A decision system based on active perception and intelligent analysis for key location security information

Jingzhao Li and Zihua Chen*

College of Electrical and Information Engineering,
Anhui University of Science and Technology,
Huainan, China
Email: jzhli@ aust.edu.cn
Email: czh1619@126.com
*Corresponding author

Guangming Cao

China University of Mining and Technology,
Beijing, China
Email: 124122692@qq.com

Mei Zhang

College of Electrical and Information Engineering,
Anhui University of Science and Technology,
Huainan, China
Email: mzhang@aust.edu.cn

Abstract: In various enterprises, the security data are manually entered by multiple security workers at different times, which can lead to disordered data in the system, and also, analysis decision files need to be manually processed. To solve these problems, this paper presents a decision system for key location security information based on active perception and intelligent analysis. First, the designed system is developed based on C/S framework. Then, a deep learning perception model is constructed. Experimental results show that this system can significantly reduce the heavy workload of security inspectors, and reduce the frequency of safety accidents.

1   Introduction

The security information supervision of key location is an extremely important part of the safety production process. It can determine the level of safe production and enterprises management. For example, the level in coal mines is related to the lives of miners. Chinese government has established many management measures and regulations for safe production (e.g., in the coal mine safety production, regulations of coal mine safety were issued).

In recent years, in response to the call of Chinese government, related enterprises attach great importance and increased investment in terms of security control, but accidents still occur frequently. After nearly half a year of research and analysis, most accidents are due to safety hazards without timely treatment (Axelsson, 2017; Yonge et al., 2017; Zhou et al., 2014). The enterprise department expects the new intelligent, security risks awareness, and decision-making early warning systems (Erban and Gorelick, 2016; Oppliger et al., 2017; Zhou et al., 2017; Viana and Sato, 2014), which can automatically extract semantic feature information, and intelligently conduct security risk decision analysis, to prevent and reduce the occurrence of security incidents in advance.

At present, the enterprise security information system still has the following problems: data description is inaccurate and non-standard, e.g., Conde-Clemente et al. (2017) explain that the data description is inaccurate; data processing analysis performance is poor, e.g., embedded hardware technology is used to process data by Heo and Jeon (2017); background database fields are not completed enough, e.g. in the DSS framework, the methods of advanced data analysis is made to cost estimate and make decision (Siryani et al., 2017).

To study how to accurately and quickly collect and extract semantic features from tens of thousands of data, proactively perceive security risks, analyse decisions and give early warnings, Anhui University of Science and Technology, Coal Mine Industry Association Productivity Promotion Center, and Tongmei Group Co. Ltd. in Shanxi jointly conducted research.

The structure of the paper is organised as follows: Section 2, the relevant work progress of the intelligent active sensing and decision system is analysed. Section 3, the composition of the security information sensing system is constructed, and the intelligent perception and decision model is analysed. Section 4, the active learning method based on deep learning is provided. Section 5, the security information analysis model based on ACO is introduced. Section 6 the experimental analysis is applied. Section 7 is the application case analysis. Finally, a summary and prospects are given in Section 8.

2   Related works

This section reviews existing work on active perception and intelligent retrieval of security information. After reviewing a large amount of literature and summarised it, our team finds that the research on the data processing and management system still has room for development (Althebyan et al., 2016; Tawalbeh et al., 2016), and intelligent analysis decision-making system of security information is still in its infancy (Agrawal et al., 2017; Hamilton et al., 2016). According to the current research methods, there are three main analysis aspects including: semantic extraction based on deep learning, semantic extraction of keyword matching and similarity retrieval and database normalisation.

2.1   Semantic extraction methods based on deep learning model

Most of the semantic extraction methods based on deep learning mainly rely on building deep learning algorithms and combining machine learning with artificial intelligence. The algorithms add the part-of-speech weights to construct the semantic feature extraction analysis model [the commonly used deep learning methods include back propagation (BP) neural network, adaptive learning algorithm, etc.]. Oh et al. (2017) introduce a semantic analysis model based on DNN. The model is used in native-language classification. He et al. (2017) create a model based on deep learning EMCNN to enhance semantic sentiment analysis. The K-NN regression is used to predict the load data of coal-fired power plants by Agrawal et al. (2016), but the lead time is too short. In addition, we propose a method based on CNN automatic perceptual semantic extraction.

2.2   Semantic extraction of keyword matching and similarity retrieval model

The semantic extraction of keyword matching method can be built by several ways. The similarity is calculated by algorithm for the semantic feature extraction of Weibo, which is studied by Zhang et al. (2017, 2018). Liu et al. (2015) introduced the research of the term semantic level extraction for Chinese text information processing in a specific field. Lo et al. (2016) provided a semi-supervised method to solve the problem of inaccurate dictionary keyword matching.

Hameed et al. provided a way to classify the recommended systems by calculating similarity (Kumar et al., 2017). Other scholars have also conducted research on keyword similarity matching models (Hameed et al., 2012 Beheshti et al., 2017; Zouaq et al., 2017; Ding et al., 2016). Among them, Kumar et al. (2017) proposed a weighted semantic information extraction algorithm idea by extracting nouns and verbs from the marker data, then providing different possibilities for each noun or verb to extract semantic-based terms; next these weights were established using cosine similarity. Finally, the weights of similar sentences were associated with the training sentences in the structured semantic weight relationship algorithm, and the keyword intelligent semantic extraction was performed by the classifier. This algorithm solves the problem of inaccurate query and slow retrieval speed for a

single keyword, but it has limitations on the keyword intelligent semantic extraction. The word retrieval order is fixed and cannot be fully applied to the keyword retrieval problem of enterprise security risk information.

2.3 *Database normalisation model*

The database normalisation model can solve the problem of non-standard data in databases. In Wang (2014), a web database optimisation recognition query model is proposed, based on intelligent semantic autocorrelation feature for the inaccurate query of web database data, which includes intelligent query, data error correction, weight elimination and integration, and greatly improved query accuracy of the data, but it does not involve the multi-attribute field interconnection analysis function in the client, and only processes on the database module, and the background server retrieval function is not strong.

In view of the above works, an intelligent semantic analysis sensing model based on CNN (Tang and Wu, 2017) and ACO frequency intelligent retrieval and decision analysis has been constructed. After large amounts of data in the database are trained, the automatic intelligent analysis and retrieval of security information data is based on this. The active perception and intelligent analysis decision-making system based on C/S for critical location security information is proposed.

3 Safety information sensing system

The security information sensing system consists of a client and a server. The client is responsible for interacting with the user. The background server is used to process data. The structure is divided into security information data management module, intelligent information sensing module, high frequency regulation display module, and intelligent decision warning module. The structure of the sensor control system is shown in Figure 1.

- The security information data management module includes these functions: data entry, data modification, and data display. The security workers input the security risk data into the background server database through the client to display on the client, which is convenient for security workers management and storage of relevant security regulations data.

- The intelligent information perception module is accorded to establish a similar keyword database about the similarity of semantics and based on the intelligent collection model of CNN. When we collect security risk information, the system automatically maps the standard semantic keywords in the standard database, which is according to the semantic similarity of the matching autonomously identify and perceive keywords in the standard library for retrieval, so as to accurately find violations of security regulations.

- The main function of the high-frequency regulation display module is to analyse the frequency of the data of a certain period of time, the details of the problem, the safety regulations of the violation, etc. Through the server-side algorithm and the retrieval information, whose intelligent extraction frequency exceeds a certain value, is displayed on the system client. When entering data, the interface enables security workers to know the frequent occurrence of related security risks through the intelligent system.

- The intelligent decision-making early warning module mainly realises users accessing the server-side database. After the system processes the data according to some multi-attribute interconnection conditions, the system obtains the data decision analysis in the form of graphs and briefings, performs pre-judgment and realises the function of intelligent decision-making early warning.

**Figure 1** Safety information sensing system structure
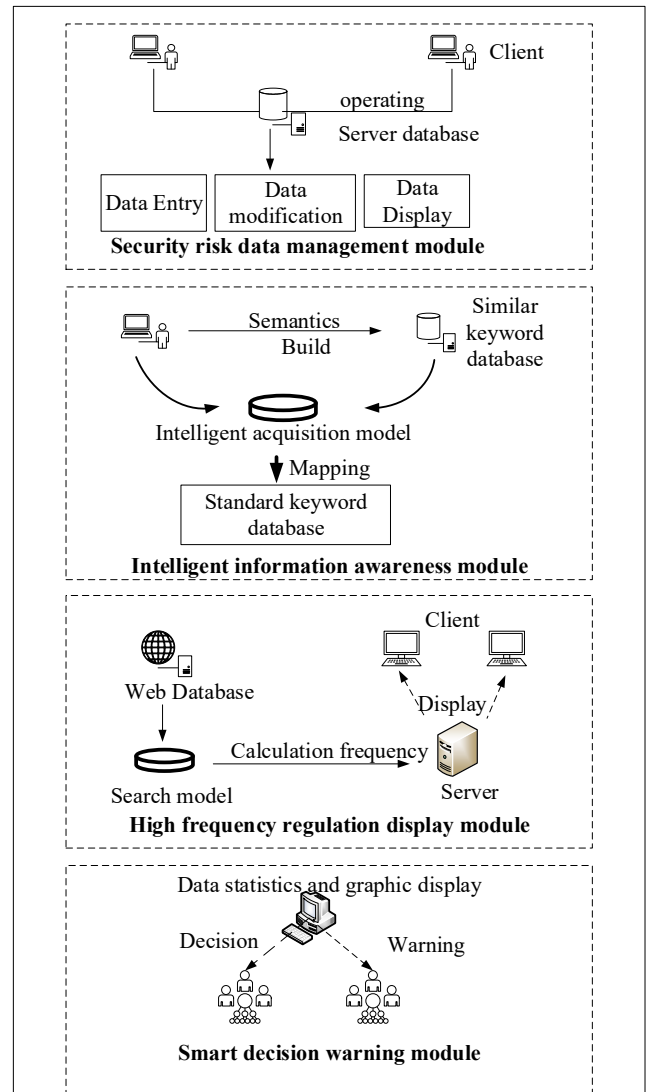
3.1 Intelligent information perception module

The security information data management module is used by the security worker to add data to the server database through the client. The basic information such as the inspection time, area, category, problem details, inspectors, and number of times of rectification is entered, and the related data is modified or deleted, and other basic operations are included in the module through the background database. The server side dynamically fills the check box option of the client interface by calling the corresponding database table through the background program, and the security worker can directly select the option in the client interface.

3.2 Intelligent information perception module

The intelligent information perception module is shown in Figure 2. When the previous data entry, the security information data in the database constitutes the physical database of the system for the staff. After establishing a series of operations such as initial object, semantic analysis, algorithm design and instantiation object, the CNN-based active perception model for training database is constructed. After the data is completed, the database is mapped to the standard keyword database in a many-to-one relationship.

Figure 2   Safety information sensing system structure (see online version for colours)
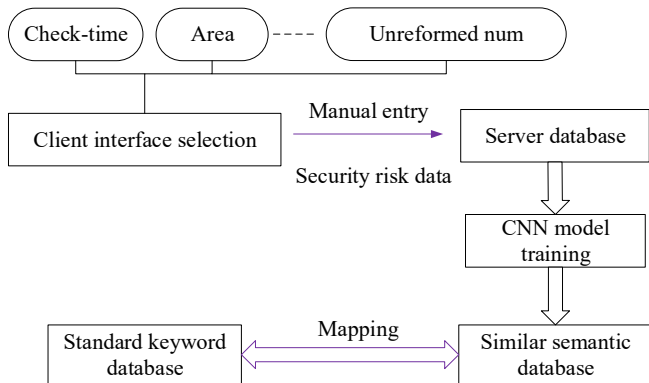
3.3 High frequency regulation display module

The high frequency regulation display module structure is shown in Figure 3. Through the improved CNN algorithm of the previous module, the active keyword model is trained to obtain the mapped standard keywords. The server calls the standard keywords inferred by the smart collection model, and performs precise query of the security regulations to obtain high-accuracy retrieval results. At the same time, it accesses the local server. In the session, through the improvident ant colony optimisation (ACO) frequency retrieval model, the security rules with higher retrieval frequency are obtained, and the positive feedback is used to strengthen the logo and display, which is convenient for the staff to carry out the next information. The selection rules provide a reference for the

decision-making work of the intelligent decision-making module.

Figure 3   Intelligent frequency retrieval structure (see online version for colours)
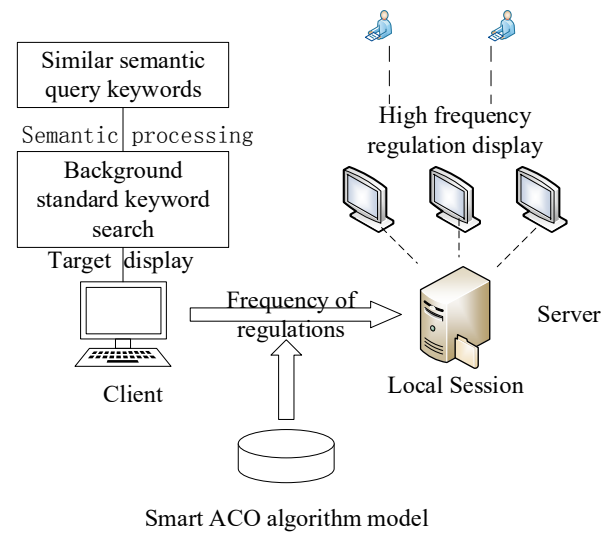
Figure 4   Wisdom analysis decision module structure diagram
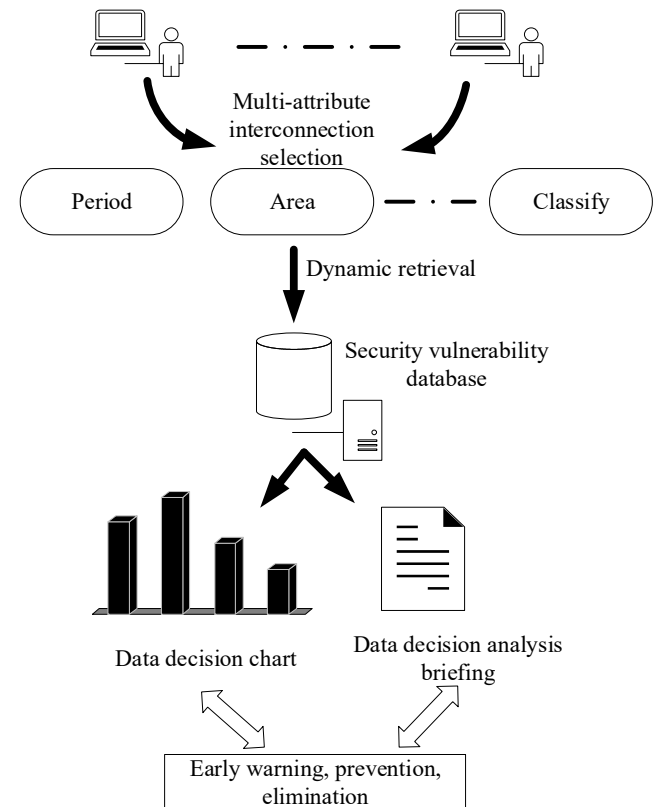
3.4 Smart decision warning module

The structure of the intelligent decision warning module is shown in Figure 4. After the enterprise security worker selects the multi-attribute interconnection query condition group in the system retrieval interface, the system performs database data retrieval, updates the data in real time, and collects the data of the hidden danger problem by real-time

statistical multi-attribute chained search data, which are used to calculate the period of time. The degree of violations includes 'general hidden dangers', 'heavy hidden dangers', and 'red line hidden dangers', using data decision statistics to clearly and concisely reflect the changing trend of security problems from tens of thousands of data (e.g., in this system, it can realise multi-attribute joint query, and can display the trend of various categories of 'year', 'the certain area' and 'some month' through data statistics chart) and the decision analysis of multiple template formats generated by the background system through the program briefing, which can achieve all aspects of intelligent analysis of various security issues violation results, timely warning, in order to develop measures to timely rectification, reduce or eliminate the occurrence of security incidents.

4 Active learning method based on deep learning

The essence of the intelligent information perception model is to establish a similar semantic keyword mapping model based on improved CNN algorithm. The main idea of the model construction is first to establish a matching matrix $(x_1, x_2, \ldots, x_n)$, and by using the representation of the vector, the input keyword and standard words in the similar-semantic database are changed into a vector consisting of 0 and 1 by using cosine similarity, and according to the rule base. Degree calculation formula to define the similarity between words and words, and then construct the spatial position of the two words into a two-dimensional structure, the process is similar to a planar image, and then use the convolutional neural network (CNN) to model, Next, design the convolution kernel for feature extraction and self-sampling, generate multiple filters in the convolution module. Sample them through the pooling module, reduce the dimension by max pooling. Finally, use the excitation function and based on multiple 2D convolution modules and pooling. The module constructs a perceptual network in the fully connected module, and represents the final data as a matrix of n*1. Also, obtains the similarity between the final input keyword and the semantic similar keyword, and compares the most similar one to perform the corresponding standard in the mapping standard database keywords, just as the e-learning systems are used to analyse learning files and provide advice for social learning (Weng et al., 2016).

The model based on the improved CNN algorithm steps are as follows:

1 Firstly, let the input keywords be changed into vector coordinate forms according to rules of the rule library, and the vector coordinate is $x(x_1, x_2, \ldots, x_n)$. After the reference keywords are processed by the rule library, the vector coordinate is $y(y_1, y_2, \ldots, y_n)$. Then, until each keyword in $\{x\}$, $\{y\}$ has been processed, the words vector similarity is expressed by the formula (1) and λ is a constant scaling factor.

$$\alpha_i = \lambda * \sum_1^n (x_i * y_i) \bigg/ \sqrt{\sum_1^n x_i^2} * \sqrt{\sum_1^n y_i^2} \qquad (1)$$

2 According to the dot product operation of equation (1) and the finest-grained matching information, a matching matrix is created as shown in Figure 5.

3 First, each node of the matching matrix is numbered, and the ith row and the jth column element of the image are represented by $\alpha_{i,j}$; each weight of the filter is labelled, and the weight of the ith row and the jth column is represented by $x_{i,j}$ and $\omega_{i,j}$ indicates the element of the ith row and the jth column of the feature map; and the activation function is represented by f, which is obtained by the formula (2).

$$\alpha_{i,j} = f\left(\sum_{m=0}^m \sum_{m=0}^m \omega_{m,n} x_{i+m,j+n} + \omega_b \right) \qquad (2)$$

Figure 5 Matching matrix

(a)   (b)
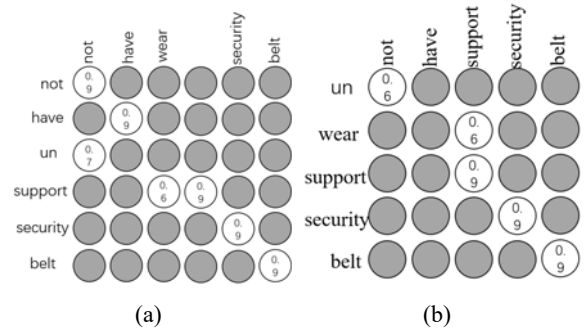
As shown in Figure 5(a), the element $a_{0,0}$ in the upper left corner, the convolution kernel is (1, 0, 1, 1, 1, 0, 1, 0, 1), the step size = 1, the bias = 0, the convolution The value is:

$$
\begin{aligned}
a_{0,0} &= f\left(\sum_{m=0}^2 \sum_{n=0}^2 \omega_{m,n} x_{i+m,j+n} + \omega_b \right) \\
&= \text{Relu}\left(\omega_{0,0} x_{0,0} + \omega_{0,1} x_{0,1} + \ldots \omega_{2,2} x_{3,3} \right) \\
&= \text{Relu}(1.2 + 1.2 + 1.7 + 0) = \text{Relu}(4.1) = 4.1
\end{aligned}
$$

Calculate all the elements of the convolutional module-feature map, then sample it in the pooling module (3 * 3), and use max pooling to obtain the feature map after sampling. Finally, after processing the fully connected module, you can get one-dimensional semantic matching similarity with comparison, which through the keyword id in the database to obtain the standard keywords to be retrieved. For example, the matching degree finally obtained in Figure 5(a) is 0.89, and the matching degree obtained in Figure 5(b) is 0.72. Comparing the two values, you can get 'not have wear support security belt' and 'not have support security belt', which the most similar standard keywords in the database are mapped on the condition that 'not have wear support security belt'.

Figure 6   Semantic similarity training chart based on CNN
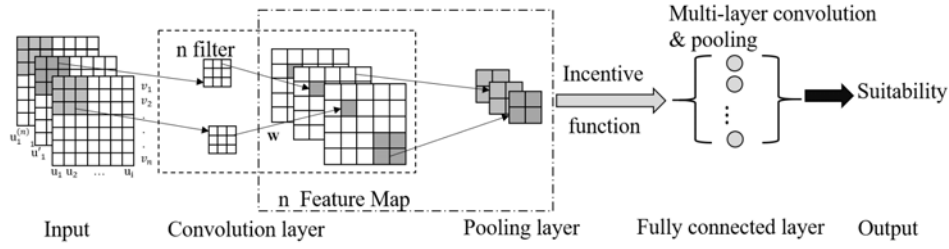
The essence of the security information analysis model is based on the improvement colony optimisation. The main idea is that, when the state is initialised, all the security regulations in the database are marked as empty. When the security worker retrieves a security rule, the system automatically add a pheromone to this security regulation, and then start from the virtual source node through the ACO. Next, broadcast to all the aggregation nodes after each retrieval, and next spread from the aggregation node to the child nodes. Finally, all information of the previous process will be recorded in the child node information table (INF-ant), which is corresponding to the aggregation node of security regulations. The information includes last hop node address, node depth, current node pheromone strength, pheromone volatilisation factor ρ, and pheromone intensity Q, the number of times of retrieval (frequency). If the current node sends negative feedback information (local optimal path) to the source node, the system will positively enhance the node path in the high pheromone, realise positive feedback, and finally get the strongest path of the pheromone. Use the high frequency retrieval rule to record the session table to the server and display it to the client front-end interface in real time. It is selected when all workers are entered, and it also provides reference conditions for later decision-making, and can make some predictions from the chart and briefing.

The model based on the improved ACO algorithm steps as follow:

1   Node discovery phase: Before the initial retrieval, build the aggregation node set $S\{S_1, S_2, \ldots, S_n\}$, the child node set is $D\{D_1, D_2, \ldots, D_n\}$. Set G as an aggregate information amount, Ma is the total number of retrieved entries, and M is the total number of current regulations data, then the amount of information is calculated as shown in equation (3).

$$G = M_a / M \tag{3}$$

Starting from the virtual source node, the initial value is $h = 0$. After each retrieval, if the node is searched last time, the depth is unchanged. Otherwise, the node depth is increased by 1, and the depth h of each child node is updated in real time; If the degree of the corresponding regulation $n_{i,j}$ is increased by 1, the

frequency $Q_{i,j}$ of the security regulations is retrieved as shown in equation (4).

$$Q = n_{i,j} / N \tag{4}$$

2   Reverse detection phase: by starting from the current maximum leaf node, through the established NF-ant table, find the node of the last hop, find the node with the largest degree at this time, repeat the operation, and continuously feedback, and finally the node information on the local optimal path is transmitted to the sink node.

3   Positive reinforcement phase: after receiving the information in the INF-ant table, the aggregation node adjusts the path with the largest amount of pheromone in the current feedback according to equation (5). $\Delta_{i,j}$ is the unit time pheromone change. And $\tau_{i,j}$ is the pheromone intensity.

$$\tau_{i,j} = \tau_{i,j}^{prev} + \rho \Delta_{i,j} \tag{5}$$

Considering the three aspects of pheromone strength, regulation frequency, and aggregate information, the calculation formula of the influence factor of the high-frequency path is according to (6):

$$\delta_{i,j} = \tau_{i,j} * Q_{i,j} * G \tag{6}$$

By constructing the model, the high-frequency search rules are acquired and displayed in the form of an array in descending order. And the model flow chart is shown in Figure 8.

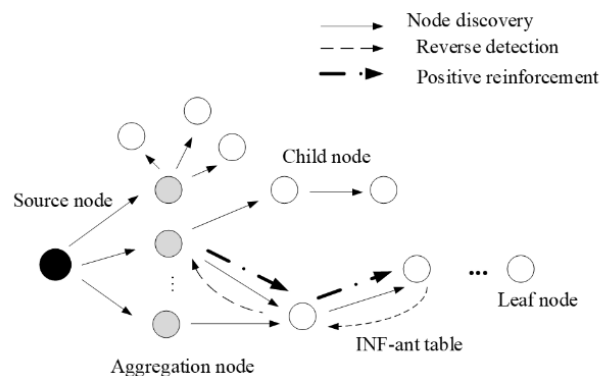**Figure 7**   ACO-based retrieval model structure diagram

Figure 8 Model flow chart based on deep colony optimisation
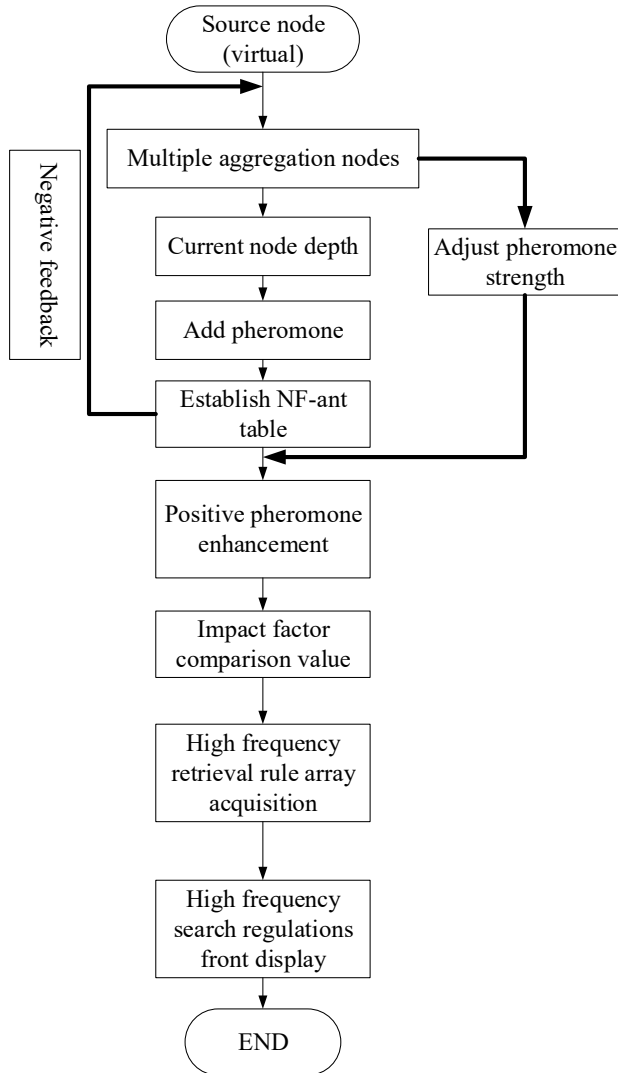
Improved ACO algorithm.

Input: Security risk data information

Output: High frequency safety regulations

Begin

1: Init();//Initialise each sink node

2: Foreach(if in S)

{

3: Init();//Initialise the currently retrieved regulations

4: Switch(){

5: Case :h=h;//Node depth is unchanged

6: Break;

7: Case :h=h+1;

8: Break;

}

9: Include();//Call the custom method Include() to save thenode on the current local optimal path

10: Save(G) ;

//Call the custom Save() method to save the information in the INF-ant chart

11: table and calculate the impact factor

}

12: Max ();//Get the maximum impact factor by Max method

13: Strengthen(S,D); //Positive reinforcement of pheromones the Strengthen() method

14: GetSession(Max);//Obtain the combination of safety regulations

corresponding to the maximum impact factor

15: Create Array[i];//Get the array of security rules combined to the corresponding frequency in order of size

16: Print { Array[i];} //Output display

End

The algorithm 1 gives the flow of the security information analysis model, and mainly includes three parts:

1 The first part, the steps 1–8, is used to initialise the structure tree and save the discovery node. It judges whether the node is included in the tree. And then calculates the depth of the node. Until the retrieval is terminated the conditional statement stops. The time complexity of the first step is $O(k)$.

2 The second part, the steps 9–13, is used to save a directed graph of all nodes the node, and get the local weight maximum path, saved in the INF-ant table. Next, strengthen the frequency retrieved value. The time complexity of the second step is $O(k^2)$.

3 The third part, the steps 14–16, is used to enter the node values on the high-weight path as an array and display it on the client interface for data decision. The time complexity of the second step is $O(k)$.

6 Experimental analysis

According to the CNN intelligent semantic perception model in the Section 4, we made two experiments:

• The experimental data collection is 20,000, which are divided into ten groups. The data will be compared by the system of active sensing of intelligent information with intelligent analysis and decision-making system and weighted semantics system based on keyword similarity matching, and information extraction algorithm. The result of first experiment is shown in Table 1.

Among them, the *Accuracy* number (*aver*) is the average number of correct retrieved regulations. Time (*aver*) is the spent time of retrieving the correct regulations. Accuracy is the ratio of the *Accuracy* number and two thousand. The formula of the *Accuracy* (*aver*) is:

$$Accuracy \ (aver) = \left(\sum_{i=1}^{10} (Average \ number)/2{,}000 \right) \Big/ 10$$

- The experiment time is 2 hours and 10 groups in total. Through the three systems, the average accuracy of the algorithms, the number of actual security regulations and actual correct security regulations are compared in the same time. The compared results of second experiment are shown in Table 2.

Among them, *Correct-accuracy* (*aver*) is the average ratio of the *Actual-correct number* (*aver*) and *Actual number* (*aver*). The *Actual* number (*aver*) is the average number of retrieved regulations. And the *Actual-correct number* (*aver*) is the data of the correctly retrieved regulations. The formula of the *Correct accuracy* (*aver*)

Correct-accuracy(*aver*)

$$= \left(\sum_1^{10} (Actual\text{-}correct number) \middle/ Actual\text{-}number \right) \middle/ 10$$

7 Application case analysis

By applying the intelligent acquisition and intelligent decision-making system to the coal mine safety inspection, the data of the mine within six months is processed. The data statistics of the safety hazard problems in January-June 2018 in the relevant area are obtained. We randomly sample two people's data from ten security workers. The actual applications are shown in Figure 9, Figure 10 and Figure 11.

Figure 9 Coal mine overall safety data chart

After nearly half a year, we collected about 28,000 data in total. The overall safety hazard data of coal mines are shown in Figure 9. The data show that the hidden dangers of coal safety accidents account for a large proportion of the safety management and the water control problems, which indicates that the management capacity of coal mine workers is not strong and insufficient. Secondly, the safety hazards in tunnelling and auxiliary transportation have also reached the problem. At the tenth level, the proportion of the remaining categories is low. After using this system, the trend is shown in Figure 10 and Figure 11. And they show that the number of potential safety hazards in the mine is

generally declining, which indicates that the intelligent system can prevent the occurrence of safety accidents and reduce the frequency of coal mine safety accidents. Secondly, it can be concluded from Figure 10 and Figure 11 that the proportion of safety hazard data in safety management of the mine is too large, and the safety hazards of water prevention and control in the mine are also frequent, although other types of problems account for a small proportion. The overall situation of the mine is also a problem that needs attention. The mine has insufficient management of the normative operation of the miners, and the staff safety management perception is insufficient. The coal mining enterprises use this intelligent collection and decision support system to analyse the problems existing in the mine, strengthen the training on relevant safety operations, regularly convene the relevant staff to meet and summarise, formulate rectification measures, standardise the safety management operations, and reduce the occurrence of related categories of hidden dangers.

Figure 10 Statistical analysis of safety hazard data

Figure 11 Statistical analysis of safety hazard data

Table 1 The average retrieval accuracy and time cost of each algorithm model (the same amount of data)

Model	Algorithm	Accuracy (aver)/%	Accuracy number(aver)/bar	Time(aver)/h
Intelligent semantic perception model	Improved CNN intelligent security information perception algorithm	94%	1,880	4.6
Weighted semantic information extraction model	'T' structure semantic weight relationship algorithm	85%	1,700	6.7
General keyword matching model	Keyword matching algorithm	78%	1,560	8.9

Table 2 The average retrieval accuracy and retrieved number of each algorithm model (the same time)

Algorithm	Model	Correct accuracy (aver)/%	Actual number(aver)/bar	Actual correct number(aver)/bar
Improved CNN intelligent security information perception algorithm	Intelligent semantic perception model	93%	490	456
'T' structure semantic weight relationship algorithm	Weighted semantic information extraction model	82%	410	337
Keyword matching algorithm	General keyword matching model	72%	320	231

8 Conclusions and future work

In this paper, our contributions mainly include the following three aspects.

1 First, through the construction of security information data management module, intelligent information sensing module, high-frequency regulation display module and intelligent decision-making early warning module, the active perception and intelligent decision-making system of security risk information is formed.

2 Second, the system uses the improved CNN algorithm to select the closest keywords from the database and map them to the standard keyword database, which solves the problems about the keyword matching accuracy of similar semantics being not high. Then the ACO model is constructed to realise that the retrieved high-frequency regulations are intelligently displayed in the interface, which is convenient for the staff to select.

3 Finally, through the real-time processing statistics of database data, high-intelligence multi-attribute interconnection query, the paper constructs clear and concise data statistics, intelligently analyses data statistics, and formulates rectification measures, which achieves timely warning of safety incidents, and prevents the occurrence of safety hazards.

The experimental results show that the proposed method can extract course semantics well. The intelligent extraction semantic algorithm can not only improve the semantic integration of a course, but also provide theoretical support for semantic searching on some online system. Artificial intelligence and machine learning have been gradually applied to computer software development. This system will continuously improve the system's self-learning ability. We believe that in the future, it will be able to realise high-end intelligent systems that integrate coal mine safety hazard collection, intelligent decision-making and early warning analysis, which can serve coal mine safety production well.

References

Agrawal, V., Agrawal, S., Nag, S. et al. (2016) 'Application of K-NN regression for predicting coal mill related variables', *International Conference on Circuit, Power and Computing Technologies*, IEEE, pp.1–9.

Agrawal, V., Panigrahi, B.K. and Subbarao, P.M.V. (2017) 'Intelligent decision support system for detection and root cause analysis of faults in coal mills', *IEEE Transactions on Fuzzy Systems*, Vol. 25, No. 4, pp.934–944.

Althebyan, Q., Alqudah, O., Jararweh, Y. et al. (2016) 'A scalable map reduce tasks scheduling: a threading-based approach', *International Journal of Computational Science & Engineering*, Vol. 14, No. 1, pp.44–54.

Axelsson, J. (2017) ''Safety in vehicle platooning: a systematic literature review', *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18, No. 5, pp.1033–1045.

Beheshti, S.M.R., Benatallah, B., Venugopal, S. et al. (2017) 'A systematic review and comparative analysis of cross-document coreference resolution methods and tools', *Computing*, Vol. 99, No. 4, pp.313–349.

Conde-Clemente, P., riviño, G. and Alonso, J.M. (2017) 'Generating automatic linguistic descriptions big data', *Inf. Sci.*, Vol. 380, pp.12–30.

Ding, J., Dong, Y., Gao, T. et al. (2016) 'Sentiment analysis of Chinese micro-blog based on classification and rich features', *Web Information Systems and Applications Conference*, IEEE, pp.61–6.

Erban, L.E. and Gorelick, SM. (2016) 'Closing the irrigation deficit in Cambodia: Implications for transboundary impacts on groundwater and Mekong River flow', *Journal of Hydrology*, Vol. 535, pp.85–92.

Hameed, M.A., Jadaan, O.A. and Ramachandram, S. (2012) 'Collaborative filtering based recommendation system: a survey', *IJCSE*, Vol. 4, No. 5, pp.859–876.

Hamilton, M.C., Doody, P., Doody, P., Bates, M.E., Bauer, N.L., Voyadgis, D.E. et al. (2016) 'Web-based geospatial multiple criteria decision analysis using open software and standards', *International Journal of Geographical Information Science*, Vol. 30, No. 8, pp.1667–1686.

He, Y-X., Sun, S-T., Niu, F-F. and Li, F. (2017) 'A deep learning model enhanced with emotion semantics for microblog sentiment analysis', *Chinese Journal of Computers*, Vol. 40, No. 4, pp.773–790.

Heo, G. and Jeon, J. (2017) 'A study on the data compression technology-based intelligent data acquisition (IDAQ) system for structural health monitoring of civil structures', *Sensors*, Vol. 17, No. 7, p.1620.

Kumar, C.S.S., Mohanapriya, M. and Kalaiarasan, C. (2017) 'A new approach for information retrieval in semantic web mining involving weighted relationship', *International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, Coimbatore, pp.1–4, doi: 10.1109/ICIIECS.2017.8276095.

Liu, L., Li, Z., Zhang, X. et al. (2015) 'A review of semantic information processing of Chinese web texts', *Journal of Computer Applications*, Vol. 32, No. 1, pp.6–10+16.

Lo, S.L., Cambria, E., Chiong, R. et al. (2016) A multilingual semi-supervised approach in deriving Singlish sentic patterns for polarity detection', *Knowl.-Based Syst.*, Vol. 105, No. 8, pp.236–247.

Oh, Y.R. et al. (2017) 'A deep-learning based native-language classification by using a latent semantic analysis for the NLI shared task', *The Workshop on Innovative Use of Nlp for Building Educational Applications*, pp.413–422.

Oppliger, R., Pernul, G. and Katsikas, S. (2017) 'New frontiers: assessing and managing security risks', *Computer*, Vol. 50, No. 4, pp.48–51.

Siryani, J., Tanju, B. and Eveleigh, T. (2017) 'A machine learning decision-support system improves the internet of things' smart meter operations', *IEEE Internet of Things Journal*, Vol. 4, No. 4, pp.1056–1066.

Tang, Y. and Wu, X. (2017) 'Scene Text detection and segmentation based on cascaded convolution neural networks', *IEEE Transactions on Image Processing*, Vol. 26, No. 3, pp.1509–1520.

Tawalbeh, L., Haddad, Y., Khamis, O. et al. (2016) 'Efficient and secure software-defined mobile cloud computing infrastructure', *International Journal of High Performance Computing & Networking*, Vol. 9, No. 4, pp.328–341.

Viana, P. and Sato L. (2014) 'A proposal for a reference architecture for long-term archiving, preservation, and retrieval of big data', *IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications*, Beijing, pp.622–629.

Wang, H. (2014) 'Database entity recognition query based on intelligent semantic autocorrelation feature', *Microelectronics & Computer*, No. 5, pp.159–162,166.

Weng, M., Hung, J.C., Weng, J.D. et al. (2016) 'The recommendation mechanism for social learning environment', *International Journal of Computational Science & Engineering*, Vol. 13, No. 3, pp.246–257.

Yonge, A.V., Swenor, B.K., Miller, R., Goldhammer, V., West, S.K., Friedman, D.S. et al (2017) 'Quantifying fall-related hazards in the homes of persons with glaucoma', *Ophthalmology*, Vol. 124, No. 4, pp.562–571.

Zhang, S., Wei, Z. et al. (2018) 'Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary', *Future Generation Comp. Syst.*, Vol. 81, pp.395–403.

Zhang, S., Zhang, S.Y. et al. (2017) 'The recommendation system of micro-blog topic based on user clustering', *MONET*, Vol. 22, No. 2, pp.228–239.

Zhou, X.Y. et al. (2014) 'The peril of fragmentation: security hazards in android device driver customizations', *Security and Privacy IEEE*, pp.409–423.

Zhou, Y., Su, W., Ding, L., Luo, H. and Love, P.E.D. (2017) 'Predicting safety risks in deep foundation pits in subway infrastructure projects: a support vector machine approach', *Journal of Computing in Civil Engineering*, Vol. 31, No. 5.

Zouaq, A., Gagnon, M. and Jean-Louis, L. (2017) 'An assessment of open relation extraction systems for the semantic web', *Inf. Syst.*, Vol. 71, No. 12, pp.228–239.