is a process node, and each ellipse represents the risk confirmation. Therefore, we propose a reversed Apriori (RApriori) algorithm to explore the association rules hide in the system logs. The association rule we want to search is defined as $[e_i, \dots, e_j, c_t]$, here e_i represents a failed rule check in the process node i and c_t is a risk confirmation node. By investigating the practical TPMS, we figure out the process nodes are arranged in a single sequence. According to it, we optimize the classical Apriori by ordering the nodes and extending the association set in reverse. The details of the proposed RApriori are shown in Algorithm 1.

In the algorithm, we assign numerical codes to both process nodes and risk confirm nodes based on their sequence in trial. Firstly, the search of latent association rules always starts from a frequent risk confirm node c_t and set it as the root of the tree T we show in Figure 5. Secondly, the frequent process nodes whose numerical codes less than c_t are reversely sorted in a candidate list $[e_q, e_p, \dots, e_i]$. Thirdly, we join each item in the list with c_t to form a set separately, such as $\{c_t, e_q\}$, and then check the corresponding support score

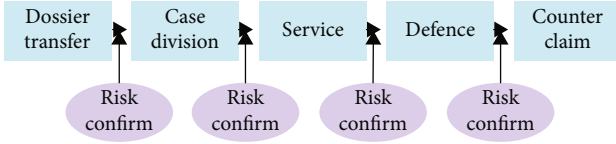


FIGURE 4: A fragment of the trial process.

```

Require: system logs generate from T1 to T2
Ensure: association rules  $[[e_i, \dots, e_j, c_t], \dots]$ 
1: total = len(logs)
2: ens = getSupportErrorNodes(logs)
3: cns = getSupportConfirmNodes(logs)
4: for  $t = 1; t < \text{len}(cns); t + \mathbf{do}$ 
5:   cn = cns[t]
6:   rlogs = getRelatedLogs(cn)
7:   rens = getRelatedErrorNode(cn)
8:   cSet = joinSet(cn, rens)
9:   cSet = getSupportSet(cSet, rlogs, total, sRecord)
10:  sList = sort(cSet)
11:  k = 2
12:  repeat
13:    fSet = union(cSet)
14:    cSet = joinSet(cSet, sList, k + 1)
15:    cSet = getSupportSet(cSet, rlogs, total, sRecord)
16:    k = k + 1
17:  until cSet is empty
18: end for
19: rules = getConfidenceRule(fSet)

```

ALGORITHM 1: RApriori

to create layer 2. The support score is defined as

$$\text{support}(A) = \frac{\text{count}(A)}{|D|}, \quad (12)$$

where A is a set and $|D|$ is the amount of the logs. Fourthly, the tree moves to the next layer by orderly combining a set in the current layer with items in the candidate list that are less than the minimum node in the set. Then, iteratively increase the height of T until there is no more satisfied new set. At last, we calculate the support score of the satisfied sets and work out the association rules. The support score is defined as

$$\text{support}(\{e_i, \dots, e_j\} \Rightarrow c_t) = \frac{\text{count}(\{e_i, \dots, e_j, c_t\})}{\text{count}(\{e_i, \dots, e_j\})}. \quad (13)$$

The RApriori method is executed regularly, and the searched association rules are added to update the indicator model of POR. The experimental results show that our algorithm decreases the computational complexity significantly.

3.3. Risk Indicator Evaluation and Public Opinion Analysis. Besides the indicator factor, we have to calculate the probability of indicator occurrence, which we call the indicator value. The data sources of value computing can be divided into three categories: (1) social media, (2) manual input,

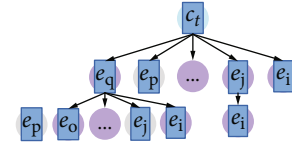


FIGURE 5: The procedures of joinSet operation.

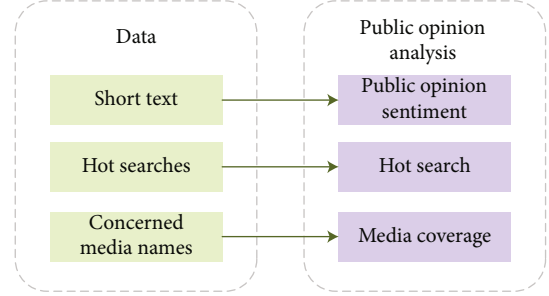


FIGURE 6: The structure of the social media analysis module.

and (3) document analysis. For indicator C3.1, the judge can report the POR during the trial. As for C1.2, C1.4, C2.1, and C3.3, the indicator values are determined by the other subsystems in the TPMS, for instance, the case division system. Apart from them, the values of indicators C1.1, C1.3, C2.2, and C3.2 are inferred from the social media analysis module. Figure 6 illustrates the structure of our module for social media analysis. It is composed of three parts listed as follows:

- (i) Analysis in Public Opinion Sentiment. This part explores how people are interested in the case and how intensely they discuss the related topics. If the public cares much about the case and shows negative sentiment in their expressions, the indicator value will be large. On the contrary, the indicator value will come near zero
- (ii) Analysis in Hot Search. The judge or the parties frequently searched in social media is an important indicator that this case may have the POT during the trial
- (iii) Analysis in Media Coverage. If the media in our maintained important-media list has taken part in the related topic, this case's media coverage will increase. The POT level increases with the coverage reaching a threshold

In this section, we mainly describe the public opinion sentiment based on topics. The comments collected from social media related to the case are divided into some topics to address this. Then, the texts and the related topics are fed into a neural network to train a classifier used to analyze the sentiment. The details are as follows.

3.3.1. Input Embedding. Firstly, a short text is split into a word sequence $W = \{w_1, w_2, \dots, w_n\}$ which contains n words. After that, we transform words to vectors by a Word2vec model [32] and obtain the embedding matrix E_W which consists of all word embeddings.

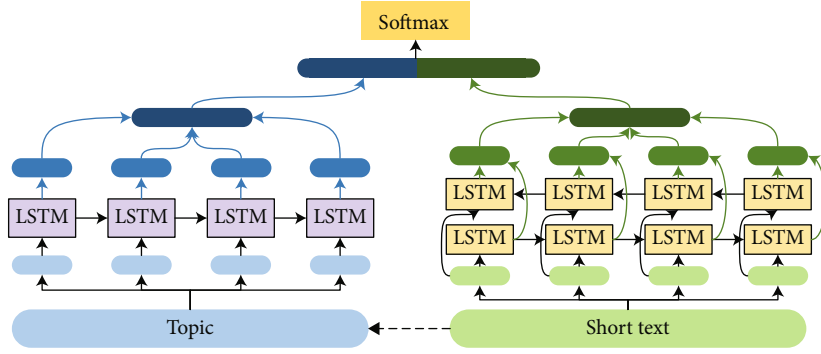


FIGURE 7: The structure of T-BiLSTM.

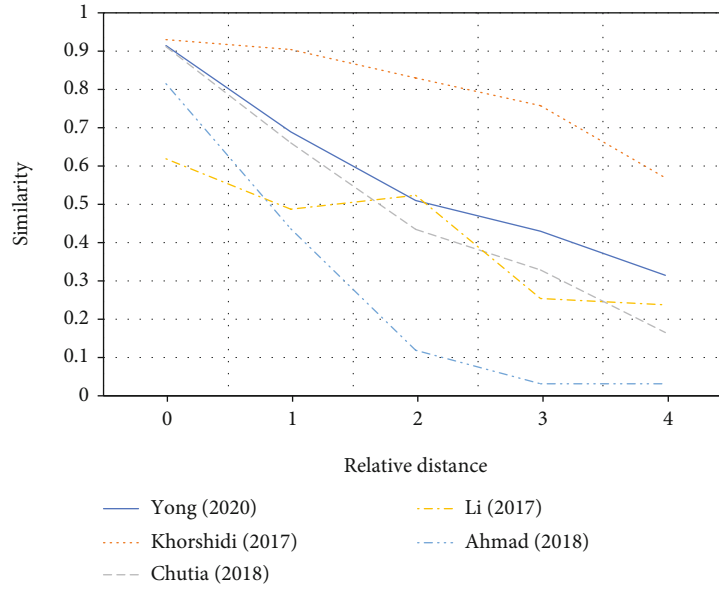


FIGURE 8: The comparison in attenuation of similarities with increasing distance for the five methods.

TABLE 3: The parameters of simulation datasets.

Parameters	Value
Count of process nodes	80
Count of confirm nodes	80
Error rate	0.15
Confirm rate	0.15

3.3.2. *Topic Clustering*. Single-pass clustering [33] with the cosine similarity is employed to iteratively partition m short texts into k clusters, the topics can be represented as $T = \{t_1, t_2, \dots, t_k\}$, and t is a set of some keywords. The similarity is calculated as

$$\text{dist}(s_i, s_j) = \frac{s_i \cdot s_j}{|s_i| |s_j|} \quad (14)$$

where s_i and s_j are vectors of two short texts. Then, the keywords in the clusters are detected to be the topics. Moreover,

we get the embedding matrix E_T , which contains all keyword embeddings of a topic through word embedding.

3.3.3. *T-BiLSTM-Based Comment Sentiment Analysis*. Since BiLSTM [34] has been proven efficient for sentiment analysis, we propose the T-BiLSTM network to train a text sentiment classifier. Figure 7 illustrates the structure of the T-BiLSTM. On the right side, we employ a BiLSTM layer to capture the contextual features of the text. On the left side, we adopt a LSTM layer to explore the contextual features of the topic. Next, we concatenate the outputs of both sides and feed it into a softmax layer. The above processes are represented as

$$\begin{aligned} H_W &= \text{BiLSTM}(E_W), \\ H_T &= \text{LSTM}(E_T), \\ H &= [H_W, H_T], \\ p &= \text{softmax}(W * H + b), \end{aligned} \quad (15)$$

	Case 1					Case 2					Case 3				
Risk levels	AL	AL	AL	AL	AL	AL	AL	AL	AL	AL	AL	AL	AL	AL	AL
	VL	VL	VL	VL	VL	⋮	⋮	⋮	⋮	⋮	VL	VL	VL	VL	VL
	L	L	L	L	L	M	M	M	M	M	L	L	L	L	L
	FL	FL	FL	FL	FL	FH	FH	FH	FH	FH	FL	FL	FL	FL	FL
	M	M	M	M	M	H	H	H	H	H	M	M	M	M	M
	[22]	[20]	[10]	[21]	[23]	[22]	[20]	[10]	[21]	[23]	[22]	[20]	[10]	[21]	[23]
	← Similarity measure algorithms →														

FIGURE 9: The comparison in results of five similarity measure methods.

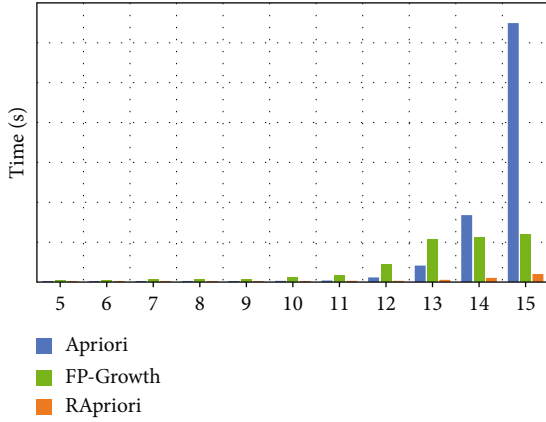


FIGURE 10: Time costs with different rule lengths.

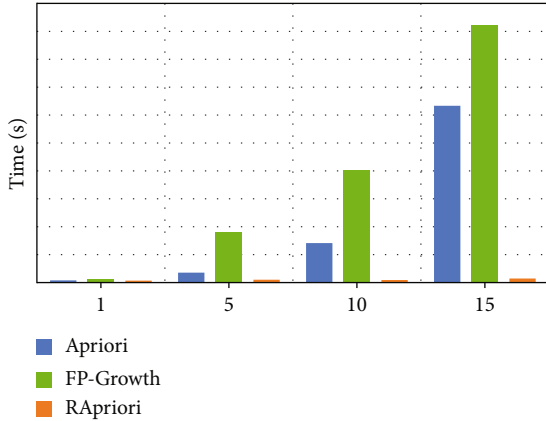


FIGURE 11: Time costs with different rule counts.

where W and b are the weight matrix and bias, respectively. In addition, we use cross-entropy loss to lead the network training.

3.3.4. *Evaluation of Indicator C1.1.* The public opinion sentiment for topics is defined as

$$f_i = \min \left(\max \left(\sum_{M=1}^{i=1} \frac{\text{neg}_i}{\tau}, 1 \right), 9 \right). \quad (16)$$

Here, neg_i is the number of negative comments in topic t_i , τ is the threshold which is used to testify whether a topic is discussed widely, and the evaluation of indicator C1.1 is calculated as

$$v = \left\lceil \frac{N_1 \times f_1 + N_2 \times f_2 + \dots + N_k \times f_k}{M} \right\rceil, \quad (17)$$

where N_i is the count of texts in topic t_i , and M is the total amount of texts in the case.

3.4. *Risk Assessment on Public Opinion for Trial Management.* In this section, we describe the fuzzy number similarity-based risk assessment module which evaluates the risk level in the nine-level risk system. At first, the risk indicator evaluations we talk about in Section 3.3 are converted into fuzzy numbers as

$$\text{Evaluate}_i = \mathbf{GFNS}_{v_i,1}. \quad (18)$$

Here, $v_i \in [1, 9]$ and $\mathbf{GFNS}_m = (It_m, \mathbf{GFN}_m)$; It_m is a linguistic term which is in $\{AL, VL, L, FL, M, FH, H, VH, AH\}$, and \mathbf{GFN}_m is a generalized fuzzy number defined in Table 2. Since the risk of public opinion has various indicators, the risk assessment module aggregating risk of each indicator by the weighted average method is

$$R = \frac{\sum_{i=1}^N \boxtimes \text{Impact}_i \otimes \text{Evaluate}_i}{\sum_{i=1}^N \boxtimes \text{Impact}_i}. \quad (19)$$

As Figure 8 shows, the selected method's similarity drops smoothly with the distance increases compared with the other algorithms. The risk level is calculated as

$$m = \underset{m}{\operatorname{argmax}} \text{Similarity}(R, \mathbf{GFN}_m). \quad (20)$$

4. Experiment

In this section, we discuss the results of the three experiments: (A) efficiency of algorithm RApriori, (B) efficiency of the classifier T-BiLSTM, and (C) the case study of the whole framework RAPOT.

4.1. Efficiency of RApriori. To validate the efficiency of RApriori, we compare it with the classical Apriori and FP-Growth. There are three subexperiments in this section: (a) time costs with different rule lengths, (b) time costs with different rule counts, and (c) time costs with different datasets. We carry on these experiments on the simulation datasets generated with the parameters shown in Table 3. In experiment (a), we employ Apriori, FP-Growth, and RApriori to work out rules with different lengths. Figure 10 shows that Apriori and FP-Growth's time costs sharply increase with more extended rules. In experiment (b), we compare the three methods for dealing with different counts of rules. Figure 11 illustrates our method's time cost grows slower than the other methods. In experiment (c), we conduct the three algorithms on three datasets with different data sizes. Figure 12 shows that our method has a better efficiency than Apriori and FP-Growth while tolerating data explosion.

4.2. Efficiency of T-BiLSTM. We train the classifier for public opinion sentiment analysis with the dataset contains 18000 positive comments and 18000 negative comments come from Weibo. The validating set has 3600 positive items and 3600 negative items. In addition, we compare the T-BiLSTM-based sentiment classifier with the KNN, maximum entropy, Bayes, SVM, and traditional BiLSTM. We adopt accuracy, positive-precision, positive-recall, and Macro-F1 as the evaluation metrics that are defined as

$$\begin{aligned}
 \text{Accuracy} &= \frac{T}{N}, \\
 \text{Precision} &= \frac{TP}{TP + FP}, \\
 \text{Recall} &= \frac{TP}{TP + FN}, \\
 \text{F1} &= 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \\
 \text{Macro-F1} &= \frac{1}{2} (\text{F1}_{\text{pos}} + \text{F1}_{\text{neg}}),
 \end{aligned} \tag{21}$$

where T is the number of correct predictions, and N is the total number of valide samples. For TP and FP , they represent the amount of the predicted "Positive" samples which are correct and incorrect, respectively, which are similar to TN and FN . As for Macro-F1, it is defined as the average of F1_{pos} and F1_{neg} and is used to evaluate the efficiency of each classifier comprehensively. Table 4 shows the comparison result, and we can see that our T-BiLSTM exceeds the other methods.

4.3. Case Study of RAPOT. In this section, we evaluate the efficiency and applicability of RAPOT with a case study. It includes three sets of short texts corresponding to three cases; the size of the three sets are 764, 306, and 156. At first, the risk indicator model of RAPOT is shown as Figure 3. There are nine indicators in the aspects of the case, the related parties, and the judge. Then, we figure out the indicator values for each case, and the mapped linguistic terms are shown in Table 5. In the next step, the linguistic terms are turned into

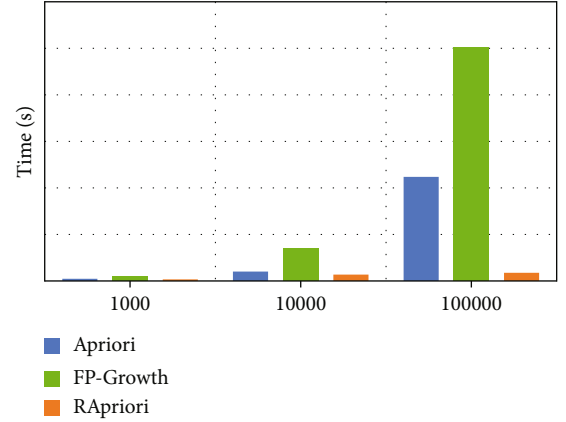


FIGURE 12: Time costs with different dataset.

TABLE 4: The comparison in accuracy of the five classifiers.

Classifier	Acc	Precision	Recall	Macro-F1
T-BiLSTM	0.88	0.90	0.85	0.88
BiLSTM	0.87	0.88	0.86	0.87
ME	0.81	0.82	0.79	0.81
Bayes	0.84	0.81	0.89	0.84
KNN	0.73	0.67	0.90	0.72
SVM	0.80	0.79	0.82	0.80

TABLE 5: Indicator evaluation for the three cases.

Indicators	Impact factors	Case 1	Case 2	Case 3
C1.1	VL	VL	FH	FH
C1.2	H	AL	AL	AL
C1.3	H	M	L	VL
C1.4	M	AL	AH	AH
C2.1	H	AH	AH	AL
C2.2	M	AH	AH	AL
C3.1	FH	AL	AH	AL
C3.2	M	AL	AL	AL
C3.3	L	AL	AL	AL

TABLE 6: The results of fuzzy similarities.

Risk level	Case 1	Case 2	Case 3
Absolutely low	0.4970	0.3506	0.5813
Very low	0.5229	0.3412	0.6386
Low	0.6917	0.4318	0.8592
Fairly low	0.9146	0.5872	0.7900
Medium	0.6866	0.7997	0.5561
Fairly high	0.4700	0.7347	0.3880
High	0.3895	0.5955	0.3246
Very high	0.3185	0.4770	0.2712
Absolutely high	0.3108	0.4881	0.2616

corresponding fuzzy numbers. Then, the impact factors and evaluations of the indicators are aggregated into a fuzzy number for each case. Finally, we compute the fuzzy number similarities to figure out the risk level.

Table 6 lists the similarities. Therefore, the POR of case 1 is fairly low, the POR of case 2 is medium, and the POR of case 3 is low. Combined with Table 5, case 3 has the least heat. Meanwhile, the judge and the parties are not unique identities. Even though the case type is at high risk, without hot discussion, the POR is low. As for case 1, the public opinion is quite positive, so the risk assessment result is “Fairly-Low”. Referring to case 2, one of the related parties has unique identities, and he has attracted much attention on social media. Nevertheless, media coverage is low, which illustrates that the issue has not been widespread yet. As we can see, the RAPOT recognizes the risk of POR successfully and distinguishes the three cases in risk measurement. To validate our framework’s efficiency, we compare five similarity measure algorithms. As we can see in Figure 9, the selected method’s output is the same as the majorities without outlier.

5. Conclusion

The accurate and fine-grained risk assessment on public opinion in the trial procedure is crucial for refined trial management. Our framework proposed in this paper provides an objective and efficient assessment for POR in the trial without using a large amount of historical data, which is quite lacking, and we propose T-BiLSTM to analyze public sentiment opinion based on topics. The method is more comprehensive than traditional BiLSTM in practice. The risk assessment framework for POR consists of three modules: (1) an adaptive multifactor indicator model for POR assessment, (2) the indicator evaluation module with an accurate public opinion analysis, and (3) the objective risk ranking module. The experimental results show the efficiency and practicability of our framework. In the future, we will work hard on the considerable amount of processing logs in the TPMS to further improve our indicator model’s adaptation and robustness.

Data Availability

The dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors gratefully acknowledge the support of the National Key R&D Program of China under grant No. 2018YFC0830500.

References

- [1] M. Bilal, H. Israr, M. Shahid, and A. Khan, “Sentiment classification of Roman-Urdu opinions using naive Bayesian, decision tree and KNN classification techniques,” *Journal of King Saud University - Computer and Information Sciences archive*, vol. 28, no. 3, pp. 330–344, 2016.
- [2] H. Htet, S. S. Khaing, and Y. M. Yi, “Tweets sentiment analysis for healthcare on big data processing and IoT architecture using maximum entropy classifier,” in *International Conference on Big Data Analysis and Deep Learning Applications*, pp. 28–38, Singapore, 2019.
- [3] J. Prusa, T. M. Khoshgoftaar, and D. J. Dittman, “Using ensemble learners to improve classifier performance on tweet sentiment data,” in *2015 IEEE International Conference on Information Reuse and Integration*, pp. 252–257, San Francisco, CA, USA, 2015.
- [4] S. Shubha and P. Suresh, “An efficient machine learning bayes sentiment classification method based on review comments,” in *2017 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC)*, pp. 1–6, Bangalore, India, 2017.
- [5] H. Kim and Y. S. Jeong, “Sentiment classification using convolutional neural networks,” *Applied Sciences*, vol. 9, no. 11, 2019.
- [6] S. Lai, X. Liheng, K. Liu, and J. Zhao, “Recurrent convolutional neural networks for text classification,” in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, AAAI’15*, pp. 2267–2273, Austin, Texas, USA, 2015.
- [7] M. Yang, T. Wenting, J. Wang, X. Fei, and X. Chen, “Attention-based lstm for target-dependent sentiment classification,” in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, AAAI’17*, pp. 5013–5014, San Francisco, California, USA, 2017.
- [8] H. Kanj and P. E. Abi-Char, “A new fuzzy-TOPSIS based risk decision making framework for dangerous good transportation,” in *2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/Smart-City/DSS)*, Zhangjiajie, China, 2019.
- [9] P. Singh, N. K. Mishra, M. Kumar, S. Saxena, and V. Singh, “Risk analysis of flood disaster based on similarity measures in picture fuzzy environment,” *Afrika Matematika*, vol. 29, no. 7–8, pp. 1019–1038, 2018.
- [10] R. Chutia and M. K. Gogoi, “Fuzzy risk analysis in poultry farming using a new similarity measure on generalized fuzzy numbers,” *Computers & Industrial Engineering*, vol. 115, pp. 543–558, 2018.
- [11] S. Kabir, C. Wagner, T. C. Havens, and D. T. Anderson, “A bidirectional subsethood based similarity measure for fuzzy sets,” in *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–7, Rio de Janeiro, 2018.
- [12] S. Cheng, S. Chen, and T. Lan, “A new similarity measure between intuitionistic fuzzy sets for pattern recognition based on the centroid points of transformed fuzzy numbers,” in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2244–2249, Hong Kong, China, 2015.
- [13] Y. Zhang, X. Yin, and Z. Mao, “Study on risk assessment of pharmaceutical distribution supply chain with bipolar fuzzy information,” *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 2, pp. 2009–2017, 2019.

- [14] A. Karasan, E. Ilbahar, S. Cebi, and C. Kahraman, "A new risk assessment approach: safety and critical effect analysis (SCEA) and its extension with pythagorean fuzzy sets," *Safety Science*, vol. 108, pp. 173–187, 2018.
- [15] M. Zaky, "Risk analysis using fuzzy system based risk matrix methodology," *Arab Journal of Nuclear Sciences and Applications*, vol. 51, no. 4, pp. 204–212, 2018.
- [16] G. F. Can and P. Toktas, "A novel fuzzy risk matrix based risk assessment approach," *Kybernetes*, vol. 47, no. 9, pp. 1721–1751, 2018.
- [17] T. Luo, C. Wu, and L. Duan, "Fishbone diagram and risk matrix analysis method and its application in safety assessment of natural gas spherical tank," *Journal of Cleaner Production*, vol. 174, pp. 296–304, 2018.
- [18] R. R. Yager, "Pythagorean membership grades in multicriteria decision making," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 4, pp. 958–965, 2014.
- [19] H. Ni, A. Chen, and N. Chen, "Some extensions on risk matrix approach," *Safety Science*, vol. 48, no. 10, pp. 1269–1278, 2010.
- [20] H. A. Khorshidi and S. Nikfalazar, "An improved similarity measure for generalized fuzzy numbers and its application to fuzzy risk analysis," *Applied Soft Computing*, vol. 52, pp. 478–486, 2017.
- [21] J. Li and W. Zeng, "Fuzzy risk analysis based on the similarity measure of generalized trapezoidal fuzzy numbers," *Journal of Intelligent and Fuzzy Systems*, vol. 32, no. 3, pp. 1673–1683, 2017.
- [22] Y. Qi, W. Jiang, and N. Liu, "Trial risk analysis based on a novel similarity measure on generalized fuzzy numbers," in *Proceedings of the 2020 4th International Conference on Management Engineering, Software Engineering and Service Sciences*, pp. 157–163, Wuhan, China, 2020.
- [23] S. A. Ahmad, D. Mohamad, N. H. Sulaiman, J. M. Shariff, and K. Abdullah, "A distance and set theoretic-based similarity measure for generalized trapezoidal fuzzy numbers," in *AIP Conference Proceedings*, vol. 1974, pp. 020–043, AIP Publishing LLC, 2018.
- [24] Z. Xu, S. Shang, W. Qian, and W. Shu, "A method for fuzzy risk analysis based on the new similarity of trapezoidal fuzzy numbers," *Expert Systems with Applications*, vol. 37, no. 3, pp. 1920–1927, 2010.
- [25] D. Yong, S. Wenkang, D. Feng, and L. Qi, "A new similarity measure of generalized fuzzy numbers and its application to pattern recognition," *Pattern Recognition Letters*, vol. 25, no. 8, pp. 875–883, 2004.
- [26] K. J. Schmucker, *Fuzzy Sets, Natural Language Computations, and Risk Analysis*, vol. 27, no. 3, 1984 Computer Science Press, 1984.
- [27] A. Darko, A. P. C. Chan, E. E. Ameyaw, E. K. Owusu, P. Ā. Erika, and D. J. Edwards, "Review of application of analytic hierarchy process (AHP) in construction," *International Journal of Construction Management*, vol. 19, no. 5, pp. 436–452, 2019.
- [28] G. Tian, M. Zhou, H. Zhang, and H. Jia, "An integrated AHP and VIKOR approach to evaluating green design alternatives," in *2016 IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC)*, pp. 1–6, Mexico City, Mexico, 2016.
- [29] A. Farkas, "Multi-criteria comparison of bridge designs," *Acta Polytechnica Hungarica*, vol. 8, p. 173, 2011.
- [30] S. A. K. Muhammad, A. Ali, S. Abdullah, F. Amin, and F. Hussain, "New extension of TOPSIS method based on pythagorean hesitant fuzzy sets with incomplete weight information," *Journal of Intelligent and Fuzzy Systems*, vol. 35, no. 5, pp. 5435–5448, 2018.
- [31] N. L. H. Mo and Y. Qi, "Trial risk index model and assessment system based on extended TOPSIS method," in *2020 International Conference on Data Intelligence and Security*, South Padre Island, TX, USA, 2020.
- [32] H. Tian and L. Wu, "Microblog emotional analysis based on TF-IWF weighted Word2vec model," in *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS)*, pp. 893–896, Beijing, China, 2018.
- [33] H. Bo, Y. Yang, A. Mahmood, and W. Hongjun, "Microblog topic detection based on LDA model and single-pass clustering," in *International Conference on Rough Sets and Current Trends in Computing*, pp. 166–171, Berlin, Heidelberg, 2012.
- [34] T. Chen, R. Xu, Y. He, and X. Wang, "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN," *Expert Systems with Applications*, vol. 72, pp. 221–230, 2016.