International Journal of Information Technology and Web Engineering
Volume 17 • Issue 1

Winning the War on Terror:
Using "Top-K" Algorithm and CNN to Assess the Risk of Terrorists

Yaojie Wang, Engineering University of PAP, China
Xiaolong Cui, Engineering University of PAP, China
Peiyong He, Engineering University of PAP, China

ABSTRACT

From the perspective of counterterrorism strategies, terrorist risk assessment has become an important approach for counterterrorism early warning research. Combining with the characteristics of known terrorists, a quantitative analysis method of active risk assessment method with terrorists as the research object is proposed. This assessment method introduces deep learning algorithms into social computing problems on the basis of information coding technology. The authors design a special "Top-k" algorithm to screen the terrorism related features and optimize the evaluation model through convolution neural network so as to determine the risk level of terrorist suspects. This study provides important research ideas for counterterrorism assessment and verifies the feasibility and accuracy of the proposed scheme through a number of experiments, which greatly improves the efficiency of counterterrorism early warning.

KEYWORDS

Characteristic of Terrorist, Convolutional Neural Network, Non-Privacy Virtual Data, Risk Assessment, "Top-K" Algorithm

1. INTRODUCTION

Terrorism poses threat to the entire society, from which many countries have been suffering. With the rapid development of global information technology, international terrorism activities are upgrading, resulting in an increasingly severe situation of international counter-terrorism. In order to effectively combat terrorism, the focus of global counter-terrorism has gradually shifted from passive "emergency" to active "counter-terrorism early warning", which can prevent terrorist activities in time and greatly reduce harmful effects (Li, 2017; Hai & Xiaofeng, 2019). As an important part of counter-terrorism early warning, terrorism risk assessment is not only an important basis for formulating various security plans, but also an effective way to avoid attacks. Scientific risk analysis methods in the fields of finance and disaster prevention have been introduced into terrorist risk assessment (Baker, 2009; Phelps, 2009). The ability to quantify risk using a common performance metric across different hazards allows the risk manager to prioritize risks and to make informed and effective resource allocation and policy decisions. Although risk assessment is widespread in the field of social computing, security aspects are more complex to understand and measure (Samrat & Mark, 2011). For example, public areas are

DOI: 10.4018/IJITWE.288038

at risk of various attacks, one of which is a suicide terrorist or a bomber, targeting components of an infrastructure system (i.e., Bridges, tunnels, etc.) or innocent crowds.

In the post-9/11 era, there have also been more and more research studies focused on terrorism risk assessment in the world, which mainly conducts systematic research from the perspective of theoretical analysis. For example, Woo's work includes the development of a theoretical stochastic terrorism risk model providing the framework for probabilistic risk analysis (Woo, 2002). Since 2014, some experts have gradually begun to conduct research on quantified counter-terrorism risks (Jason & Gregory, 2011). The main objects of the research are important infrastructure and terrorist incidents themselves (Viscusi & Zeckhauser, 2017; Aaron & Ryan, 2013). However, there are relatively few researches on the risk assessment of terrorist organizations or terrorists, and the research methods are relatively single. Most of them still use Bayesian theory (Yongnan & Jianwei, 2017; Ming & Ren, 2020), social network analysis(SNA) and other methods for reasoning analysis (Ralph, 2007).

The main work of this paper is to assess the risk of terrorist suspects and design an active risk assessment modeling method based on deep learning, which can provide decision-making reference for the counter-terrorism department to actively prevent terrorist activities. The second part is an overview of terrorist risk assessment. The third part introduces the framework and process of the proposed method, using "Top-K" algorithm and CNN to assess the level of terrorist risk. The fourth part is the simulation analysis based on non-privacy virtual data, and finally the conclusion.

2. RELATED WORK

Once terrorist activities occur, they will bring major social risks and irreparable losses, which are a huge challenge for counter-terrorism operations (Jennifer & Louise, 2009; Pooja & Archana, 2021). There are great differences in the research contents of counter-terrorism threat risk assessment. From the perspective of the role attributes of research object, it can be divided into global risk assessment, fixed object risk assessment and active risk assessment (Scott, 2001; Novikov & Koshkin, 2019; Karl & John, 2008). The latter two are more focused and easy to quantify, and generally serve as the basis of global risk assessment. Due to the different objects of risk assessment, the methods adopted are also very different, resulting in great differences in the difficulty, effect and scope of application of the assessment.

Global risk assessment, which starts from a macro perspective, gradually shifts from theoretical research to multi-factor quantification. Pate-Cornell and Guikema (2002) adopted a systems approach and developed a theoretical probabilistic model for prioritizing terrorist threat and counter-terrorism strategies. Woo (2009) proposed to use event-trees for estimation of success probabilities of attacks and development of terrorism loss exceedance curves. In September 2002, Risk Management Solutions (RMS) released the first version of its "Terrorism Risk Model." The RMS model calculates expected annual consequences (human and economic) from varied terrorist threats. The methodology relies on the elicitation of particular attack scenarios at different targets using expert judgment, and assessing the capabilities for different attack modes, overall likelihood of attack, and ability to stage multiple coordinated attacks (Willis, 2007). Garrick *et al*. (2004) identifified the importance of processing intelligence information and developed a framework for scenario-based probabilistic terrorism risk assessments for assets and facilities. According to the results of RAND study, Willis (2007) also suggested that dividing risks into categories in terms of individual and population may help in making risk management decisions. Ezell and Winterfeldt (2009) acknowledged the use of probabilistic risk analysis and event-trees for terrorist risk assessment. Shafieezadeh *et al*. (2015) proposed an asset-level security risk management framework to assist stakeholders of critical assets with allocating limited budgets for enhancing their security against terrorist attack. More recently, Benxian *et al*. (2016) introduced the game theory into the analysis and assessment of counter-terrorism risk, assessed the long-term risk of terrorism based on policy, and then put forward more effective risk management measures. Junnan *et al*. (2020) introduced advanced knowledge graph technology

to terrorism events, and proposed a three-layer event model, which can deeply analyze the logical relationship of event elements.

The target of fixed object risk assessment is mainly important infrastructure or government officials, including bridges, railway stations, religious sites, and emerging cyberspace. Most of them exist in real physical world and belong to "fixed targets". At present, the research form of fixed object risk assessment mainly adopts quantitative methods. Unlike other risk assessments that focus on event likelihood and consequence, the Department of Homeland Security (DHS) believes that the components of terrorist risk of fixed object are threefold: (1) threat to a target, (2) target vulnerability, and (3) consequence of a successful attack (Samrat & Mark, 2011). The first two components of terrorist risk are considered to be probabilistic in nature, while consequence, as for other risk assessments, is considered to be deterministic. Using the results from the RAND study, Willis et al. (2005) proposed the RT method of terrorist attack risk expressed as: $RT=T\times V\times S$, and gives suggestions for resource allocation based on risk. To analysis risk indicators at all levels and to improve the scientificity, rationality and accuracy of the results, Yi (2019) designed a hierarchical evaluation model and analyzed various factors of terrorist attacks in religious sites using FAHP-SWOT. Xuan *et al*. (2016) proposed a bridge-oriented terrorist attack risk assessment simulation model based on the network analysis method (ANP), and built a terrorist attack risk assessment system in specific scenarios. Taking the whole airport or an aircraft as the research object, Feng *et al*. (2018) designed a model to assess the overall risk of the aircraft. On the basis of machine learning, Yan (2020) proposed a quantitative risk assessment method for civil aviation passengers by combining expert scoring and robot portrait. Some scholars also discuss the effect and application scope of fixed object risk assessment from legal, political and other factors. Generally speaking, the models proposed in most of the above studies have too many indicators and lack quantitative analysis of feature selection, resulting in a narrow scope of application and poor actual counter-terrorism risk assessment.

Active risk assessment mainly focuses on terrorist organizations or terrorists, and it can also be used in combination with fixed object risk assessment. Dr. Kathleen of the Carnegie Mellon university software institute and her team using dynamic network analytics to study how to effectively fight against the terrorist organisation network (Carley, 2006; Carley et al., 2003) and develop a text mining tool AutoMap (Carley et al., 2013) and a SNA tools ORA (Carley, 2016). Through the social network analysis software SNA and ORA, Xiaopeng (2018) proposed a terrorist network model, revealing the characteristics of terrorist networks. Guo *et al*. (2019) analyzed the GTD data and summarized the spatial distribution and temporal distribution characteristics of terrorist incidents. Majeske and Lauer (2012) studied the classification of passengers by Bayesian decision method, so as to improve the efficiency of airport security. Combined with the characteristics of terrorists in Asia, Yongnan and Jianwei proposed a quantitative analysis method of terrorist risk assessment by Bayesian theory. According to the data of China Airlines, Yujun *et al*. (2017) used deep neural networks to classify the risk of passenger based on their personal booking information. Yongbao *et al*. (2021) analyzed the risk level of terrorist suspects through the legality of the risk matrix, then used the "Borda" sequence value method to quantitatively calculate the values of different risk factors at the same risk level, which can give feasible suggestions for terrorist risk management.

All in all, although substantial progress has been made to quantify terrorism risk, the method of global risk assessment is qualitative in nature. Fixed object risk assessment has a narrow scope of application, resulting in poor application effects of actual terrorist risk assessment. The research method of active risk assessment is not enough to support the development of risk assessment, but as an ideal terrorist risk assessment method, it has a great development space. Nevertheless, this prior research has provided important insights for developing a more quantitative active risk assessment model.

3. RISK ASSESSMENT SCHEME DESIGN

3.1 Preliminaries

Convolutional neural networks (CNN), belonging to the category of deep learning, is a type of feedforward neural network model that includes convolution calculations (Lecun et al., 2010). CNN is a very representative network structure in deep learning, and has made breakthroughs in image analysis and feature extraction. Convolution calculation is essentially a one-way mapping, which can automatically learn a large number of mapping relationships between inputs and outputs, but does not require precise mathematical expressions. As long as the convolutional network is trained with known mapping logic, the network model has the ability to map between input and output. CNN is mainly composed of input layer, convolutional layer, activation function, pooling layer, fully connected layer, etc. The network depth and parameters can be flexibly set. Compared with other deep learning algorithms, the advantage of CNN is that they can perceive locally, similar to the structure of the visual system in biology, which can avoid the preprocessing process of manual labeling and classification. It also uses weight sharing mechanism, which greatly reduces the training parameters. CNN can process data of different dimensions. For example, one-dimensional CNN is mainly used for sequence processing, two-dimensional CNN is often used for image recognition, and three-dimensional CNN is mainly used for medicine and video (Abbasi et al., 2018), among which two-dimension CNN have the most prominent performance.

In this paper, a two-dimensional CNN is used to assess the risk level of terrorist suspects. It is worth noting that the feature matrix of terrorist suspects in this paper is regarded as a binary image, which can play the advantage of the two-dimensional CNN for image processing. For other deep learning methods, such as LSTM, RNN, etc., these methods are not applicable under the current setting conditions, and we will conduct in-depth research in the future.

3.2 Proposed Framework

Through the "Top-K" algorithm and CNN, an active risk assessment method with terrorists as the research object is proposed in this paper. This method first summarizes the general terrorism-related features of known terrorist organizations, and each feature has a corresponding binary code. Secondly, the "Top-K" algorithm was designed to screen out the top k terrorism-related features with the highest degree of relevance, and their corresponding codes were formed into a feature sequence in order. Then the feature sequence in the form of matrix, which is equivalent to a binary image, is input into convolutional neural network for training. After several iterations of optimization, the assess model can be obtained, which can assess the risk of the suspects. The specific framework includes the following three phases:

3.2.1 Summary of General Features

The summary of terrorist organization characteristics is the basic work of active risk assessment. Under the new severe situation, terrorist organizations around the world have shown the characteristics of "obvious regional characteristics, concentrated religious beliefs, frequent cross-domain communications, and highly networked" (Siqueira & Arce, 2020). Therefore, we analyze and summarize the risk characteristics of terrorists based on terrorist incidents that have occurred in the world. It is worth noting that we also refer to the latest public literature or regulatory documents, such as "cyber terrorism and cyber terrorism", "actions of world terrorist organizations", "75 specific manifestations of religious extremist activities" (Gumz, 2017; Choi et al., 2018).

In order to clearly illustrate the method of feature summarization, this chapter takes the most harmful terrorist organization "East Turkestan Islamic Movement (ETIM)" as an example, the ETIM has carried out terrorist activities for more than 20 years and has become an important part of the international terrorist network, which poses a huge threat to the international community. In the course of a series of strikes, the international community has also grasped the behavioral characteristics

of many ETIM terrorists. The general terrorism-related characteristics of the ETIM that have been mastered are summarized, as shown in Table 1.

Table 1. General terrorist feature summary and corresponding codes

	Feature label	Attribute category	Binary	Decimal
2 bit coding	Gender	Male or female	01/10	1/2
	Marital status	Male or female	01/10	1/2
	Is the job stable	Yes/no	01/10	1/2
	Fixed property	Yes/no	01/10	1/2
	Suspected of money laundering	Yes/no	01/10	1/2
	Have you ever been abroad	Yes/no		
	…	…	…	…
	Whether to buy a gun	Yes/no	01/10	1/2
3 bits coding	Education	College degree or above/junior and high school level/other levels	001/010/011	1/2/3
	Browse military news	Frequently browsed/occasionally browsed/ never browsed	001/010/011	1/2/3
	Browse cross-domain websites via proxy or VPN	Frequently browsed/occasionally browsed/ never browsed	001/010/011	1/2/3
	Use clothing or accessories with special signs	Wear specially marked clothes/wear specially marked items/both/none	001/010/011/101	1/2/3/4
	Whether to buy violent audio and video	Watch violent audio and video/Have violent audio/video/both/none	001/010/011/101	1/2/3/4
	…	…	…	…
	Illegal assembly	Participate frequently/participate occasionally/never participate	001/010/011	1/2/3

When the information of a certain feature label is missing, it is represented by the same bit of "0". For example, when the job information is missing, "00" is used for binary code, and "0" is used for decimal code.

Since the characteristics of terrorist organizations are relatively concentrated, we only need to perform the above summary work once, and there is no need to repeat screening for a long period of time, which greatly reduces the workload. In the actual coding process, a suitable coding rule is formulated for multi-label features. When a certain feature information is missing, fuzzy information can also be used to fill in. So we can build a database of characteristics of specific terrorist organizations and classify risk levels of terrorist suspects. In this paper, the risk level is divided into three categories, corresponding to red, orange, and yellow respectively, as shown in Table 2.

3.2.2 Design Of "Top-K" Algorithm

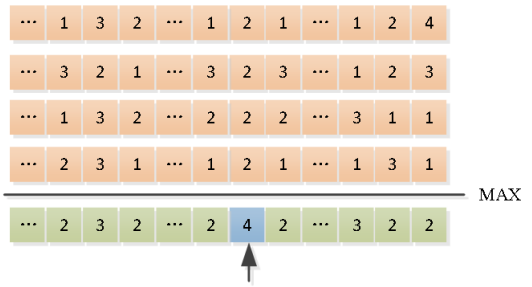
Terrorist characteristic database provides general data, which can not be directly used for active terrorist risk assessment. In order to study the terrorist organizations in a certain region, it is necessary to further screen out the key features of the database. For example, the terrorist characteristics of ETIM in different regions have obvious differences, and the effect of using general features directly is very poor. Therefore, a "Top-K" algorithm is designed in the paper, and its goal is to output the top k features label with the highest terrorist relevance on the basis of general features, so as to delete the irrelevant feature factors and improve the computational efficiency. The processing process of the "Top-K" algorithm is as follows:

Table 2. Three level classification of terrorist risk

Level sort	definition	corresponding code
Level 1	Once the risk occurs, the target index of the project will drop seriously, and the impact will be extremely bad	R(red)
Level 2	Once the risk occurs, the project will be moderately affected, but the project objectives can be partially achieved	O(orange)
Level 3	Once the risk occurs, the project will be slightly affected	Y(yellow)

① ***Define the Max function.*** The Max function can obtain the most relevant feature label among terrorists. Taking the decimal code as an example, the general characteristics of each terrorist form a decimal sequence in order. The Max function sums all the decimal sequences by column conditions. The largest frequency number in each column is the corresponding outcome. By comparing the results of each column, then we can get the Max position in the sequence (the Max position is finally marked as 1, and the rest are marked as 0). The specific principle is shown in Figure 1.

Figure 1. Schematic diagram of max function



In the actual programming environment, the Max function can be realized by simple circuit. Assuming that the general feature sequences of two terrorists are known, we input their respective n-bit feature sequences $\mathbf{x} = (x_1,...,x_n)$ and $\mathbf{x} = (x_1,...,x_n)$, and the Max function outputs $\mathbf{b} = (b_1,...,b_n)$:

$$b_i = \mathrm{EQ}(\mathrm{MAX}(\mathrm{ADD}(x_1, y_1),..., \mathrm{ADD}(x_n, y_n)), \mathrm{ADD}(x_i, y_i))$$

Where $\mathrm{ADD}(\cdot)$ performs conditional addition on the input sequence, $\mathrm{MAX}(\cdot)$ outputs the maximum value of all input data, and $\mathrm{EQ}(\cdot)$ performs an XOR operation on the two input sequences. $\mathrm{ADD}(\cdot)$ is equivalent to the adder in the circuit. The function of $\mathrm{ADD}(\cdot)$ is to find the most value circuit, and the function of $\mathrm{EQ}(\cdot)$ is equivalent to the comparison circuit. The whole process belongs to addition or XOR operation, and the calculation efficiency is very high through hardware acceleration.

,***Recursive algorithm.*** If the sequence $\mathbf{x}_i = (x_{i,1},..., x_{i,n})$, $i \in \{1,...,h\}$ of h terrorists is input, $\mathbf{b} = (b_{top-k,1},..., b_{top-k,n})$ is output. By using the Max function k times in the loop, the top k feature labels with the highest degree of relevance can be output. The recursive algorithm is simple and easy to use, and its computational complexity is low. The pseudo code of Algorithm 1 is shown below.

Algorithm 1: "Top-K" algorithm

Input: the bit string $\mathbf{x}_i = (x_{i,1},\ldots,x_{i,n})$ of the n-dimensional sequence of h terrorists, $i \in \{1,\ldots,h\}$

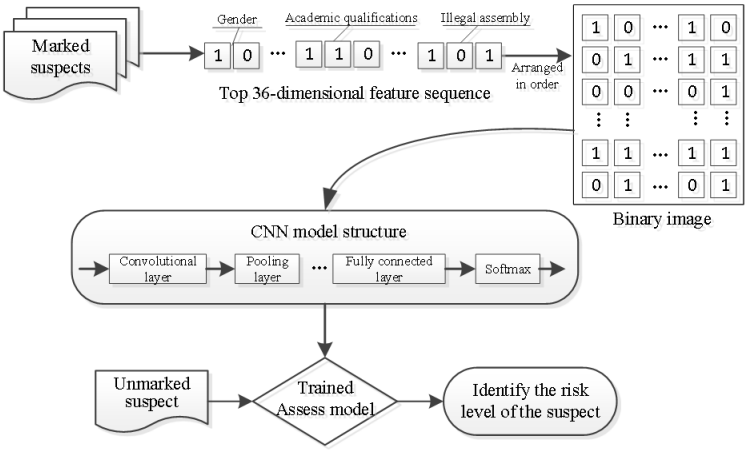
Output: $\mathbf{b} = (b_{top-k,1},\ldots,b_{top-k,n})$ (the top k features label with the highest terrorist relevance)

1. $\overline{x}_{top-1,i} = \mathrm{ADD}(\mathbf{x}_{1,i},\ldots,\mathbf{x}_{h,i})$, and $1 \le i \le n$
2. ***For r =1...k-1 do***
3. $\overline{b}_{top-r,i} = \mathrm{EQ}(\mathrm{MAX}(\overline{x}_{top-r,1},\ldots,\overline{x}_{top-r,n}),\overline{x}_{top-r,i})$, $1 \le i \le n$
4. $\overline{x}_{top-(r+1),i} = \prod_{j=1}^{r}(1-\overline{b}_{top-j,i})\overline{x}_{top-r,i}$, $1 \le i \le n$
5. ***end For***
6. $\overline{b}_{top-k,i} = \mathrm{EQ}(\mathrm{MAX}(\overline{x}_{top-k,1},\ldots,\overline{x}_{top-k,n}),\overline{x}_{top-k,i})$, $1 \le i \le n$
7. $b_{top-k,i} = 1-\prod_{j=1}^{k}(1-\overline{b}_{top-j,i})$, $1 \le i \le n$
8. **return** $\mathbf{b} = (b_{top-k,1},\ldots,b_{top-k,n})$

3.2.3 Training Model

According to the obtained k feature labels, we reconstruct the feature sequence in order. Then the feature sequence in the form of matrix, which is equivalent to a binary image, is input into the CNN network for training. After several iterations of optimization, the assess model can be obtained, which can assess the risk of the suspects. The basic process is shown in Figure 2.

Figure 2. The basic process of the proposed assess model



For suspicious personnel of different risk levels, we should adopt different security measures, especially in the deployment of police resources. In general, we need to focus on red or orange suspicious persons, and adopt travel monitoring, network tracking, social user portrait and other security measures. Limited by the police resources, the way of random sampling and regular re-

examination can be adopted for the yellow suspicious personnel. Once the risk level of a suspicious person changes, we should immediately upgrade to the corresponding security measures to prevent the occurrence of terrorist activities. In practical applications, it is possible to flexibly set the level division, security scheme formulation, and the resource budget according to the regional risk.

4. SIMULATION ANALYSIS OF NON-PRIVACY VIRTUAL DATA

4.1 Assumptions

Taking into account the factor of confidentiality, the feasibility of the proposed method is verified by non-privacy virtual data in this paper. The analysis process of non-privacy virtual data can be applied to counter-terrorism teaching applications, which facilitates the exchange and training of intelligence analysis. The case assumption is as follows: according to the statistics of counter-terrorism agencies, a total of 1200 suspects have been found in *R-Area* from 2010 to 2019, including 125 red suspects, 450 orange suspects and 625 yellow suspects. The detailed information characteristics of suspects are shown in Table 2. At *P* railway station, the security agency found suspicious outsiders *Alice* and *Bob* from *R-Area*. The preliminary information is shown in Table 3. How to judge the risk level of the two men and what precautions need to be taken?

4.2 Simulation Experiment Settings

During the simulation experiment, the non-privacy virtual data in Table 2 has been coded and classified, so we can skip the step of summarizing general features. Although the virtual data in this paper seems simple, it has been able to clarify the core point of the proposed method. In practice, the coding rules of feature data are very complex and numerous. The smaller the granularity of the feature label, the better the assess effect. We mainly explain the settings of the "Top-K" algorithm and CNN as follows:

① ***The design of "Top-K" algorithm circuit.*** Before the design of the circuit, the general feature sequence corresponding to each suspect in *R-Area* is established in order. We select the top 36 feature labels with the highest degree of relevance, that is, K=36. Through XOR and addition operations, the core circuit design of "Top-36" is shown in Figure 3.

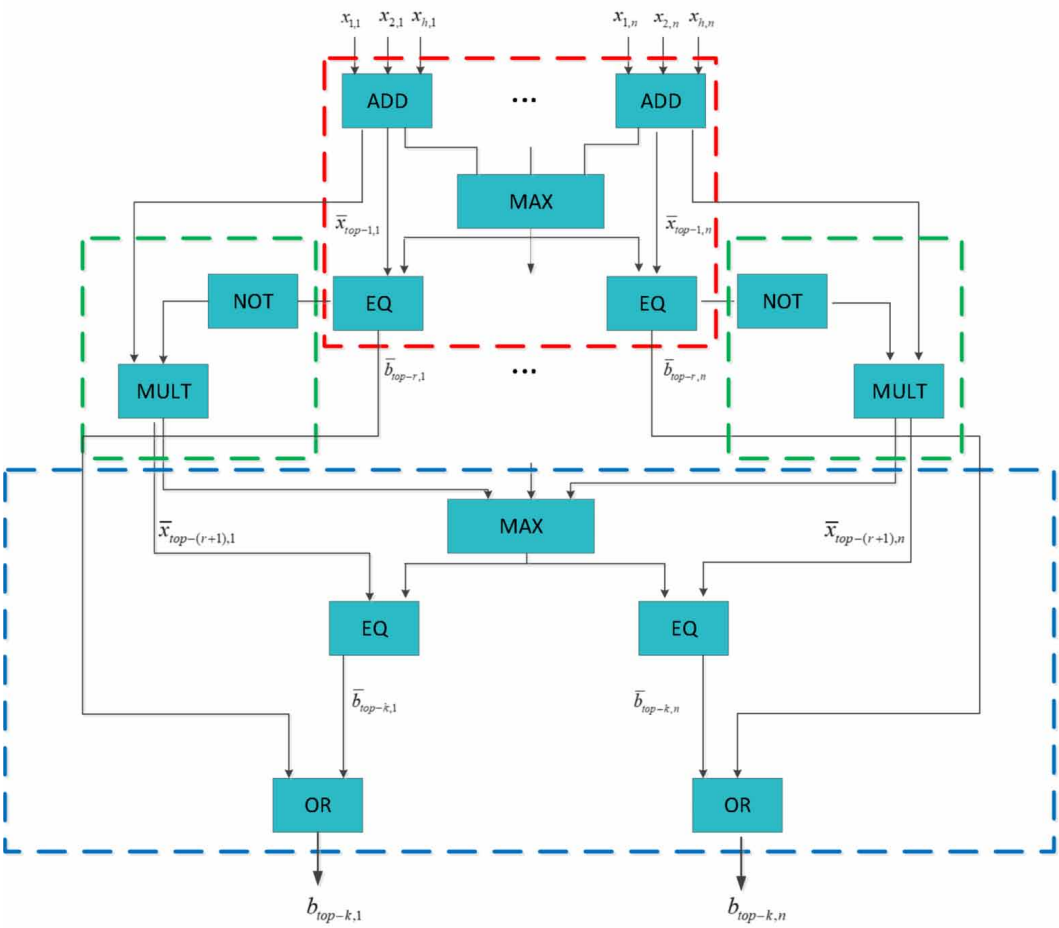
,***The setting of network structure and parameters.*** For the training process of the model, 1,000 samples were randomly selected from the 1200 sample data in *R-Area* as the training set, and the remaining 200 samples were the test set. The experimental platform uses TensorFlow V0.12, NVIDIA 1080 graphics card and 16G memory. In this experiment, the convolutional neural network uses Adam's optimization algorithm, the learning rate is 0.001, and the output is three categories, corresponding to the three risk levels of red, orange, and yellow respectively. Other specific parameter settings are shown in Figure 4.

4.3 Analysis of Experimental Results

The accuracy of the assess model in the simulation experiment is shown in Figure 5. With the training period increases, the error rate of the proposed model continues to decrease, especially when training for 200 periods, the average error rate of the assess model is less than 7%, which has been able to meet the needs of daily counter-terrorism applications.

In this simulation, we used the same network settings to verify the test set under different training periods. In the training period in the range of (100, 250), we take every 5th, and calculate the average value of the recognition accuracy of red, orange, and yellow, a total of 30*3=90 situations. Draw each mean value as a three-dimensional histogram, as shown in Figure 6(a) (the coordinate points 1, 2, and 3 on the X axis represent three risk levels of red, orange, and yellow respectively). In Figure 6(a), the recognition accuracy is obviously increasing within the training period range of (100, 200); The recognition accuracy rate maintains a steady state within the training period of (200, 250). In

Figure 3. The core circuit design of the "Top-36" algorithm



order to further clarify the performance of the proposed model, taking 200 training periods as an example, the recognition accuracy rate under the test set conditions is shown in Figure 6(b). The recognition accuracy of the three risk levels are all over 90%, especially the recognition accuracy of the red level is 95%. Therefore, the above simulation experiments can prove that the method proposed in this paper is feasible and has great application value, which is worth continuing to study in depth.

In the classification phase, we input the feature information of *Alice* and *Bob* in Table 3 into the trained assess model, and then classify the risks of the two people. We can realize that *Alice* is a red level suspect and needs the highest level of early warning and surveillance measures against her. *Bob* is a yellow-level suspect, which can be randomly selected according to local actual condition.

Table 3. Information on suspicious persons in R-Area from 2010 to 2019

Feature label(abbreviation)	Red level (125 suspicious persons)					Orange level (450 suspicious persons)					Yellow level (625 suspicious persons)				
Name(Na)	*A1*	*A2*	*A3*	…	*A125*	*B1*	*B2*	*B3*	…	*B450*	*C1*	*C2*	*C3*	…	*C625*
Gender(Ge)	10	10	10	…	10	11	10	11	…	10	10	01	10	…	10
Suspected of money laundering (Su)	01	01	01	…	00	1	0	0	…	10	01	10	01	…	10
Is the job stable(Js)	10	10	01	…	10	01	01	10	…	10	10	00	10	…	10
Whether to buy a gun(Bg)	10	00	10	…	10	10	00	10	…	10	10	10	10	…	10
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
Have you ever been abroad(Ab)	01	10	10	…	10	10	10	10	…	10	10	10	10	…	10
Fixed property(Fp)	10	11	10	…	10	11	10	10	…	10	00	01	01	…	00
Browse cross-domain websites via proxy or VPN(Bw)	001	010	001	…	011	010	001	001	…	011	001	001	011	…	011
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
Education(Ed)	001	010	001	…	011	010	001	001	…	011	010	001	011	…	011
Whether to watch violent audio and video(Wv)	001	010	001	…	011	010	001	001	…	011	001	001	011	…	011
Browse military news(Mn)	001	010	001	…	011	010	001	010	…	001	001	001	011	…	011
Illegal assembly(Ia)	001	010	001	…	011	010	001	001	…	011	011	001	001	…	011

Table 4. Information of suspicious persons to be classified

Na	Ge	Su	Js	Bg	…	Ab	Fp	Bw	…	Ed	Wv	Mn	Ia
Alice	10	10	10	10	…	10	10	010	…	010	010	010	010
Bob	01	00	10	00	…	10	01	011	…	011	011	001	000

Figure 4. The structure and parameters of the proposed CNN model

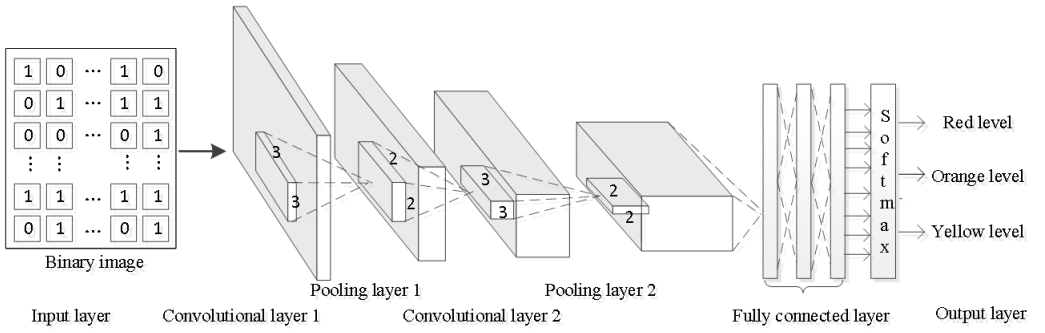


Figure 5. Training process of the proposed CNN model (1000 samples)

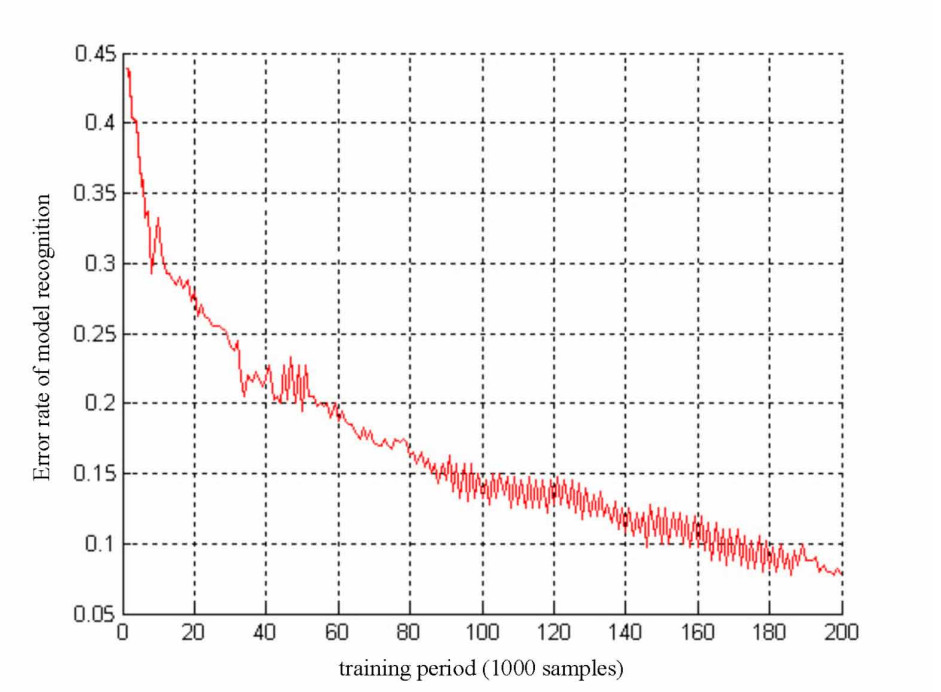
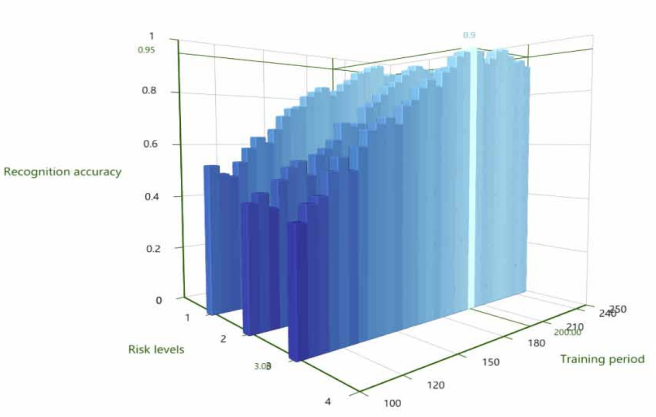


Figure 6. The recognition accuracy of the model on the test set



5. CONCLUSION

The study of counter-terrorism early warning is the top priority of global counter-terrorism, and its essence is the process of risk analysis and assessment of intelligence. From the perspective of active counter-terrorism, this research conducts risk level assessment around terrorists or suspects. In this paper, we proposed an active method to assess the risk of terrorist by using the "Top-K" algorithm and CNN, and detailed descriptions in three aspects: general feature summary, "Top-K" circuit design, and assess model setup. In daily terrorist risk assessment applications, we can flexibly set the "Top-K" algorithm according to the actual condition, which has high practicability. We also verify the feasibility of the proposed method with non-private virtual data, which facilitates the communication and learning between security intelligence agencies. Through this research, our team hopes to provide a new perspective for counter-terrorism early warning and play a positive role in effectively combating terrorist activities.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (NSFC) (No. U1603261), National Key R&D Program of China (Grant No.2017YFB0802000), Innovative Research Team in Engineering University of PAP(No.KYTD201805), Natural Science Foundation of Xinjiang Uygur Autonomous Region (No.2016D01A080), the basic research foundation project of Engineering University of PAP (No.WJY201920).

REFERENCES

Aaron, C., & Ryan, W. (2013). Estimating the historical and future probabilities of large terrorist events. *The Annals of Applied Statistics*, *7*(4), 1838–1865. doi:10.1214/12-AOAS614

Baker, H. U. S. (2005). *Urges "All Hazard" Approach in Disaster Management*. Available at: http://www.iwar.org.uk/ news-archive/2005/01–21.htm

Benxian, L., & Wei, P. (2016). Quantitative analysis methods and key technologies of anti-terrorism research. *Journal of Xinjiang Police College*, *36*(2), 4–9.

Carley, K. M. (2006). Destabilization of covert networks. *Computational & Mathematical Organization Theory*, *12*(1), 51–66. doi:10.1007/s10588-006-7083-y

Carley, K. M. (2016). *ORA-NetScenes quick start guide*. Available online at: http://www.casos.cs.cmu.edu/ index.php

Carley, K. M., Columbus, D., & Landwehr, P. (2013). *Automap user's guide*. Available online at: http://www.casos.cs.cmu.edu/ index.php

Carley, K. M., Dombroski, M., & Tsvetovat, M. (2003). Destabilizing dynamic covert networks. *Proceedings of International Command and Control Research and Technology Symposium*.

Choi & Lee. (2018). The Present and Future of Cybercrime, Cyberterrorism, and Cybersecurity. *International Journal of Cybersecurity Intelligence & Cybercrime*, *1*(1), 2-2.

Ezell, B., & von Winterfeldt, D. (2009). Probabilistic risk analysis and bioterrorism risk (Commentry). *Biosecurity and Bioterrorism*, *7*(1), 108–110. doi:10.1089/bsp.2009.0010 PMID:19379110

Garrick, B., Hall, J., Kilger, M., McDonald, J., O'Toole, T., Probst, P., Parker, E., Rosenthal, R., Trivelpiece, A., Van Arsdale, L., & Zebroski, E. (2004). Confronting the risks of terrorism: Making the right decisions. *Reliability Engineering & System Safety*, *86*(2), 129–176. doi:10.1016/j.ress.2004.04.003

Gumz, J. (2017). Combatting cybercrime and cyberterrorism: Challenges, trends and priorities. *Computer Review*, *58*(2), 99–100.

Hai, Z., & Xiaofeng, X. (2019). Research on Trend Analysis Model of Movement Features Based on Big Data. *International Conference on Advanced Hybrid Information Processing*, 279, 187-194. doi:10.1007/978-3-030-19086-6_21

Jason, M., & Gregory, S. P. (2011). A Comparative Analysis of PRA and Intelligent Adversary Methods for Counterterrorism Risk Management. *Risk Analysis*, *31*(9), 1488–1510. doi:10.1111/j.1539-6924.2011.01590.x PMID:21418080

Jennifer, E., & Louise, L. (2009). A Social-Cognitive Perspective of Terrorism Risk Perception and Individual Response in Canada. *Risk Analysis*, *29*(9), 1265–1280. doi:10.1111/j.1539-6924.2009.01264.x PMID:19650811

Junnan, L., Haiyan, L., & Xiaohui, C. (2020). *Terrorism Event Model by Knowledge Graph. Journal of Wuhan University*. doi:10.13203/j.whugis20190428

Karl, R., & John, G. (2008). Risk Assessment and the Terrorist. *Perspectives on Terrorism*, *2*(8), 3–10.

Lecun, Y., Kavukcuoglu, K., & Farabet, C. (2010). Convolutional Networks and Applications in Vision. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*. IEEE. doi:10.1109/ISCAS.2010.5537907

Li, W. (2017). The solution of anti-terrorism intelligence early warning information system based on human intervention factors. *Journal of Intelligence*, *3*(4), 33–40.

Majeske, K. D., & Lauer, T. W. (2012). Optimizing airline passenger prescreening systems with Bayesian decision models. *Computers & Operations Research*, *39*(8), 1827–1836. doi:10.1016/j.cor.2011.04.008

Ming L., Ren Z. (2020). Evolving a Weighted Bayesian Network for Consequence Assessment of Terrorist Attack. *IEEE Access*. .10.1109/ACCESS.2020.2993016

Novikov, A. V., & Koshkin, A. P. (2019). Identification and analysis of major risk factors causing national terrorism. *Journal of Aggression, Conflict and Peace Research*, *11*(3), 225–239. doi:10.1108/JACPR-01-2019-0402

Pate-Cornell, M., & Guikema, S. (2002). Probabilistic modeling of terrorist threats: A systems analysis approach to setting priorities among countermeasures. *Military Operations Research*, *7*(4), 5–23. doi:10.5711/morj.7.4.5

Phelps, R. (2005). *All Hazards vs Homeland Security Planning*. Available at: http://www.ems-solutionsinc.com/pdfs/ MeettheExperts.pdf

Pooja, N. J., & Archana, S. V. (2021). Analysis of Social Media Based on Terrorism— A Review. *Vietnam Journal of Computer Science*, *8*(1), 1–21. doi:10.1142/S2196888821300015

Ralph, L. (2007). Modeling values for anti-terrorism analysis. *Risk Analysis*, *27*(3), 585–596. doi:10.1111/j.1539-6924.2007.00910.x PMID:17640209

Samrat, C., & Mark, D. A. (2011). A Methodology for Modeling Regional Terrorism Risk. *Risk Analysis*, *31*(7), 1133–1140. doi:10.1111/j.1539-6924.2010.01565.x PMID:21232062

Scott, M. (2001). Coronary plaque as a replacement for age as a risk factor in global risk assessment. *The American Journal of Cardiology*, *88*(2), 8–11. doi:10.1016/S0002-9149(01)01712-X PMID:11473737

Shafieezadeh, A., Cha, E., & Ellingwood, B. (2015). A decision framework for managing risk to airports from terrorist attack. Risk Analysis, 35(2), 292-306.

Siqueira, K., & Arce, D. (2020). Terrorist training: Onsite or via the Internet? *European Journal of Political Economy*, *63*, 101878. Advance online publication. doi:10.1016/j.ejpoleco.2020.101878

Viscusi, W. K., & Zeckhauser, R. J. (2017). Recollection Bias and Its Underpinnings: Lessons from Terrorism-Risk Assessments. *Working Paper Series*, *37*(5), 969–981.

Wengang, F., & Jing, H. (2018). Early Warning for Civil Aviation Security Checks Based on Deep Learning. *Data Analysis and Knowledge Discovery*, *2*(10), 46–53.

Willis, H., Morral, A., Kelly, T., & Medby, J. (2005). *Estimating Terrorism Risk*. Available at: http://www.rand.org/pubs/ monographs/2005/RAND MG388.pdf

Willis, H. H. (2007). Guiding resource allocations based on terrorism risk. *Risk Analysis*, *27*(3), 597–606. doi:10.1111/j.1539-6924.2007.00909.x PMID:17640210

Woo, G. (2002, April). Quantitative terrorism risk assessment. *The Journal of Risk Finance*, *4*(1), 7–14. doi:10.1108/eb022949

Woo, G. (2009). *Quantifying insurance terrorism risk*. Prepared for the National Bureau of Economic Research meeting, Cambridge, MA. Available at: http:// www.rms.com/NewsPress/Quantifying Insurance Terrorism Risk.pdf

Xiaopeng, L. (2018). *Method of building terrorist network model based on social network analysis.* University of Defense Science and Technology. Available at: https:https://xueshu.baidu.com/usercenter/paper/

Xuan, G., Fei, X., Xiaozhi, T., Hongguo, Y., & Xiaoyuan, Y. (2019). Winning the War on Terror:Using Social Networking Tools and GTD to Analyze the Regularity of Terrorism Activities. *International Journal of Information Technology and Web Engineering*, *10*(4), 422–437.

Xuan, G., Zhiting, X., & Wenhu, W. (2016). ANP simulation model of bridges as potential terrorist attack targets. *Jisuanji Yingyong Yanjiu*, *34*(5), 1342–1345.

Yan, Z. (2020). *Comprehensive Risk Assessment for Civil Aviation Passengers Based on Machine Learning.* Civil Aviation University of China. Available at: https://kns.cnki.net/kns8/defaultresult/index

Yi, Z. (2019). Research on the Risk Assessment and Pprevention of Terrorist Attacks in Religious Site Based on FAHP-SWOT. *Journal of Hunan Police Academy*, *31*(2), 99–106.

Yongbao, W., Junqi, W., & Xiaohui, H. (2021). The Application Research of Matrix and Count Method in the Risk Assessment of Terrorist Suspects. *Journal of Public Security*, (1), 83–91.

Yongnan, L., & Jianwei, M. (2017). A Counter-terrorism Risk Analysis Method Based on Bayes Theory. *Journal of Intelligence*, *36*(9), 14–18.

Yujun, Z., Weiguo, S., & Xingming, S. (2017). Airline Passenger Profiling Based on Fuzzy Deep Machine Learning. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(12), 2911–2923. doi:10.1109/TNNLS.2016.2609437 PMID:28114082