Electronics, Communications and Networks
A.J. Tallón-Ballesteros (Ed.)

doi:10.3233/FAIA241381

A Knowledge Graph-Based Approach to Anti-Smuggling Intelligence Analysis

Zhibang DING[1], Pengbo ZHAO, Shuangjie LIANG and Xinmeng WANG
School of Information Technology, Nanjing Police University, Nanjing 210023, China

Abstract. Customs anti-smuggling intelligence as an important link in the anti-smuggling chain. From the massive case information in time to find effective intelligence clues, all kinds of anti-smuggling case data for automated intelligence extraction, high efficiency according to the case of time, location and other elements of the case thread; anti-smuggling case intelligence under the environment of big data for intelligent correlation, anti-smuggling cases in the intelligence of intelligent research and analysis. This is especially important for the construction of China's public security informatization in recent years, through big data technology, entity and attribute as nodes, relationship as the edge, the establishment of a semantic network based on the structure of the knowledge graph, through the integration of a large number of anti-smuggling intelligence knowledge graph structure, and further to realize the semantic relationship network of potential criminal individuals or groups predicted to stimulate the value of the existing public security big data to create a combination of the experience of public security business A knowledge graph-based anti-smuggling intelligence analysis method.This paper introduces the following aspects: firstly, we construct the spatio-temporal based expression of anti-smuggling intelligence elements to realize the logical relationship representation of anti-smuggling cases. At the same time, we construct an information extraction model combining deep learning and conditional random field driven by elements of anti-smuggling intelligence under big data environment. In addition, we further design and propose the association method of anti-smuggling intelligence based on graph convolutional neural network, so as to construct the three-dimensional anti-smuggling intelligence research and judgment model based on space and time.

Keywords: automation; big data; knowledge graph; anti-smuggling intelligence association methods

1. Synthesis of research

1.1 Technical problems in extracting and analyzing information on anti-smuggling cases under the current policing model

In the context of globalization, the value of data is more apparent, and in today's criminal smuggling activities are often accompanied by a large amount of text information data, including but not limited to customs declarations, transport

[1]Corresponding author: Zhibang Ding, School of Information Technology, Nanjing Police University, Nanjing 210023, China, Email: 3086439071@qq.com

documents, packing lists and weight lists, and commercial invoices. At the same time, these criminal activities are also involved in the smuggling of criminal suspects of basic identity information, telephone records, money flow records, as well as all kinds of clues closely linked to the data. In view of the case involves a variety of data types and the amount of information is huge and complex, the customs anti-smuggling police need to have efficient and rapid ability to extract from the massive amount of data and the case is closely related to the effective intelligence. At the same time, they also need to quickly integrate the case of the flow of funds, transportation of goods and the flow of documents and other key evidence chain, in order to quickly clarify the case line, to provide strong support for the detection of cases. Each of these key aspects of our ability to extract and analyze anti-smuggling information in the era of big data has put forward strict and necessary requirements.

1.2 The necessity of improving the efficiency of intelligence extraction and research and judgment in anti-smuggling cases

Existing intelligence analysis systems face challenges when dealing with anti-smuggling intelligence, because anti-smuggling intelligence has a huge amount of data and contains complex and diverse information factors, which makes the acquisition of intelligence clues extremely difficult. For a long time, how to automatically extract and correlate intelligence clues from huge information sources has been a difficult problem in the research of anti-smuggling intelligence intelligence. Therefore, with the rapid growth of anti-smuggling case information, improving the efficiency of intelligence extraction and research and judgment has become a key technical problem in the field of anti-smuggling intelligence that needs to be solved urgently.

2 Technological route

2.1 Constructing an efficient anti-smuggling intelligence extraction model

At present, the extraction of unstructured data faces certain difficulties. In order to obtain more contextual information in the text of various cases, in this session, we mainly focus on the textual data, and propose a method combining deep learning model and conditional random field for extracting intelligence elements in Chinese text with respect to the natural language expression characteristics of the key elements, such as time, place, case, and character-related relationships.

2.1.1 Selection of textual information on various types of anti-smuggling cases and construction of a database

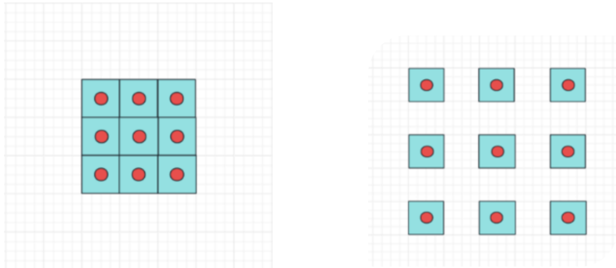
Determine the scope of the sample pool, such as the types of anti-smuggling cases to be included, the time frame, the geographical scope, etc. When collecting data, cooperate with anti-smuggling departments to obtain data and consult specialized legal databases, academic literature and anti-smuggling-related statistical yearbooks. After in-depth organization and cleansing of data, it is necessary to choose an appropriate data storage solution. We can consider designing a collection function with data source type of database, supporting mainstream databases including Oracle, MySQL, SQLServer, etc., and supporting a variety of mainstream big data platforms, including Hadoop, FusionInsight, MaxCompute and other three and more mainstream big data platforms [1]. This step is to make full preparation for the next lexical annotation work.

2.1.2 Perform lexical processing of textual content in case events and label their lexical properties

Firstly, the data are manually labeled with intelligence elements to obtain high-quality lexical labeling, and the labeled element categories are based on the proposed intelligence element expression model, thus obtaining the labeled dataset. Based on the labeled training set and the unlabeled case text corpus, various tools are used to perform unsupervised training and learning on the word-split text, and the word vectors and word vectors obtained from the preprocessing of the intelligence elements are used as inputs to the model.

2.1.3 Expanded Convolutional Neural Networks (ENNs) for obtaining more contextual information in text.

In deep learning, the value of each unit of the feature map is determined by a region of the convolutional input, which is called the receptive field. The perceptual field is the size of the pixels in each layer of the CNN's feature map that are mapped onto the original image. The neurons of a CNN cannot perceive all the information of the original image because in these structures both convolutional and pooling layers are used, and the two layers are locally connected to each other. The larger the convolution kernel of the model, the more holistic and comprehensive information the neurons can extract; however, varying the convolution kernel also implies a dramatic increase in the number of parameters. Dilated convolutional neural network can solve this problem well [2]. Dilated convolution is more advantageous relative to ordinary convolution as in Fig. 1(a), especially due to the nature of dilated convolution in terms of sensory field and multi-scale feature capture. Dilated convolutional neural networks are designed by introducing a series of voids such that each convolutional kernel by introducing a series of voids, each convolutional kernel is able to cover a wider area, providing a larger receptive field as shown in Fig. 1(b). This advantage is very important for us to preprocess the word vectors and word vectors of intelligence elements, because the target may be affected by the surrounding environment. Expansive convolution, on the other hand, usually does not introduce additional parameters while increasing the sensory field. This helps to keep the model size small and provide better computational efficiency in the face of a large amount of textual information, while the lightweight and efficient computation of the model is crucial for real-time performance.



(a) (b)

Fig. 1. Structure of ordinary convolution and dilation convolution

2.1.4 Construct convolutional neural network and conditional random field model to realize entity recognition and feature information extraction.

The process of entity recognition, feature information and correlation extraction through convolutional neural network and conditional random field model is shown in Figure 2. Through the convolutional neural network can automatically learn the text representation, and during the training process, the convolutional neural network will learn to automatically extract useful feature information from the original information to further improve the learning and extraction efficiency. The text analysis method based on convolutional neural network can be applied to different kinds of text analysis tasks, such as text categorization, sentiment analysis, translation and so on. It can be selected for different anti-smuggling information extraction and research and judgment needs, which provides sufficient flexibility for different anti-smuggling cases. Meanwhile, with the introduction of Conditional Random Field (CRF) technique, the model can contain arbitrary features, and it does not require significant adjustments in the model configuration to achieve competitive results [3]. The convergence speed of the generated model is significantly improved. When dealing with a large number of samples, the model can more quickly approach the real data distribution. This significant advantage is especially obvious when dealing with a large number of anti-smuggling samples obtained in the pre-processing step. Due to the faster convergence of the model, the entity recognition and feature information extraction will become more accurate and reliable. This undoubtedly lays a solid foundation for the construction of knowledge graph and analysis of intelligence, which makes the whole process of intelligence processing more efficient and accurate.

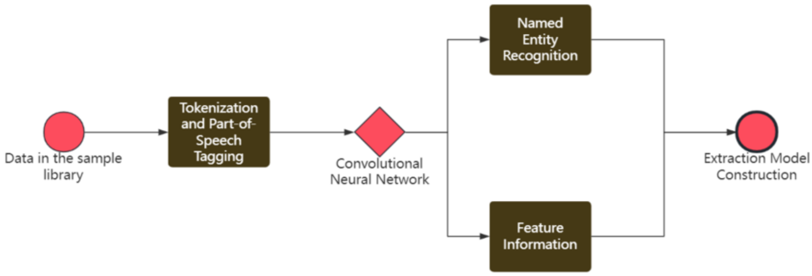


Fig. 2. Flow chart of anti-smuggling intelligence extraction model construction

2.2 Construction of an anti-smuggling knowledge map

Modern knowledge graph techniques mainly include data integration, data cleaning, data storage, data querying, data mining and knowledge reasoning. The generation of the entire knowledge graph requires the collection of data from different sources into a whole. The next step is to preprocess, transform and validate the data. After data cleaning, the cleaned data is further stored in databases, file systems or other storage devices, and the stored data is retrieved, searched and analyzed. Through data mining, from a large amount of data to find the hidden patterns, laws and relationships contained therein (i.e., used to find all kinds of anti-smuggling cases behind the mode of operation, the law of operation and the relationship between the suspects). Finally, after knowledge reasoning, new clues are derived from the knowledge graph. The advantage of modern knowledge mapping techniques is that they can handle large-scale, multi-source and uncertain knowledge. In connection with the analysis of anti-smuggling intelligence, the entity identification and feature information extraction process described above can be used to generate a unique anti-smuggling knowledge map for visualization and analysis.

2.2.1 Generating an anti-smuggling knowledge map based on graph convolutional neural networks

Currently, we can combine graph convolutional neural networks to build knowledge graphs [4]. Typically, Convolutional Neural Networks (CNNs) receive as input graph data such as images with a Euclidean structure, which can be characterized by a convolutional kernel (or filter) in Euclidean space. Given the regular structural characteristics of an image, the convolutional kernel is able to translate over it in order to extract features point by point. The core of CNN is the use of convolutional kernels, which act like small windows, sliding over the input image to capture image features through convolutional operations. However, when faced with an irregular topology, The number of nodes varies, the neighbors of each node are different, and the nodes are often associated with each other - the traditional CNN can not effectively deal with this kind of graph structure. At this time, we use the graph convolutional network (CGN).

A comparison of knowledge graph generation was carried out using two models, CNN and CGN. After experimental validation and based on two key evaluation metrics, recall and F1, the results show that CGN exhibits higher applicability in generating knowledge graphs compared to CNN. As shown in Figures 3 and 4:

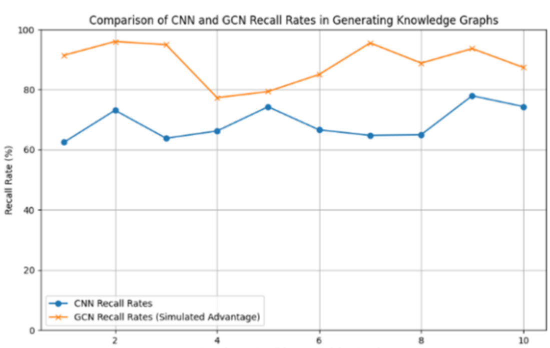


Fig. 3. Comparison of recall rates

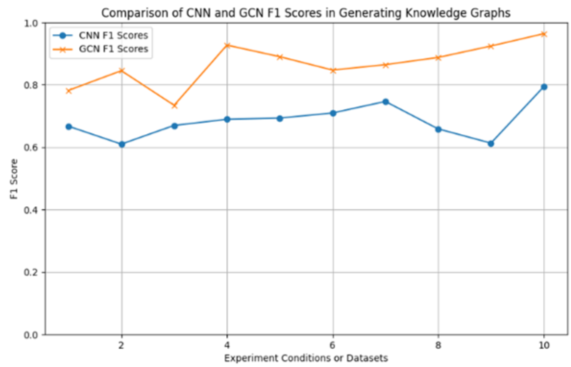


Fig. 4. F1 Score Chart

This allows each node in our graph to change its state from time to time due to the influence of its neighbors and points further away until it is finally in equilibrium, and the closer the neighbors, the greater the influence. The closer the neighbor, the greater the influence. The nodes learn the characteristics of each node through the influence of each other and the training of the W-parameter matrix. Applied to the actual analysis of anti-smuggling intelligence, we can generate multiple entity sub-nodes containing different anti-smuggling intelligence entity features as shown in Figure 5,

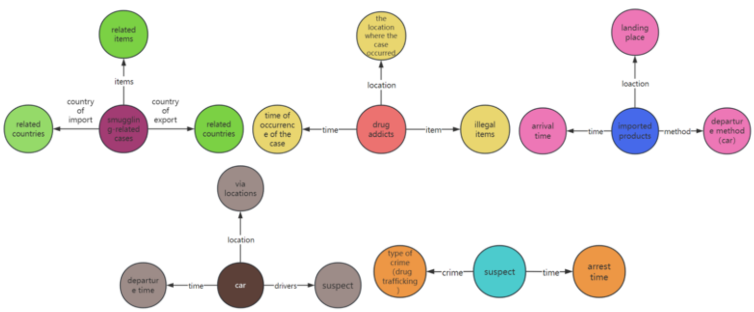


Fig. 5. Entity sub-nodes containing different anti-smuggling intelligence entity features

The joining of different anti-smuggling intelligence entity feature sub-nodes will continuously exchange the neighbourhood information to update the features of each node until a stable equilibrium is reached, at which time the feature vectors of all the sub-nodes contain the information of their neighbouring nodes, which can then be used for clustering, node classification, link prediction and so on. When examining the vertices and the rest of the nodes in the subgraph, these nodes carry the spatio-temporal attributes of the anti-smuggling knowledge units and the case-specific feature information. The following section further elaborates on the anti-smuggling knowledge acquisition and graph generation methodology with a sample example, the sample data is as follows.

Time	Space	Event Description
XX/X/20XX	not quite clear	Time of commission, suspects used multinational smuggling routes for drug smuggling
XX/X/20XX	International Container Terminal	Cargo ship XXXX docked
XX/X/20XX	customs(i.e.border crossing inspection)	Customs clearance of cargo ships
Afternoon of XX/X/20XX	not quite clear	Commencement of unloading of transport (suspect vehicles)
Unknown	one location	Captured the suspect vehicle stopping information and arrested the suspect in the vehicle
XX/X/20XX	one location	Several arrests of drug users

(Colours marked in the table are different sub-entities)

A graph convolutional neural network was trained to predict the connectivity between vertices in subgraphs using tabular textual information data of the sample type (see Fig. 6 for details).

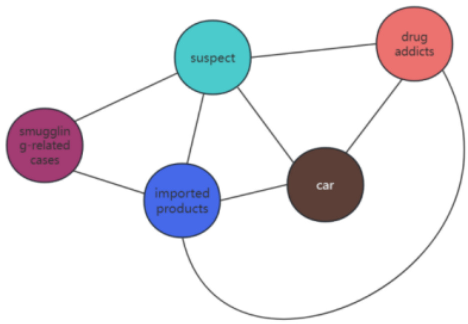


Fig. 6. Graph of connectivity between vertices in the predicted subgraphs

Immediately after that, by synthesizing and associating the information of the vertices in the subgraphs, we successfully built a knowledge graph focusing on anti-smuggling cases (see Fig. 7 for details). This result shows that this kind of anti-smuggling knowledge map can clearly display anti-smuggling related knowledge at multiple levels of abstraction, such as individuals, objects and events, and can outline the trajectory of each element over time according to the changes of time and place, thus effectively serving the practical needs of anti-smuggling intelligence analysis.

Fig. 7. Knowledge map of cases in the anti-smuggling category

2.3 Criminal Intelligence Analysis and Mining Using Knowledge Graphs

This project utilizes knowledge map and through a series of orderly steps, realizes the intelligent processing of anti-smuggling case and event data in the process of anti-smuggling intelligence research and judgment, including "overall planning - intelligence collection - intelligence assessment - intelligence organization - intelligence analysis - intelligence research and judgment - intelligence feedback". Intelligence analysis - intelligence research and judgment - intelligence feedback". In this process, we will be the time and space elements as the core of the event, centered on the case itself, especially focusing on documents, goods and funds and other dimensions of the evidence chain of comprehensive analysis and research and judgment [5].

Figure 8 below shows the process of anti-smuggling intelligence research and judgment based on the knowledge map proposed by this project. This project firstly focuses on the event data of the whole smuggling case, extracts useful knowledge and text information from it, and imports it into the anti-smuggling intelligence database. Through the above process, it realizes the intelligent association and fusion of big data in multiple dimensions, and further constructs a visualized knowledge graph that associates a large amount of textual information of past smuggling cases. Visual presentation of the knowledge map, improve the quality and efficiency of planning decision-making in the past to store data is mainly in the form of data tables, data visualization can enhance the efficiency of data processing and organization, but this way is difficult to structure the storage of good knowledge of the type of data. In addition to conveying factual information, knowledge visualization can be used to transfer, reconstruct, memorize and apply knowledge. In wartime, using images to express semantic relationships and network structures in long texts can help commanders quickly clarify the pattern laws of battlefield data, analyze the content of intelligence and its intrinsic correlation, and improve the speed and quality of planning and decision-making [6]. This visualization of the knowledge map not only facilitates the anti-smuggling department police to quickly access the case and event information needed for research and judgment, but also to further guide and clarify the direction of the subsequent investigation. Subsequently, the police can be based on the verification

Fig. 8. Flowchart of intelligence research and judgment based on knowledge graph

process to obtain the documents, goods and money clues, the knowledge map for continuous updating and supplementation, until the complete sorting out of the case of the flow of funds, goods and documents flow and other evidence chain. This process greatly contributes to the rapid clarification of the case, providing strong support for the successful detection of the case.

3 Concluding remarks

Based on the construction of anti-smuggling knowledge map, this paper designs a more efficient anti-smuggling intelligence analysis and judgment method. The method is designed to effectively solve the complexity and diversity of anti-smuggling intelligence information factors, and analyze and judge the process is long and other shortcomings, to build a set of intuitive knowledge map based on the intelligence research and judgment process.

This paper utilizes a variety of databases combined with expansion neural network and introduces conditional random field model to identify multiple entities and extract multiple feature information from a large amount of anti-smuggling intelligence knowledge. These entities and feature information are prepared for the construction of knowledge graph in the next step. Subsequently, a graph convolutional neural network method is applied to predict the connection relationship between vertices in subgraphs and display the anti-smuggling related knowledge at multiple levels of abstraction. At the same time, according to different time and place, it can give the spatial and temporal development trajectory, and carry out comprehensive intelligence analysis and judgment. This method effectively solves the pain point problem of the research and judgment of smuggling crime information, and provides a better research and judgment method for the possible future occurrence of multi-factor smuggling cases and events, so that the anti-smuggling police can fully and efficiently carry out the research and judgment of information, saving manpower and material resources. At the same time, through continuous improvement and optimization, this technology will be further developed, and its potential will be explored, contributing scientific and technological power to the development of public security anti-smuggling cause.

Acknowledgements

This research was supported by the 2024 College Students Innovation and Entrepreneurship Training Program (Grant No. 202412213002Z).

References

[1] Gu MS. Research on constructing government information resource base library based on non-relational database [J]. Modern Information Science

[2] Gao ZJ, Cao HD, Han ZH. Anomaly detection model based on expansive convolutional neural network[J]. Journal of Shenyang University of Architecture (Natural Science Edition),2024,40(04):738-744.

[3] Yu BG, Fan ZZ. A review of research on conditional random field modeling for natural language processing[J]. Journal of Information Resources Management,2020,10(05):96-111. DOI:10.13365/j.jirm.2020.05.096.

[4] Shen XK, Li Y, Chen JW, et al. A recommendation algorithm fusing collaborative knowledge graph and graph convolutional network[J]. Computer Technology and Development,2024,34(01):150-157.
[5] Zhang MJ. Application of knowledge mapping in criminal intelligence analysis[J]. Law and Society,2021,(04):82-83+112. DOI:10.19387/j.cnki.1009-0592.2021.02.038.
[6] Wen W. Research on the application of knowledge graph in intelligent intelligence analysis[J]. China Military to Civilian, 2024, (10):25-27.