IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 13, 2020, accepted April 27, 2020, date of publication April 29, 2020, date of current version May 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2991305

A Personalized and Practical Method for Analyzing the Risk of Chemical Terrorist Attacks

RONGCHEN ZHU, XIAOFENG HU, XIN LI, AND HAN YE
School of Information Technology and Network Security, People's Public Security University of China, Beijing 102628, China

Corresponding author: Xiaofeng Hu (huxiaofeng@ppsuc.edu.cn)

This work was funded by the National Key Research and Development Program of China under Grant 2018YFC0809702 and the National Natural Science Foundation of China under Grant 71704183.

ABSTRACT The chemical terrorist attack is a type of unconventional terrorism that threatens the safety of cities. This kind of terrorist attack is highly concealed and difficult to be detected. Once the attack is successful, the consequences will be severe and the scope of impact will be enormous. Therefore, public security and emergency departments need to perform risk analysis and dynamic knowledge update to reduce risk or mitigate the effects of accidents. In order to quickly and effectively analyze the risk of chemical terrorist attacks, this article proposed a hybrid approach (B-R model) to analyze the risk of chemical terrorist attacks. First, a modular and customizable Bayesian network (BN) model library was built, which can satisfy users to select multi-dimensional risk factors. Based on the personalized BN, a risk knowledge graph (RKG) is constructed with multi-source data to realize the combination of risk analysis and knowledge acquisition. Then the threat degree of terrorist organizations, the strength of defensive forces, and the risk value of targets is calculated and displayed. The BN-RKG method provides data and theoretical support for defenders' resource allocation and emergency decision-making. Finally, a case study was conducted for a hypothetical scenario analysis. The result shows that the hybrid method can help with risk control and have the potential to support practical policymaking.

INDEX TERMS Chemical terrorist attack, Bayesian network, knowledge graph, risk analysis.

I. INTRODUCTION

Chemical terrorism, involving the use of toxic drugs, is intended to cause large numbers of casualties and could overwhelm the capacity of regional emergency medical services [1], [2]. Between 1970 and 2015, GTD recorded 156,772 terrorist incidents, of which 292 (0.19%) met the criteria as chemical terrorist attacks [3]. Although the proportion is small, the threat of chemical terrorism is reported to be increasing globally [4]. Examples are Sulphur mustard used by the Da'ish terrorist group [5] and the Tokyo subway sarin attack in 1995 [6]. To effectively respond to chemical terrorist attacks, people need to develop comprehensive strategies that include emergency response, long-term health care, risk communication, and other research [7].

Many pieces of research on chemical terrorism have been carried out on chemical weapon types [3], [8], emergency response [9], [10] and countermeasures [11]–[13]. Different approaches to analyze patterns and relationships of terrorist activity has been conducted by [14]–[19]. The above research not only provides the basis for the characteristics of chemical weapons and the harm of chemical attacks but also enriches emergency decision-making and prevention strategies for chemical terrorist attacks.

Extensive research on counter-terrorism strategies based on big data analysis has gradually received attention [20], [21]. Boyd used historical data in the Global Terrorism Database to study the frequency of attacks by 224 terrorist organizations against their own countries and other countries, and provided a basis for the development of international terrorism prevention strategies [22]. Kaur proposed a research framework for the analysis and prediction of terrorist activities using real-time data such as Facebook, Twitter, and Google, providing a complete example for the research of anti-terrorism strategies based on big data [23]. These traditional database-based methods have some limitations: poor visualization of knowledge, difficulty in showing the connections between data, and inability to

The associate editor coordinating the review of this manuscript and approving it for publication was Malik Jahan.

reason about knowledge. Furthermore, the traditional method does not work well for multi-source data fusion. Usually, each database is stored separately and lacks connections. In this article, the knowledge graph is utilized to solve this problem.

The essence of the knowledge graph is to establish the relationship between knowledge. The "entity-relation-entity" semantic data structure describes the concepts, entities and their relationship in the objective world [24], [25]. Knowledge graph originated from the Semantic Network. It was originally proposed by Google [26] and used to optimize search results [27] and has been applied to various vertical fields so far [24], [25], [28], [29]. The domain knowledge graph can be regarded as an "industrial knowledge base based on semantic technology." Its construction is based on industry data and usually has strict and rich data models [30]. Several different research methods on the knowledge graph of terrorist attacks have been proposed. Jha and Jin proposed a method based on knowledge graphs to discover potentially high-value hidden information under massive corpora and used it in the research of counter-terrorism big data analysis and counter-terrorism decision-making [31]. Xia and Gu built a terrorist knowledge graph (TKG) from GTD and Wikipedia. Compared to GTD, TKG strengthens links between terrorist organizations and enriches the description by absorbing Wikipedia's data. TKG can better help humans and machines understand terrorist attacks [32]. However, these knowledge graphs are not combined with risk analysis. Often, users are trapped in a large amount of data, so that unable to make effective decisions.

The combination of knowledge graph and risk analysis technology can play a significant role in practical application. Risk analysis of terrorist attacks using qualitative or quantitative methods appears in several cases in the literature, such as Hazard and operability study [33], Layer of protection analysis (LOPA) [34], Event tree analysis (ETA) [35], Fault tree analysis (FTA) [36], Fuzzy set theory [37], Markov chain model [38]. However, we choose the Bayesian network (BN) as the risk analysis tool. Bayesian network is a directed acyclic graph (DAG), which describes the relationship between nodes (a set of variables) in the form of directed edges with a conditional probability distribution [39]. The nodes in BN are discrete variables, and the node contains many possible states. Each child node has a set of parent nodes [32]. The relationship between a child node and its parents is expressed as a directed edge with a conditional probability table (CPT). CPT is the confidence (expressed as a probability) that a node will be in a specific state given the state of the parent node (Figure 1). Compared with other risk analysis methods, Bayesian network has some advantages: (1) BN can answer hypothetical questions. For example, users can enter defensive measures and evaluate which measures can reduce the risk of terrorist attacks. (2) BN can fuse multi-source information. Various risk factors for terrorist attacks can be effectively considered. (3) Users can easily add or reduce terrorist attack risk factors, or modify the conditional probability table. (4) As shown in Figure 1, Both Bayesian network and Knowledge graph are network structures and composed of nodes and relationships. Furthermore, nodes or values in both can be reused. (5) BN can obtain the risk probability distribution, which can be mapped into the knowledge graph to assign risk values to nodes. (6) BN can handle non-linear relationships and store non-numeric states of nodes. Such states can be represented as instances in the knowledge graph.

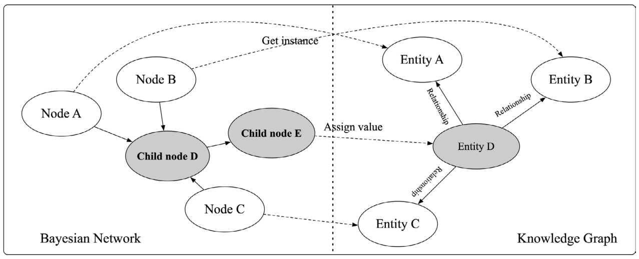


FIGURE 1. The relation between Bayesian networks and knowledge graph. Bayesian network and knowledge graph share nodes and influence each other.

Many scholars have made useful attempts in the field of terrorist attacks with the Bayesian network. For example, Wei *et al.* proposed a multi-module Bayesian network terrorist attack threat assessment model and gave a calculation method of the threat degree of terrorist attacks [40]. Fu *et al.* built a Bayesian network model using terrorist attack samples of other countries. Then, based on the principle of case suitability, combined with China's actual data, the EM (Expectation-Maximum) algorithm is used to update the parameter learning and modify the model [41]. Olama *et al.* proposed a BN based on terrorist threat anticipatory model, which takes the physical, social and economic aspects into consideration [42]. However, the risk factors considered in these studies are not comprehensive. In particular, Zhu *et al.* established a Bayesian network for chemical terrorist attacks for risk analysis [43], conducting attacking risk analysis from multiple dimensions including terrorist organizations, target attractiveness, defensive forces, emergency response forces, and danger level of the weapon. Nevertheless, this model is not able to obtain specific attack risk indicators or values, which further limits the practicality of the risk analysis.

Through the review above, the contributions of this article are as follows: (1) This article proposes a hybrid approach called "Bayesian Network – Risk knowledge graph" model (B-R model). Bayesian network is used to represent the critical nodes of the attack and assign risk value; the risk knowledge graph plays a critical role in summarizing domain knowledge, fusing multiple source data and showing dynamic knowledge visually. (2) An interactive, customizable modular Bayesian network library was developed to discover and access the risk of chemical terrorist attacks actively. With our method, users can perform chemical terrorist attack risk analysis faster and more effectively. The results of the case study and scenario analysis prove the effectiveness of the proposed method.

The structure of this article is as follows. Section 2 details the materials and process of our method. Section 3 and 4 describe the detailed construction process of Bayesian network model library and risk knowledge graph. Section 5 conducts a case study and illustrates the applications that demonstrate the potential of our approach. Some limitations are discussed in Section 6. Finally, Section 7 summarizes the article.

II. METHOD AND MATERIALS

The framework of personalized terrorist attack risk analysis model (B-R model) includes four phases: (1) Search data sources; (2) Construction of BN risk analysis model; (3) Construction of risk knowledge graph; (4) Risk analysis and application (see Figure 2). Phase 2 and 3 are the most critical. Users can choose different algorithms to generate customized BN models. Only the phase 2 is the user run-time, and the other phases are the build-time. Both the BN model library and terrorist attack data need to be updated dynamically. The library continuously collects the risk factors, BN models. Moreover, data sets that users have created or used are stores in the library after expert review. The news and reports are automatically collected from the website by the crawler script.

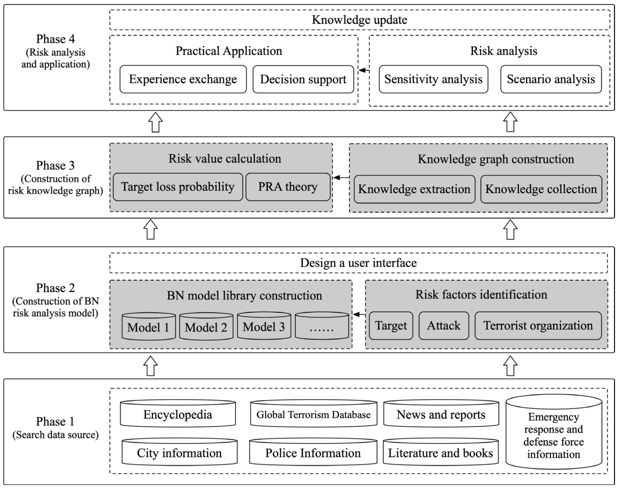


FIGURE 2. **The framework of Bayesian network–risk knowledge graph (B-R) model.**

TABLE 1. **Data source and classification.**

Data source	Type of data	From
GTD	Terrorist organization, Attack information	Website
Encyclopedia	Terrorist organization, Attack information, Chemical weapon	Website
News and reports	Terrorist organization, Attack information	Website
City information	Target, Chemical plant, Weather condition	Manual collection
Emergency response and defense force information	Police station, Fire brigade, Hospital	Manual collection
Police Information	Key person, Dynamic track of key person	Police Information System

As shown in Table 1, most of our data sources are publicly available except for police information. Different departments can use specific data as needed. Data sources for this article include (1) Global Terrorism Database (GTD) terrorist attack data. GTD [44] is an open database containing information on global terrorist attacks from 1970 to 2018, and currently has more than 190,000 cases. Each case provides information on the date, location, type of chemical weapon, number of casualties, and terrorist organizations. We screen a total of 336 chemical terrorist attack cases as the core data set and the remaining terrorist attack cases as the supplementary data set. (2) Wikipedia and Baidu Encyclopedia. Keywords are collected and searched in encyclopedias, such as chemical weapons: mustard gas, sarin gas; attacks: Japanese subway sarin gas, Matsumoto sarin incident; terrorist organization: ISIS, Taliban; Based on a total of 50 keywords, we use the breadth-first search algorithm to obtain more encyclopedia information. (3) News and reports. Web crawlers are used to crawl reports and news of terrorist attacks on news sites such as China Caixin.com. (4) City information. For any specific city, information about chemical plants is considered because terrorist organizations may steal chemical raw materials. Also, this article identifies hotspot targets in cities through the Bayesian network and manually collect objective conditions such as weather and traffic conditions *et al.* (5) Emergency response and defense force information include some basic information of police station, fire brigade and hospital in the city. (6) Police information includes key person information and its dynamic track, which was recorded by surveillance, hotel and traffic information system.

III. BN RISK ANALYSIS MODEL CONSTRUCTION

First, we search article with specific keywords on the web of science. Then the retrieved articles are screened based on three criteria. The final articles are used as the source of risk factors and BN networks. The keywords include (1) Bayesian network + terrorist attack (2) Bayesian network + terrorism (3) Bayesian network + terrorist. Criteria include: (1) Only journal articles or conference articles, not patents (2) Articles need to be written in English (2) Articles contain the structure of Bayesian networks.

A. RISK FACTORS IDENTIFICATION

The scope of risk factors needs to be defined first. Faced with different risk factors input by users, it is necessary to make a proper classification and determine the scope before accurate risk analysis. Other risk factors related to terrorist attacks are obtained from books and terrorist attack risk analysis articles [3], [40]–[43], [45]. For example, Khakzad *et al.* [46] listed some information which needs to be considered from the perspective of terrorist organization: (1) the general history of threats and attacks against similar

TABLE 2. Part of risk factors.

Emergency Response	Reference	Terrorist organization	Reference	Terrorist attack	Reference
Police emergency response	[43]	Religious background	[43]	Country	[40]
Fire emergency response	[43]	Region of the perpetrator	[40, 43]	Whether the attack is successful	[40, 41, 43]
Hospital emergency response	[43]	Number of members	[43]	Casualties	[41-43, 45]
Target emergency response	[43]	Average educational level	[43]	Property loss	[40-42]
Climate	Reference	Technical background	[43]	Economic impact	[42]
Wind direction	[43, 45]	Social relations and organizational components	[43]	Psychological influence	[42]
Wind speed	[43, 45]	Whether they have been reported recently	[43]	Number of people involved in crime	[40, 41]
Precipitation	[43]	Whether they ever launched terrorist attack	[43]	Attack type	[40, 41, 43]
Target	Reference	Whether they made a statement or threat	[43]	Whether deliberately expand the impact	[40]
Target type	[40]	Source of weapon	[43]	Whether political or economic purpose	[40, 41]
Traffic condition	[43]	Whether they have technical support	[43]	Whether ask for ransom	[40]
Population density	[43, 45]	Whether they have the capabilities of storing and transporting the weapon	[43]	Whether it is a suicide attack	[40, 41]
Population movements	[43]	Whether they have the ability to launch the attack	[43]	Delivery method	[43]
Symbolic Value	[42]	Has a Charismatic Leader	[42]	Weapon type	[40, 42]
Whether it is a high-value target	[43]	Ingroup/Outgroup Sentiments	[42]	Time	[45]
Location	[42, 45]	Collectivist Ideology/ Loyalty to Regime	[42]	Size of weapon	[43, 45]
Defensive power	Reference	Individual	Reference	Individual	Reference
Security check	[43, 45]	Age	[42]	Close Family Ties	[42]
Patrol	[43, 45]	Gender	[42]	Educational level	[42]
Surveillance	[43, 45]	Occupation	[42]	Ever used chemical weapons	[42]
Police investigation	[43, 45]	Religious background	[42]	Recidivism	[42]

targets, (2) location-specific attack records, (3) attacker's capabilities and potential behaviours, (4) the attractiveness of chemical facilities in the eyes of attackers. To model chemical terrorist attacks from a comprehensive perspective, In this study, risk factors are classified into seven classes: Terrorist organization, Terrorist attack, Target, Individual, Climate, Defensive power, Emergency Response (as seen in Table 2).

B. BN MODEL LIBRARY CONSTRUCTION

It is worth noting that users' needs for risk models in different scenarios are different. For example, one user may only want to study the relationship between defence forces and casualties. If users want to analyze the risk of a specific suspect, they often need to make judgments based on the suspect's portrait and past experience. Therefore, we need to provide users with a risk analysis model that can be customized rather than a fixed model. We developed an interaction interface, in which users can set risk factors according to the actual situation, design the network structure and upload data sets for training. At the same time, according to the needs of users, BN library will recommend models to users. The library was built by different structure learning and parameter learning algorithms. There are many classic Bayesian network learning algorithms, which can be roughly divided into three categories: score-based search methods, dependency analysis-based methods and hybrid learning methods. The algorithms include (1) K2 algorithm. The algorithm achieves the purpose of finding the network topology with the best scoring function under the conditions of the given node order. The K2 algorithm can effectively incorporate prior knowledge in its structure search process and have excellent time performance. It is a practical and most representative score search learning algorithm. (2) Simulated annealing method. The simulated annealing algorithm adopts an entirely random search strategy. When the temperature is high, the algorithm can accept the solution with the low value of the partial scoring function to avoid falling into the local optimization. Among the parameter learning algorithms, some are well-known: (1) EM algorithm. The EM algorithm is an iterative optimization strategy because each iteration in its calculation method is divided into two steps, one is the expected step (E step), the other is the maximum step (M step). Therefore,

TABLE 3. Information about some terrorist attack BN models.

Model id	Model name	Number of nodes	Model content	Structural learning	Probabilistic learning	Reference
Model 1	Fu	8	Terrorist attack warning model	K2	EM	[41]
Model 2	Wei	12	Terrorism threat assessment model	K2+Expert experience	Data+ Expert experience	[40]
Model 3	Zhu	42	Chemical terrorist attack model	Expert experience	EM + Expert experience	[43]
Model 4	Tang	13	Urban dirty bomb attack model	D-S theory	D-S theory	[45]
Model 5	Olama	66	Physical, social and economic BBN threat anticipation model	Expert experience	Expert experience	[42]
……	……	……	……	……	……	……

the algorithm is called the Expectation-Maximization Algorithm. The EM algorithm was originally designed to solve the problem of parameter estimation in the case of missing data. (2) MAP algorithm. Maximum posterior estimation is the point estimation of a quantity that is difficult to observe based on empirical data. It is similar to the maximum likelihood estimation. However, the maximum difference is that the maximum posterior estimation is integrated into the prior distribution of the estimator. The specific process is shown in Figure 3.

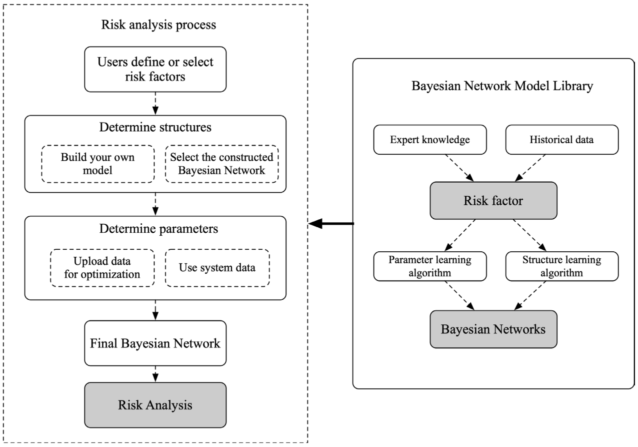


FIGURE 3. User conducts risk analysis process. The pre-stored risk factors and network models play a very important role in practical application.

Firstly, after identifying the risk factors, we manually sorted out the state and state characteristics of each factor. For example, the state of the season concludes "spring, summer, autumn and winter" and the data type of it is character. After that, various data was collected based on Section 2 and organized them into data sets. For example, the attack information was compiled from GTD. Terrorist organization information was collected from Wikipedia. When the user selects and inputs risk factors, if there exist no such risk factors in the library, similar risk factors and related data sets will be recommended according to the state space of the factors.

Secondly, we determine the one-to-one correspondence between risk factors and the BN model in the module library. If the risk factors are similar, then the BN network should be similar. The library will recommend the user with the BN model that contains the most risk factors. If the library does not include the risk factors entered by the user, the user needs to construct the BN model by himself. A brief introduction to the BN models is shown in Table 3 and Figure 4.

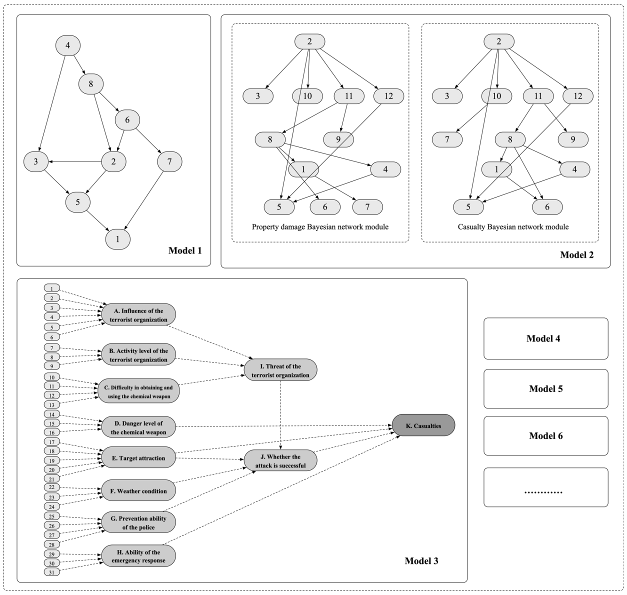

Model 1
Property damage Bayesian network module
Casualty Bayesian network module
Model 2
A. Influence of the terrorist organization
B. Activity level of the terrorist organization
C. Difficulty in obtaining and using the chemical weapon
D. Danger level of the chemical weapon
E. Target attraction
F. Weather condition
G. Prevention ability of the police
H. Ability of the emergency response
I. Threat of the terrorist organization
J. Whether the attack is successful
K. Casualties
Model 3
Model 4
Model 5
Model 6
…………

FIGURE 4. Some network structures in BN model library. Due to space limitations, we only show 3 network structures.

The BN model used by Fu has fewer nodes than Model 2 and Model 3. Fu used the K2 algorithm to learn the network structure from the GTD database. As a result, the explanation of the network structure is not very high. For example, in common sense, "Successful attack" will affect "casualties", but in model 1 the situation is opposite. Furthermore, when it comes to scenario analysis and forward reasoning, this structure is not as intuitive as Model 3. A significant feature of Model 2 is the module-based construction method, which is reflected in the construction of bn networks for property losses and casualties. This module-based construction method is advantageous and can better divide different node types. Users are also recommended to construct the network by the module. A hybrid method based on sample learning and the expert experience was utilized in Model 2, which is systematically adopted. The library stores part of the data, and users can choose to use the library data or their own data for network training and parameter learning. The main feature of Model 3 is a clear structure. For users, this construction method is simple and easy to use, as in most cases, users have a specific understanding of the risk factors they want to analyze. The state of the child nodes of Model 3 is simple. Three main types are described as high, middle, and low. This assignment method is easy to migrate. However, the disadvantage of Model 3 is the determination of a priori probability and conditional probability. Once there are more nodes, the conditional probability table will become large, and it is very natural to produce subjectivity based on the experience of only one person. Therefore, users need to balance the number of nodes, network interpretability and network complexity.

C. USER INTERFACE DESIGN

A visual interface is designed for users to select or build Bayesian networks (Figure 5). Users can add risk factors or submit the existing network model. The interface also provides several risk factors that users can refer to. After submitting risk factors, users can choose whether to provide data for network optimization.

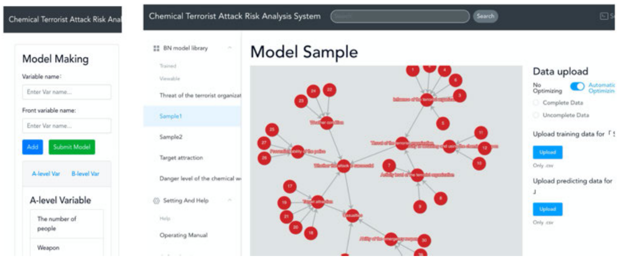


FIGURE 5. User interface.

IV. RISK KNOWLEDGE GRAPH CONSTRUCTION

The knowledge graph and Bayesian network can be combined to form a risk knowledge graph (Figure 6). The Bayesian network provides a calculation method of risk value for knowledge graph. The knowledge graph provides a knowledge base for Bayesian network risk analysis, which allows users to more intuitively assess the risk level and take measures to promote risk evolution.

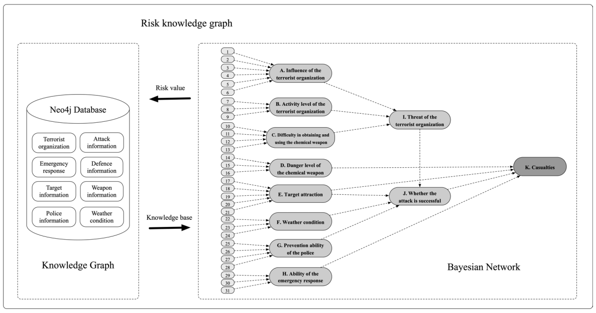


FIGURE 6. The internal mechanism of risk knowledge graph.

A. KNOWLEDGE GRAPH CONSTRUCTION

The process of knowledge graph construction is mainly divided into knowledge representation and modeling, knowledge acquisition, knowledge extraction, knowledge reasoning, knowledge storage, and application of knowledge graph [28], [30]. The tools involved in the construction are Scrapy crawlers (Data collection), Chinese Academy of Sciences Natural language processing (NLP) tool [47] (Knowledge extraction), Neo4j [48] and MongoDB database (Knowledge storage), Elasticsearch search (Knowledge retrieval).

The structure of the knowledge graph can provide guidance for knowledge extraction. Through literature review and historical data collation, 11 chemical terrorism categories are summarized in Figure 7, which includes Target, Police station, Chemical plant, Fire brigade, Hospital, Climate, Terrorist organization, Attack information, Key person, Dynamic track of key person, Chemical weapon. The white part of Figure 7 indicates the attributes included in the knowledge graph, and the green part means the existing risk factors. As a result, knowledge graph and Bayesian network can effectively share nodes. It should be noted that each class has the id attribute so that different categories of knowledge graph can be connected.

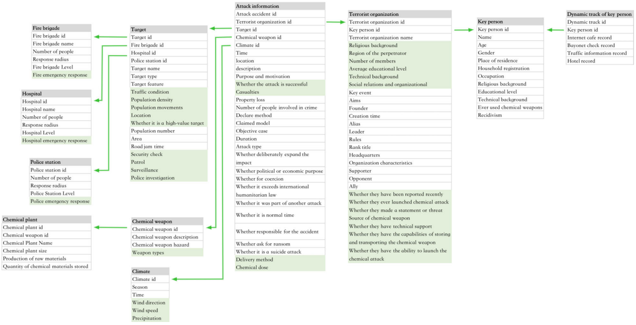


FIGURE 7. Knowledge structure of knowledge graph.

Then, based on the knowledge structure and the data collected in section 2, ‘‘Entity-Relationship-Entity’’ triples need to be extracted to construct a knowledge graph. Among the data source, structured data like encyclopedia entries can be easily imported to the graph database. However, unstructured data, such as text and news reports, contained some properties which need to be extracted. Take a sentence as an example: ‘‘On June 7, 2017, ISIS assailants exploded four mustard gas bombs at civilians in Zanjili neighbourhood, Mosul, Iraq. At least 13 people were injured in the attack.’’ Since NLP tool can identify the date, the number, and the part-of-speech of each word, it is utilized to extract the triples in the sentence. For example, ‘‘ISIS’’ is a noun, and ‘‘explode’’ is a verb. After determining the part of speech and completing the word segmentation, we match them in the thesaurus of each attribute as Table 4 shown. Finally, the triples are stored as CSV documents. It should be noted that the descriptions of most cases are irregular, which need to be cleaned up

TABLE 4. An example of knowledge extraction.

Attribute	Time	Location	Terrorist organization
Value	06/07/ 2017	Mosul, Iraq	ISIS
Chemical weapon	Delivery method	Target	Casualties
mustard gas	explode	neighborhood	13 injured

and screened manually. The triples can be imported from CSV to Neo4j using several simple scripting languages. Labels and indexes are created in Neo4j to increase search efficiency.

B. CALCULATION OF RISK VALUE

In this section, three kinds of values are calculated. The first is the target's attraction value, defensive force level value, emergency force response value, and terrorist organization threat value. Such values can help users understand potential risks and gather information. The second is the target loss probability [49], which represents the probability of the target being attacked. The police can flexibly deploy defence forces based on this value. The third is the target's risk value. This risk value reflects the expected consequences against which the benefit of existing or potential terrorism strategies can be evaluated and estimated [50].

1) ASSIGN VALUES TO KEY NODES IN KNOWLEDGE GRAPH

Bayesian network is used to assign values to the knowledge graph. For example, the Bayesian network contains nodes "Danger level of the chemical weapon," "Target attraction," "Prevention ability of the police," "Ability of the emergency response," and "Threat of the terrorist organization," and the status values of these nodes are all described as "High, Medium, and Low" (Appendix A, Table 10). Therefore, when the prior knowledge is introduced into the Bayesian network, we can assign values to each state, and then get the specific value of each node.

Use "Target attraction" as an example. We first select the candidate targets, and then collect the "population density", "population movement", "traffic situation", "location", "whether is it a high-value target", a total of five attributes of all candidate targets as input to the Bayesian network, and get the probability distribution of the "Target attraction" node. In Bayesian network, the node "target attraction" has three states: "High, Medium, and Low", so we assign high = 5, medium = 3, and low = 1, and use (1) to calculate the target attraction value.

$$Target\ attraction = 5 \times P_{High} + 3 \times P_{Medium} + 1 \times P_{Low} \tag{1}$$

P_{High} is the probability that the status of "Target Attraction" is "High". If the value of "Target attraction" is bigger than 4, the target is attractive, and its detailed information needs to be collected in the knowledge graph.

2) CALCULATE TARGET LOSS PROBABILITY

For a specific target, if a terrorist launches an attack with certain resources, the defender will allocate resources to defend the target [51], [52]. In the situation of the game between the attacker and the defender, the target loss probability can be determined, which can be reasonably explained by the target loss probability model proposed by Major [49]:

$$p_i = exp\left(-\frac{d_i}{c_i A_i}\right) \times \frac{A_i^2}{A_i^2 + c_i} \quad (i = 1, 2, 3, \ldots, n) \tag{2}$$

In (2), di is the target's defence resources, and we use the value of node "Prevention ability of the police" in Bayesian network to represent di; ci is the inherent risk loss of the target, which is the value of target attraction; Ai is the attack resource, we assume the attack resource is 5, the value has no effect on target loss probability.

Furthermore, without knowing which target the attacker chooses, the defender should formulate a strategy to get the least expected loss of balance. No matter which targets the attacker chooses, the balance expected loss (EL) of the targets need to be equal.

$$EL = P \times V \tag{3}$$

In (3), P is the target loss probability, and V is the target value. If the EL of one target is higher than other targets, the defender will transfer defence resources from the lower EL target to the higher EL target until they have equal EL. The target loss probability is stored as target attributes in the knowledge graph so that government officials can adjust resources to reduce losses.

3) DETERMINE RISK VALUES FOR DIFFERENT TARGETS

Probabilistic risk analysis (PRA) [50] is utilized to get the risk value of different targets. Probabilistic risk analysis is a standard method to study the risk of terrorist attacks on infrastructure. PRA holds that:

$$Risk = Threat \times Vulnerability \times Consequence \tag{4}$$

The threat refers to the probability of a particular attack; the vulnerability refers to the probability of the success of the attack; the Consequence refers to the losses caused by a successful attack, including human casualties and economic losses. Target loss probability is used to represent the threat. Vulnerability is the probability of node "J. Whether the attack is successful" in the Bayesian network, and Consequence is the probability distribution of the "K. Casualties" in the Bayesian network multiplied by the assignment. The equation used to assess casualties is as follows:

$$Casualties = K_1{*}P_{Minor} + K_2{*}P_{Middle} + K_3 * P_{Major} \tag{5}$$

As shown in Appendix A, Table 10, K_1, K_2, and K_3 represent the status of the node "Casualties". P_{Minor}, P_{Middle} and P_{Major} represent the corresponding probability of each attribute of "Casualties" in the Bayesian network. Equation (5) can be used to calculate the worst, best and average results of the attack. We use the maximum value of each state to calculate the worst result and the median to calculate the average, and the minimum to calculate the minimum casualties. For example, in the K_{Minor} (0-10 dead) state, the maximum value is ten, and the average value is 5.

V. PRACTICAL APPLICATIONS

Although chemical terrorist attacks are difficult to prevent, they also have weaknesses. The manufacturing, storage, transportation, and release of chemical weapons require high technical capabilities of terrorist organizations and usually require the participation of multiple people. The chain of behaviour is relatively long, so it is easier to leave "Clues," and the capture of these clues depends on the support of multi-source data. Traditional clue-capturing methods cannot minimize these risks because the data is fragmented and lacks useful connections. In this section, the effectiveness and practicability of this method are verified through case study and scenario analysis.

A. CASE STUDY

The location of our case study is HZ District, LYG City, China. HZ District is the political, economic, and cultural center of LYG City. Firstly, three targets (Suning shopping center, Municipal government, and Phoenix mountain park) with different characteristics in HZ District are selected as the research object of the chemical terrorist attack. Secondly, we select risk factors, most of which have appeared in [43]. The network parameters are partly derived from data learning (the GTD database that comes with the system) and partly from the expert experience conditional probability table (Figure 8). The specific states of each node are in Appendix A, Table 10.

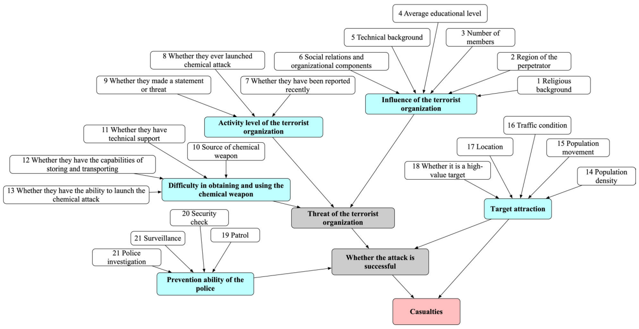

4 Average educational level
5 Technical background
3 Number of members
8 Whether they ever launched chemical attack
6 Social relations and organizational components
2 Region of the perpetrator
9 Whether they made a statement or threat
7 Whether they have been reported recently
1 Religious background
Influence of the terrorist organization
Activity level of the terrorist organization
16 Traffic condition
11 Whether they have technical support
10 Source of chemical weapon
17 Location
15 Population movement
12 Whether they have the capabilities of storing and transporting
18 Whether it is a high-value target
14 Population density
Difficulty in obtaining and using the chemical weapon
13 Whether they have the ability to launch the chemical attack
Target attraction
20 Security check
Threat of the terrorist organization
21 Surveillance
19 Patrol
21 Police investigation
Whether the attack is successful
Prevention ability of the police
Casualties

FIGURE 8. User-built Bayesian network.

Some objective information of each target is shown in Table 5.

According to (1), the values of the three targets attraction and defence ability are calculated and shown in Table 6. The city government has the strongest defence force, followed by shopping malls. The level of attraction of shopping malls and the municipal government is comparable. According to (1), the value of attractiveness ranges from 1-5. We take the top 25% of the attractiveness value as the high attractiveness target, that is, if the value is greater than 4, it is considered as a high-value target. The value of the park is 1.32. Since 1.32 is less than 4, the park is not a hot target and will not appear in the knowledge graph. Through this calculation method, the user can evaluate the attractiveness level from the objective state of the target and can reduce the dependence on subjective experience. Similarly, after basic information about terrorist organizations from databases or investigation is obtained, this method is used to calculate the threat degree of terrorist organizations. According to the objective state obtained in Table 5, the defensive ability is substituted into the Bayesian network and calculated by the same method. Table 6 shows that the city government's defence value is the strongest, reaching 4.94. The saturation of basic defensive levels reflects the government's emphasis on political objectives. The Suning shopping center's defence rating of 3.92 is above average. Shopping centers are crowded with people and have a certain amount of preventive capacity, but due to the large number of them, they cannot all carry out high-intensity protection. Moreover, high levels of protection can affect convenience, so a balance needs to be struck.

TABLE 5. Objective situation of three targets.

Parent Nodes	Suning shopping center	Municipal government	Phoenix mountain park
Population density	> 1000 / km2	500 ~ 1000 / km2	<500 / km2
Population movement	High	Medium	Low
Traffic condition	Good	Good	Bad
Location	Commercial area	Residential area	Open space
Whether it is a high-value target	Yes	Yes	No
Patrol	More than 2 times	More than 2 times	Less than 2 times
Security check	No	Yes	No
Surveillance	24 hours	24 hours	Non-24 hours
Police investigation	No	Yes	No

TABLE 6. Target loss probability for different targets.

Target	Defensive resource	Target attraction	Target loss probability
Suning shopping center	3.02	4.88	0.0433
Municipal government	4.94	4.78	0.0181
Phoenix Mountain Park	1.08	1.32	0.1684

The target loss probability is calculated according to (2). As shown in Table 6, the government has the lowest target loss probability. The results can be inferred from (2) that the higher the defensive resources, the lower the target loss probability. Furthermore, the impact of the degree of attraction on the target loss probability is relatively low.

We classify casualties as maximum, moderate and minimum. Maximum means that each of the state values takes the maximum, as shown in Table 7. Using this method to calculate the maximum and minimum casualties can provide users with more decision plans. According to (5), the maximum casualties of Suning shopping center are 22 deaths, the average is 14, and the minimal is 6. The number of casualties calculated is rounded. The results of casualties in Table 7 are in line with the actual situation. The shopping center is crowded with people. Once an attack occurs, the casualties will be huge. The city government is located in a residential area, which can also cause some casualties.

TABLE 7. **Calculation results of casualties.**

States of "Casualties"	Minor	Middle	Major	Max	Average	Min
Value	0–10	10–30	30-70	10-30-70	5-20-50	0-10-30
Suning shopping center	0.7	0.16	0.04	22 deaths	14 deaths	6 deaths
Municipal government	0.85	0.11	0.04	15 deaths	8 deaths	3 deaths
Phoenix Mountain Park	0.92	0.07	0.01	12 deaths	7 deaths	1 death

Through the above steps, risk value can be got through (4). Take moderate casualties as an example. Table 8 shows the PRA calculation for different targets. The government has the lowest risk value because it has the most defence resource and the target loss probability is the lowest. Although the park has the highest risk value and is the easiest to be attacked, in reality, terrorist organizations may be more likely to attack the government or shopping malls because they want to cause the most casualties or to achieve the purpose of threats.

TABLE 8. **PRA calculation for different targets.**

Target	Threat	Vulnerability	Consequence	Risk
Suning shopping center	0.0433	0.52	14	0.315
Municipal government	0.0181	0.11	8	0.015
Phoenix Mountain Park	0.1684	0.75	7	0.884

With the correlation values obtained, a risk knowledge graph is shown in Figure 9.

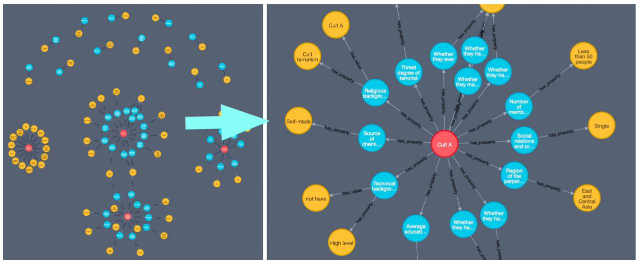


FIGURE 9. **Part of the risk knowledge graph.**

B. SCENARIO ANALYSIS

Decision-makers can use the risk value, the estimated casualties, the target loss probability to get more information in the knowledge graph to evaluate and make the best decision. The B-R method is used for critical risk point detection, early warning, emergency decision making and defense resource allocation.

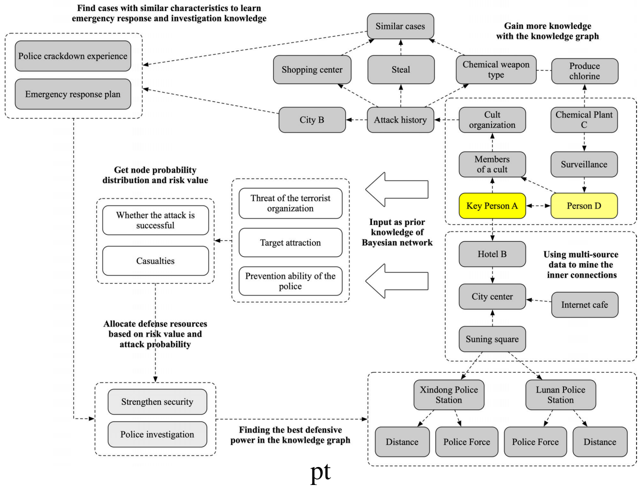
Scenario: Key Person A comes to City B to attempt for a chemical weapon attack.

First of all, through the hotel information system, the police get that A comes to city B and lives in a hotel in the city center. At the same time, person D is found in surveillance next to chemical plant C. Through the police information system query, we find that key person A and person D were members of a cult. The historical case shows the cult had stolen chemical materials in City E and lunched a chemical terrorist attack. So, person D is suspected to steal chemical materials, and A may complete the launch. At the same time, calculated by the portrait of the cult and Bayesian network, the threat degree of the terrorist organization is 2.98 (0-5), the value is relatively high and thus the cult should be paid attention to. Then from the knowledge graph, we obtain the members' information and discover that the organization has attacked the shopping center before. Moreover, Suning shopping center is within 1 km of the hotel where A lives in. The record of Internet cafe in Suning shopping center shows A used to come here, so we speculate that the shopping center is a possible target.

We bring the information from Suning shopping center into the Bayesian network and get the probability distribution of successful attacks and casualties, which shown in Table 6,7,8. The consequences are severe, so we consider how to reduce the risk of the attack. According to the sensitivity analysis of Bayesian Network, that security check is the most effective means of prevention. Through adding security check and police investigation, the probability of a successful attack reduces to less than 5%.

From the knowledge graph, the Lunan police station is the closest one with only 15 police officers, while the Xindong police station with 50 police officers is a little far away. Therefore, five police officers are dispatched from the Xindong police station to support the Lunan police station. When defense is strengthened, both the target's risk value and the target's probability of loss have decreased. Decision-makers can also further allocate resources based on the dynamic changes of the values in the knowledge graph.

The detailed analysis process of the case is shown in Figure 10. The grey nodes represent the data in the knowledge graph, while the white nodes are the probability distribution Bayesian network and the calculated risk index. The yellow node represents the initial event. As the cult has attacked the City E, the knowledge graph stores the experience of the investigation and emergency response measures in the previous attack, which can provide step-by-step guidance for prevention and help law enforcement agencies to organize similar emergency plans. Table 9 shows the part of the information which can be got from the knowledge graph. Due to space limitations, we have not listed all the information. In this scenario, the critical risk point is


Find cases with similar characteristics to learn emergency response and investigation knowledge
Similar cases
Gain more knowledge with the knowledge graph
Police crackdown experience
Emergency response plan
Shopping center
Steal
Chemical weapon type
Produce chlorine
City B
Attack history
Cult organization
Chemical Plant C
Members of a cult
Surveillance
Get node probability distribution and risk value
Key Person A
Person D
Threat of the terrorist organization
Whether the attack is successful
Input as prior knowledge of Bayesian network
Target attraction
Casualties
Hotel B
Using multi-source data to mine the inner connections
Prevention ability of the police
City center
Internet cafe
Allocate defense resources based on risk value and attack probability
Suning square
Xindong Police Station
Lunan Police Station
Strengthen security
Police investigation
Finding the best defensive power in the knowledge graph
Distance
Police Force
Police Force
Distance

pt

FIGURE 10. Scenario analysis flow chart.

TABLE 9. Information on key people A and organization B contained in the knowledge graph.

Organization information	Cult organization	Key person information	Key person A
Religious background	Cult terrorism	Name	XXX
Region of the perpetrator	East and Central Asia	Age	35
Number of members	Less than 50 people	Gender	Male
Average educational level	High	Place of residence	Cangwu Community
Technical background	High level	Household registration	Xuzhou City
Social relations and organizational components	Single	Occupation	Unemployed
Whether they have been reported recently	No	Religious background	Once believed in a cult
Whether they ever launched chemical attack	No	Educational level	High school education
Whether they made a statement or threat	No	Technical background	No
Source of chemical weapon	Self-made	Ever used chemical weapons	Yes
Whether they have technical support	Yes	Recidivism	Yes
Whether they have the capabilities of storing and transporting the chemical weapon	Yes	Internet cafe record	Place, time, duration
Whether they have the ability to launch the chemical attack	No	Hotel record	Place, time, duration
Threat degree of terrorist organizations (0-5)	2.98 (medium level)	Traffic information record	Place, time, vehicle

“Key person A appears in city B” and “Person D appears in the surveillance of chemical plant C.”

VI. DISCUSSION

The proposed risk analysis method combines Bayesian network and risk knowledge graph. We collect risk factors from a multi-dimensional perspective and develop an interface that users can customize the generation of BN. The interface facilitates users to add risk factors and customize a personalized terrorist attack risk analysis model. More importantly, the highly visual nature of the knowledge graph allows users to observe the level of risk intuitively. By combining the knowledge of multiple departments, such as the emergency response department, decision-makers can efficiently and dynamically allocate resources according to the relationship between nodes, which is also reflected in our case study. Risk assessment is the overall process of risk identification, risk analysis and risk evaluation [53]. Our proposed “B-R model” is a risk assessment model for the whole process.

If the user enters risk factors that are not included in the BN library, the library can only make recommendations based on data characteristics, which may lead to some problems. By building an ontology model of risk factors, or by synonymous matching of synonyms, the problem of comparing risk factor characteristics can be solved. Furthermore, the training data contained in the BN library is a small amount. However, high-quality and public data sets in terrorism are limited. The information about terrorists, defence forces, and emergency forces is often the government’s confidential information. Once this information is made public, terrorist organizations will attack weak points of defence. Therefore, the construction of this Bayesian network relies heavily on expert experience and lacks actual data support.

The risk knowledge graph described in section 4 has great potential. For example, based on the existing risk knowledge, the user can predict the possible high-risk areas and organizations. Clustering and data mining on the targets with similar risk values to find out the common rules and characteristics is the next step in system development. The calculation of risk value based on game theory and PRA methods needs to be optimized, such as the two-level optimization models [54]. A robust decision analysis should be developed for risk management because attackers may know more about attack options than defenders [55].

The resource consumed by the method proposed in Figure. 2 is not abundant and can be easily deployed locally. However, as the data increases, the server may be needed to improve performance. When the BN library continues to accumulate data sets and BN models in the use of a large number of users, its practicality may also be significantly enhanced. On this basis, it is an important direction to develop a method similar to meta-learning to generate the best risk analysis model for users automatically. It is worth noting that the B-R model lacks an evaluation indicator. Evaluation of the rationality of risk factors, the accuracy of Bayesian networks, and the reliability of risk values are all critical. The effectiveness of the model depends on user evaluation, but user evaluation, accuracy, and recall rate are both critical components of method capabilities.

VII. CONCLUSION

This work focuses on chemical terrorist attacks in cities. The proposed B-R method can effectively integrate a large amount of data from various departments and help with the risk analysis and assessment of critical urban targets. The results can provide decision support for investigations, early warning of law enforcement and emergency department. Specifically, we built a Bayesian network library, which can effectively reduce the learning cost of users’ risk analysis of the chemical terrorist attack and enable users to construct a Bayesian network model more quickly and effectively. We quantified the value of risk, the level of defence *et al.*, not

TABLE 10. States of BN nodes.

No.	Nodes	States of Bayesian Nodes
1	Religious background	(1) Cult terrorism; (2) Islamic terrorism; (3) Christian terrorism; (4) Jewish terrorism; (5) Other
2	Region of the perpetrator	(1) Middle East and North Africa; (2) Europe; (3) Americas; (4) South and Southeast Asia; (5) East and Central Asia; (6) Central and North Africa
3	Number of members	(1) Less than 50 people; (2) 50-500 people; (3) 500-5,000 people; (4) More than 5,000 people
4	Average educational level	(1) High; (2) Medium; (3) Low
5	Technical background	(1) High level; (2) Middle level; (3) Low level
6	Social relations and organizational components	(1) Complex and diverse; (2) Medium; (3) Single
7	Whether they have been reported recently	(1) Yes; (2) No; (3) Unknown
8	Whether they ever launched chemical attack	(1) Yes; (2) No
9	Whether they made a statement or threat	(1) Yes; (2) No
10	Source of chemical weapon	(1) Self-made; (2) Occupied inventory or armory; (3) Steal from elsewhere; (4) Black market
11	Whether they have technical support	(1) Yes; (2) No
12	Whether they have the capabilities of storing and transporting the chemical weapon	(1) Yes; (2) No
13	Whether they have the ability to launch the chemical attack	(1) Yes; (2) No
14	Population density	(1)> 1000 / km2; (2) 500 ~ 1000 / km2; (3) <500 / km2
15	Population movement	(1) High; (2) Medium; (3) Low
16	Traffic condition	(1) Good; (2) Bad
17	Location	(1) Residential area; (2) Commercial area; (3) Open space
18	Whether it is a high-value target	(1) Yes; (2) No
19	Patrol	(1) More than 2 times; (2) Less than 2 times
20	Security check	(1) Yes; (2) No
21	Surveillance	(1) 24 hours; (2) Non-24 hours
22	Police investigation	(1) Yes; (2) No
	Influence of the terrorist organization	(1) Large; (2) Medium; (3) Small
	Activity level of the terrorist organization	(1) Inactive; (2) Active; (3) Very active
	Difficulty in obtaining and using the chemical weapon	(1) Low; (2) Medium; (3) High
	Target attraction	(1) High; (2) Medium; (3) Low
	Ability of the emergency response	(1) High; (2) Medium; (3) Low
	Threat of the terrorist organization	(1) Large; (2) Medium; (3) Small
	Whether the attack is successful	(1) Yes; (2) No
	Casualties	(1) Minor (0 to 10 deaths or 0 to 50 injuries); (2) Medium (11 to 30 deaths or 50 to 100 injuries); (3) Major (more than 30 deaths or more than 100 injuries)

just the probability distribution of the nodes in BN. Based on these values, the risks of different targets can be compared, and defensive resources can be effectively allocated. A risk knowledge graph is also constructed to facilitate users to acquire knowledge and take the next action after the risk analysis. A web application is designed that allows users to dynamically adjust the risk analysis model and perform visual analysis, which also proves the feasibility of our method.

APPENDIX

See table 10.

REFERENCES

[1] S. J. Balk, B. A. Gitterman, and M. D. Miller, "Chemical-biological terrorism and its impact on children: A subject review," *Pediatrics*, vol. 105, no. 3, pp. 662–670, Mar. 2000.

[2] F. M. Henretig, T. J. Cieslak, and E. M. Eitzen, "Biological and chemical terrorism," *J. pediatrics*, vol. 141, no. 3, pp. 311–326, 2002.

[3] C. Santos, T. El Zahran, J. Weiland, M. Anwar, and J. Schier, "Characterizing chemical terrorism incidents collected by the global terrorism database, 1970-2015," *Prehospital Disaster Med.*, vol. 34, no. 4, pp. 385–392, Aug. 2019.

[4] S. R. Emmett and P. G. Blain, "Chemical terrorism," *Medicine*, vol. 48, no. 3, pp. 182–184, Mar. 2020, doi: 10.1016/j.mpmed.2019.12.008.

[5] C. Strack, "The evolution of the Islamic state's chemical weapons efforts," *CTC Sentinel*, vol. 10, no. 9, pp. 19–23, 2017.

[6] T. Okumura, N. Takasu, S. Ishimatsu, S. Miyanoki, A. Mitsuhashi, K. Kumada, K. Tanaka, and S. Hinohara, "Report on 640 victims of the Tokyo subway sarin attack," *Ann. Emergency Med.*, vol. 28, no. 2, pp. 129–135, Aug. 1996.

[7] K. C. Hyams, F. M. Murphy, and S. Wessely, "Responding to chemical, biological, or nuclear terrorism: The indirect and long-term health effects may present the greatest challenge," *J. Health Politics, Policy Law*, vol. 27, no. 2, pp. 273–292, Apr. 2002.

[8] H. Hu, "Tear gas-harassing agent or toxic chemical weapon?" *JAMA, J. Amer. Med. Assoc.*, vol. 262, no. 5, pp. 660–663, Aug. 1989.

[9] A. S. Khan, A. M. Levitt, and M. J. Sage, "Biological and chemical terrorism: Strategic plan for preparedness and response," *MMWR*, vol. 49, pp. 1–14, Jan. 2000.

[10] M. E. Keim, N. Pesik, and N. A. Y. Twum-Danso, "Lack of hospital preparedness for chemical terrorism in a major US city: 1996–2000," *Prehospital Disaster Med.*, vol. 18, no. 3, pp. 193–199, Sep. 2003.

[11] A. S. Khan, D. L. Swerdlow, and D. D. Juranek, "Precautions against biological and chemical terrorism directed at food and water supplies," *Public Health Rep.*, vol. 116, no. 1, pp. 3–14, Jan. 2001.

[12] R. J. Kendall, S. M. Presley, and S. S. Ramkumar, *New Developments in Biological and Chemical Terrorism Countermeasures*. Boca Raton, FL, USA: CRC Press, 2016.

[13] S. Perry, R. Apel, G. R. Newman, and R. V. Clarke, "The situational prevention of terrorism: An evaluation of the Israeli west bank barrier," *J. Quant. Criminol.*, vol. 33, no. 4, pp. 727–751, Dec. 2017, doi: 10.1007/s10940-016-9309-6.

[14] S. Tutun, M. T. Khasawneh, and J. Zhuang, "New framework that uses patterns and relations to understand terrorist behaviors," *Expert Syst. Appl.*, vol. 78, pp. 358–375, Jul. 2017, doi: 10.1016/j.eswa.2017.02.029.

[15] I. Akgun, A. Kandakoglu, and A. F. Ozok, "Fuzzy integrated vulnerability assessment model for critical facilities in combating the terrorism," *Expert Syst. Appl.*, vol. 37, no. 5, pp. 3561–3573, May 2010, doi: 10.1016/j.eswa.2009.10.035.

[16] R. J. B. Garcia and D. von Winterfeldt, "Defender-attacker decision tree analysis to combat terrorism," *Risk Anal.*, vol. 36, no. 12, pp. 2258–2271, Dec. 2016, doi: 10.1111/risa.12574.

[17] G. L. Keeney and D. von Winterfeldt, "Identifying and structuring the objectives of terrorists," *Risk Anal.*, vol. 30, no. 12, pp. 1803–1816, Dec. 2010, doi: 10.1111/j.1539-6924.2010.01472.x.

[18] G. White, M. D. Porter, and L. Mazerolle, "Terrorism risk, resilience and volatility: A comparison of terrorism patterns in three southeast Asian countries," *J. Quant. Criminol.*, vol. 29, no. 2, pp. 295–320, Jun. 2013, doi: 10.1007/s10940-012-9181-y.

[19] A. Rezazadeh, L. Talarico, G. Reniers, V. Cozzani, and L. Zhang, "Applying game theory for securing oil and gas pipelines against terrorism," *Rel. Eng. Syst. Saf.*, vol. 191, Nov. 2019, Art. no. 106140, doi: 10.1016/j.ress.2018.04.021.

[20] I. Toure and A. Gangopadhyay, "Real time big data analytics for predicting terrorist incidents," in *Proc. IEEE Symp. Technol. Homeland Secur.*, May 2016, pp. 1–6.

[21] S. Nie and D. Sun, "Research on counter-terrorism based on big data," in *Proc. IEEE Int. Conf. Big Data Anal. (ICBDA)*, Mar. 2016, pp. 1–5.

[22] K. A. Boyd, "Modeling terrorist attacks: Assessing statistical models to evaluate domestic and ideologically international attacks," *Stud. Conflict Terrorism*, vol. 39, nos. 7–8, pp. 712–748, Jul. 2016, doi: 10.1080/1057610x.2016.1141003.

[23] K. Kaur, "Development of a framework for analyzing terrorism actions via Twitter lists," in *Proc. Int. Conf. Comput. Techn. Inf. Commun. Technol. (ICCTICT)*, Mar. 2016, pp. 19–24.

[24] L. Ehrlinger and W. Wöß, "Towards a definition of knowledge graphs," in *Proc. 12th Int. Conf. Semantic Syst. (SEMANTiCS)*, Leipzig, Germany, vol. 48, Sep. 2016, pp. 1–4.

[25] H. Paulheim, "Knowledge graph refinement: A survey of approaches and evaluation methods," *Semantic Web*, vol. 8, no. 3, pp. 489–508, Dec. 2016.

[26] A. Singhal. *Introducing the Knowledge Graph: Things, Not Strings*. Accessed: Dec. 4, 2012. [Online]. Available: http://googleblog.blogspot.be/2012/05/introducing-knowledge-graph-things-not.html

[27] J. S. Eder, "Knowledge graph based search system," Google Patents 13 404 109, Jun. 21, 2012.

[28] P. A. Bonatti, S. Decker, A. Polleres, and V. Presutti, "Knowledge graphs: New directions for knowledge representation on the semantic Web (Dagstuhl seminar 18371)," *Dagstuhl Rep.*, vol. 8, no. 9, pp. 29–111, 2019.

[29] Q. Wang, Z. Mao, B. Wang, and L. Guo, "Knowledge graph embedding: A survey of approaches and applications," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 12, pp. 2724–2743, Dec. 2017.

[30] J. Yan, C. Wang, W. Cheng, M. Gao, and A. Zhou, "A retrospective of knowledge graphs," *Frontiers Comput. Sci.*, vol. 12, no. 1, pp. 55–74, Feb. 2018.

[31] K. Jha and W. Jin, "Mining hidden knowledge from the counterterrorism dataset using graph-based approach," in *Proc. Int. Conf. Appl. Natural Lang. to Inf. Syst.* Springer, 2016, pp. 310–317.

[32] T. Xia and Y. Gu, "Building terrorist knowledge graph from global terrorism database and wikipedia," in *Proc. IEEE Int. Conf. Intell. Secur. Informat. (ISI)*, Jul. 2019, pp. 194–196.

[33] S. Bajpai, A. Sachdeva, and J. P. Gupta, "Security risk assessment: Applying the concepts of fuzzy logic," *J. Hazardous Mater.*, vol. 173, nos. 1–3, pp. 258–264, Jan. 2010, doi: 10.1016/j.jhazmat.2009.08.078.

[34] F. Argenti, G. Landucci, V. Cozzani, and G. Reniers, "A study on the performance assessment of anti-terrorism physical protection systems in chemical plants," *Saf. Sci.*, vol. 94, pp. 181–196, Apr. 2017, doi: 10.1016/j.ssci.2016.11.022.

[35] H. H. Willis and T. LaTourrette, "Using probabilistic terrorism risk modeling for regulatory benefit-cost analysis: Application to the western hemisphere travel initiative in the land environment," *Risk Anal.*, vol. 28, no. 2, pp. 325–339, Apr. 2008.

[36] A. Volkanovski, M. Čepin, and B. Mavko, "Application of the fault tree analysis for assessment of power system reliability," *Rel. Eng. Syst. Saf.*, vol. 94, no. 6, pp. 1116–1127, Jun. 2009, doi: 10.1016/j.ress.2009.01.004.

[37] Z. L. Yang, J. Wang, S. Bonsall, and Q. G. Fang, "Use of fuzzy evidential reasoning in maritime security assessment," *Risk Anal.*, vol. 29, no. 1, pp. 95–120, Jan. 2009, doi: 10.1111/j.1539-6924.2008.01158.x.

[38] J. L. Regens, N. Mould, C. J. Jensen, M. A. Graves, and D. N. Edger, "Probabilistic graphical modeling of terrorism threat recognition using Bayesian networks and Monte Carlo simulation," *J. Cognit. Eng. Decis. Making*, vol. 9, no. 4, pp. 295–311, Dec. 2015.

[39] D. Husmeier, *Introduction to Learning Bayesian Networks from Data*. London, U.K.: Springer, 2005, pp. 17–57.

[40] J. Wei, J. Wang, and H. Yu, "Terrorism threat assessment with multi-module Bayesian network," *J. Univ. Chin. Acad. Sci.*, vol. 32, no. 2, pp. 264–272, 2015.

[41] Z. Fu, R. Xu, and W. Liu, "Research on terrorist attack warning model based on Bayesian network," *J. Catastrophol.*, vol. 31, no. 3, pp. 184–189, 2016.

[42] M. O. Mohammed, O. A. Glenn, M. D. Kristen, and C. S. Jack, "A Bayesian belief network of threat anticipation and terrorist motivations," in *Proc. SPIE*, Orlando, FL, USA, vol. 7666, 2010, doi: 10.1117/12.849464.

[43] R. Zhu, X. Hu, X. Li, H. Ye, and N. Jia, "Modeling and risk analysis of chemical terrorist attacks: A Bayesian network method," *Int. J. Environ. Res. Public Health*, vol. 17, no. 6, p. 2051, Mar. 2020, doi: 10.3390/ijerph17062051.

[44] G. LaFree and L. Dugan, "Introducing the global terrorism database," *Terrorism Political Violence*, vol. 19, no. 2, pp. 181–204, Apr. 2007.

[45] Z. Tang, Y. Li, X. Hu, and H. Wu, "Risk analysis of urban dirty bomb attacking based on Bayesian network," *Sustainability*, vol. 11, no. 2, p. 306, 2019.

[46] N. Khakzad, I. S. Martinez, H.-M. Kwon, C. Stewart, R. Perera, and G. Reniers, "Security risk assessment and management in chemical plants: Challenges and new trends," *Process Saf. Prog.*, vol. 37, no. 2, pp. 211–220, Jun. 2018.

[47] L. Zhou and D. Zhang, "NLPIR: A theoretical framework for applying natural language processing to information retrieval," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 54, no. 2, pp. 115–123, Jan. 2003.

[48] J. Webber, "A programmatic introduction to Neo4j," in *Proc. 3rd Annu. Conf. Syst., Program., Appl., Softw. Humanity (SPLASH)*, 2012, pp. 217–218.

[49] J. A. Major, "Advanced techniques for modeling terrorism risk," *J. Risk Finance*, vol. 4, no. 1, pp. 15–24, Apr. 2002.

[50] B. C. Ezell, S. P. Bennett, D. Von Winterfeldt, J. Sokolowski, and A. J. Collins, "Probabilistic risk analysis and terrorism risk," *Risk Anal., Int. J.*, vol. 30, no. 4, pp. 575–589, 2010.

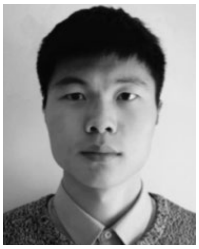
[51] J. Zhang, S. Shen, and R. Yang, "Asymmetric information in combating terrorism: Is the threat just a bluff?" *Tsinghua Sci. Technol.*, vol. 15, no. 5, pp. 604–612, Oct. 2010.

[52] T. Sandler, "Terrorism & game theory," *Simul. Gaming*, vol. 34, no. 3, pp. 319–337, Sep. 2003, doi: 10.1177/1046878103255492.

[53] S. N. Luko "Risk management principles and guidelines," *Quality Eng.*, vol. 25, no. 4, pp. 451–454.

[54] J. Cox and L. Anthony, "Some limitations of 'Risk= threat × vulnerability × consequence' for risk analysis of terrorist attacks," *Risk Anal. Int. J.*, vol. 28, no. 6, pp. 1749–1761, 2008.

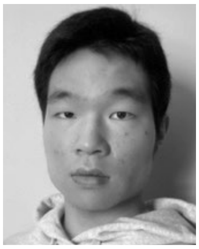
[55] G. G. Brown and L. A. T. Cox, Jr., "How probabilistic risk assessment can mislead terrorism risk analysts," *Risk Anal.*, vol. 31, no. 2, pp. 196–204, Feb. 2011.



RONGCHEN ZHU received the bachelor's degree in cyber security and law enforcement from the People's Public Security University of China, where he is currently pursuing the master's degree. He has published some academic articles and participated in some projects. His research interests include risk analysis and assessment, Bayesian network methods and applications, and knowledge graph.

XIAOFENG HU received the Ph.D. degree in nuclear science and technology from Tsinghua University, in 2014. He is currently an Associate Professor with the People's Public Security University of China. His research areas include risk assessment techniques, public safety and emergency management, machine learning, and big data analysis.

XIN LI received the Ph.D. degree from the Department of Computer Science, Zhejiang University, in 2007. He has been engaged in research on cyber security, big data, and artificial intelligence. He is currently an Associate Professor with the People's Public Security University of China, Beijing. He is also with the College of Information Technology and Cyber Security, People's Public Security University of China. He has published more than 30 articles in prestigious peer-reviewed journals and conferences.



HAN YE is currently pursuing the bachelor's degree in security engineering with the People's Public Security University of China. He has participated in several academic projects. His research interests include security prevention, risk assessment, and knowledge graph.