Design of Database Teaching Model Based on Computational Thinking Training

https://doi.org/10.3991/ijet.v14i08.10495

Xue-peng Huang (✉), Jing Leng
Hubei University of Police, Wuhan, China
Hubei Collaborative Innovation Center of Digital Forensics and Trusted Application, Wuhan, China
hxp@hbpa.edu.cn

Abstract—This paper proposes to combine knowledge acquisition and thinking training in database course teaching, and designs a database-teaching model based on computational thinking training. This model takes computational thinking as the core, and through the close combination of multiple levels of thinking ability and database teaching content, it promotes the improvement of students' thinking ability, trains students to feel and experience the role of thinking in the process of knowledge learning, and guides students to exercise and train their thinking ability independently, spontaneously and consciously in learning. A method of pre-processing before training model implementation is designed, and similar items are merged to design a processing scheme.

Keywords—Database teaching; model design; computational thinking

1 Introduction

Database technology represents one of the fastest growing fields in computer science and technology. Since the 1960s, database has attracted great attention owing to its structured data, minimal redundancy, high program and data independence, easy expansion and easy programming. Meanwhile, the computational thinking contained in database technology provides an effective way for researchers to expand thinking and innovate methods.

However, in database study, we shall not consider thinking alone, but also consider the receiver, that is, considering knowledge system structure of the undergraduate students and knowledge needs of the receiver at the same time. To this end, how to organically combine knowledge learning with thinking training should be the first issue to be considered. A balance between the two should be found. On the one hand, we should not only pay attention to students' ability to master knowledge points in books, but neglect their thinking training and improvement, because this will make students lose the ability to improve their computational thinking through course learning. On the other hand, we should not only focus on how to train students'
thinking, but ignore the basic requirements for students to master knowledge points. Only by organic combination of knowledge acquisition and thinking training in the course teaching can we achieve our final curriculum design goals.

2 Database Course Teaching Analysis

2.1 Database Course Analysis

From the perspective of course content, database itself is by no means an easy-to-learn course, which involves a number of intertwined subject theories and applications. It contains quite rich knowledge contents, including theoretical knowledge of data model, cognition of database system structure, basic theory and methods of relational databases, structured query language, basic knowledge of relational systems and theoretical methods of query optimization, relational data theory and application, database design and database system implementation, database recovery theory and technology, concurrency control, database security, distributed database, parallel database and a variety of object-oriented basics. Meanwhile, database learning involves analysis and design of information systems, use of programming languages and other related contents. Since it involves a wide range of contents, it poses high requirements for students' learning ability, learning interest and logical thinking ability.

Database is by no means a course of great interest. Most theories contained in it are abstract, so for undergraduate study in the initial stage, most theories are not only very specific and trivial when applied to practice, but also show poor applicability, which easily results in students' boredom and fear of difficulty. As a result, students feel baldness and obscurity in learning, especially for some students lacking solid knowledge base. In this way, students lack learning interest and enthusiasm in this course and have fear for difficulty.

Therefore, in the process of teaching, knowledge system should be combined with thinking system. On one hand, correlation between knowledge systems should be considered. On the other hand, graduality of thinking system should be considered to help students integrate isolated knowledge points and scattered ways of thinking, thereby forming an organic whole. In this way, students not only have a higher level of understanding in knowledge structure, but also constantly experience the role of computational thinking in computer science learning in the process of knowledge learning.

2.2 Clue Analysis in Database Course Knowledge

Database principle can be divided into foundation, design and implementation. The general clue is basic principles of database as well as design and application knowledge learning in practice.

**Database Basics and Abstract Training:** The basic knowledge of database gives a basic description of data model and conceptual model. In particular, model
definition is introduced to define and analyze three-level model structure of the database system and the derived concept of second-level image model, and data independence is proposed. From this, we can clearly see that abstract computational thinking mode plays a key role in it. In this part of the study, it is clear that the key point of abstraction should be focused. In knowledge progression, accumulation and understanding should be made on abstract knowledge ability such as patterns and models. In thinking, the emphasis should be put on training of students' abstract ability, especially involving concepts such as patterns and models. It is not only necessary to let students learn how to move from concrete to abstract, but more importantly, let students' thinking shift from the real world to the abstract world, so that students look at our data world in "abstract vision".

Regardless of basic relational model, relational model, or relational algebra or relational calculus, it is not only necessary to let students learn basic concept of relational models, patterns and relational algebra, but also consider two aspects of problems: First, its connection with the previous pattern and model; second, further training of abstraction from the perspective of thinking. In addition, in introduction of relational algebra, it is not only necessary to make accurate reasoning and calculation from mathematical point of view, but also relate relational algebra to the core theoretical knowledge of relational database, make analogy, improve their mathematical thinking and computational thinking.

For the relational database language SQL, on the one hand, it is necessary to train students' abstract ability, abstract the two-dimensional table in the student's eyes into a conceptual entity; on the other hand, train the students' automation ability. That is, using high-speed processing power of the computer, try to abstract actions into concrete instructions for the computer to execute. At the same time, it is necessary to train students to explore computing power of computers and always think "What is computable?"

**Relational System and Computational Thinking:** The definition of relational system is based on basic concepts and operations of database. As for the criterion of full-relationship system, it is generally taken as optional course content from the point of view of teaching content. However, we find in the teaching process that its learning and understanding can not only improve students' comprehensive understanding of relational system, but more importantly, play an important role in elevating students' level of thinking to the mode of computational thinking. The content of its twelve criteria is unified in a certain sense with basic requirements and methods of computational thinking. Similarly, for thinking training, mathematics or abstract thinking training helps students build and improve their computational thinking ability.

The learning effect of query optimization should be a concentrated expression of the training degree of student's computational thinking. As a teaching content, query optimization plays an extremely important role in relational database systems. It can be said that the wide acceptance and application of relational databases, the great success of SQL language are closely associated with relational query optimization. Optimization not only constitutes the basic premise for users to accept the relational system, but also enables relational system to directly analyze query semantics through
expression of higher semantic level. Regardless of any perspective, query optimization is highly abstract and highly automated.

Relational data theory focuses on database design, or directly speaking, on the issue of logical design, which is an issue highly relevant with abstraction. In this environment, students must use abstraction to solve problems in an abstracted world based on a high degree of abstraction ability, and their thinking ability will be improved. On the background and platform of relational model, students form standardization theory step by step, which finally promotes their computational thinking in building standardization theory of relational database. The theoretical construction process is actually the improvement and application process of thinking.

**Database Design and Computational Thinking Ability:** Database design is to define a certain practical application environment, train students to construct an optimal database model using the above knowledge, then establish a database system, and finally efficiently save and maintain data by designing and implementing an application system. In database design, the core content is to establish a database and apply database system, which represents a core technology in the development and construction of information systems.

Due to the complexity of database application system, it is not only to support the operation of related programs, but also to display and link the abundant data, symbols, needs and actions of the real world in the abstract world. Therefore, database design can never be completed overnight, and more is a process of “repeated exploration and gradual optimization”. Such planning and structuring process is a training process of thinking. For example, a simple ER diagram is actually correspondence between data in the real world and the virtual world. It is more necessary to comprehensively consider subsequent data relations, data processing and a series of other contents. Different conversion results directly lead to different efficiency and correctness of the entire database system. The maturity degree of computational thinking is the key to ensuring correctness and completeness of the conversion.

Similarly, in the process of physical database design, it is necessary to select different access methods for the relational model to ensure data processing efficiency under the premise of understanding and mastery of factors affecting physical database design. The index access method, hash access method, or aggregate access method can be selected according to different processing environments. Training of thinking skills must run through the process.

Physical storage structure of database is the next problem to be solved. We need to understand the role of database storage devices, files and file records, Hash files, and index files in database design and implementation. Abstract data dictionary is an important concept to realize the intermediate link and media role in database analysis, and provide basic query content and interpretation, so as to make the data flow diagram display more clearly in structured analysis. Due to the shareability of database, databases are often used by multiple users. These user programs can be executed serially one by one. It is noting that only one user program runs at a time to perform access to the database, meanwhile other user programs must wait until the end of the user program to access the database. If a user program involves input/output exchange of abundant data, the database system is idle for most of the
time. Therefore, in order to make full use of database resources, multiple users should be allowed to access the database in parallel. However, this will result in concurrent access of the same data. If the concurrent operations are not controlled, incorrect data may be accessed and stored, destroying database consistency. Therefore, database management system must provide concurrency control mechanism to resolve concurrency issues through a blocking mechanism.

Finally, database recovery technology should be considered. After failure occurs; database recovery technology can restore the database data from the error state to a certain logically consistent state. The key to this method is how to establish redundant data and use redundant data to implement database recovery.

2.3 Analysis of Computational Thinking in Database Learning

**Decomposition and reduction:** When dealing with computer problems, we need to consider that the whole and the part are a pair of opposite, but not a rigid concept. The whole is reflected as part in some aspects, and if the part is the object of decomposition, it should be considered as a whole. That is, we should know that “whole” and “part” can be transformed into each other. It is necessary to learn to make full use of conditions, decompose problems, and divide the “whole” into “parts”. Meanwhile, it is necessary to learn to elevate so that "part" can also be solved by "whole".

**Abstraction:** To model the relevant aspects of a problem, it is necessary to construct a mathematical model, let constructive thinking play a realistic role in dealing macro aspect of the problem, concisely and descriptively depict system characteristics and behavior.

**Conversion and simplification:** There are two forms of conversion: one is to narrow the target by dimensionality reduction; the other is to reversely magnify the target by raising dimension. The idea of narrowing the target includes high-dimensional reduction to low-dimensional, general reduction to specific, abstraction reduction to concrete, overall reduction to local and data relationship simplification. In conclusion, data relationships can be simplised based on the original data structure.

**Classification and division ideas:** Classification is a logical method that suggests concept extension and also a basic method of thinking for solving problems. Division represents a special classification, which is usually associated with recursion.

**Reverse thinking:** The two forms of reverse thinking usually manifest as reverse thinking of result reason, and reverse thinking from negative to positive.

Reverse thinking of result reason means that if the condition A is satisfied, the conclusion B can be obtained, then the inference of B by A is result reason; conversely, the process of inferring A by B is known as reverse thinking of result reason.

**Assumptions:** The first is similar association, i.e. analog association, which refers to association with similar existing knowledge and problem-solving experiences aroused by its form and nature. By virtue of understanding of an object, it compares similarity with another object to achieve speculative understanding of the latter, so it
is a cognitive method of transition from understanding of one type of object to another.

The second is inductive association, which refers to a form of reasoning that leads to general conclusions through analysis of special cases. First, we will list some special simple small-scale data, observe the law of the data, guess the general conclusion according to some data, and finally prove the correctness of the conjecture by reasoning or mathematical induction. In a sense, this is also a partial abstraction.

The third is backtracking association. Backtracking direction should be clarified, that is, determine the current algorithm is wrong or not optimal, while remaining the importance of the original algorithm. In this direction, we return to preliminary analysis, start rich association and rigorous rational thinking, redefine question and find a breakthrough. Finally, based on the analysis of the similarities and differences between old and new problems, we guess the improvement method and verify correctness of the new solution.

3 Data Preprocessing

3.1 Teaching content preprocessing

The content of database course is quite rich. Speaking of the knowledge points, it covers knowledge of many related disciplines and courses. For different knowledge points, the requirements and contents of computational thinking or other scientific thinking are also different. In the handling process, if the relevant knowledge points are not pre-processed, it will not help scientifically grasp and effectively cultivate students' computational thinking. Therefore, before implementing the training model, pre-processing must be carried out according to the characteristics of the relevant knowledge points, so that similar items are merged, and then the processing scheme is designed.

Considering the "occurrence" characteristics of knowledge points, in the implementation process, relevant algorithms can be designed and relevant contents can be merged and processed by analyzing its statistical characteristics.

For teaching contents that may be involved in undergraduate database system course, knowledge characteristics are preliminarily defined from the perspective of computer science knowledge system. The primary content of thinking training is analyzed from the perspective of scientific thinking, and its time characteristics are analyzed from the perspective of knowledge acquisition and thinking training. Teachers can make independent assembly and analysis according to school characteristics, student characteristics, course objective, teaching effect, then refine personalized contents and goals, and eventually process and optimize them according to the training model.
3.2 Data preprocessing algorithm

For each specific teaching project, we have to decide whether choose thinking training or not in the teaching process; while for the teaching receiver, that means whether teaching content “appears” or not. Similarly, due to changes in teaching objects, teaching objectives, teaching contents, etc., when different teachers deal with the same teaching knowledge points, the focused thinking training points will change accordingly, so it can be known that the sample space formed by arrangement and combination of above two tables has quite rich connotations. However, we cannot customize a specific arrangement applicable to all objects and targets. Instead, the teacher makes judgment and selection according to the actual situation.

From this perspective, before implementing the model, it is necessary to process the data, merge and simplify the data, which is also an important manifestation of computational thinking.

**Related definitions:**

**Definition 1:** A sequence of events that occur successively at random times is referred to as a random event stream. For the time length $t$, it is divided into $n$ equal parts, and each time interval is recorded as $\Delta t_i, i = 1, 2, ..., k$.

Suppose the probability of occurrence of event $A$ in each time interval $\Delta t_i$ is $p$, and the probability of non-occurrence is $1-p$, then the number of occurrences $X(t)$ of event $A$ within the time length $t$ is known as binominal event stream, which is recorded as: $X(T) \sim b(n, p)$.

For binomial distribution, there is:

$$P\{\xi = k\} = C_n^k p^k (1 - p)^{n-k}, \quad k = 0,1,2,3,\cdots, n$$

(1)

Assume that $b(k, n, p) = P\{\xi = k\}$, then the distribution satisfies:

1) Non-negativity, i.e. $b(k, n, p) \geq 0$

2) Normativity, i.e.:

$$\sum_{k=0}^{n} b(k, n, p) = \sum_{k=0}^{n} C_n^k p^k (1 - p)^{n-k} = [p + (1 - p)]^n = 1$$

**Definition 2:** If the event stream is stationary, no aftereffect and common, then it is referred to as a Poisson event stream (Poisson stream).

If the probability of success $p$ in the test $B$ is small, and the rare event $n$ is sufficiently large, then in $n$ times of repeated $B$ tests, the number of occurrences of the rare event approximately obeys a distribution, which is the Poisson distribution.

The integer part $[\lambda]$ of the parameter $\lambda$ in Poisson distribution is exactly the most likely number of occurrence of rare events.
For a Poisson stream, the number of occurrences $X(t)$ of an event Accidence during the interval $t$ should obey Poisson distribution with parameter $\lambda t$ ($\lambda > 0$, $\lambda$ is constant).

That is:

For all $\lambda (\lambda > 0$, $\lambda$ is constant),

$$P(X(t) = k) = \frac{\lambda^{k} e^{-\lambda t}}{k!}$$

For Poisson stream, the number of occurrences of events (such as traffic accidents) in any time interval $(0, t)$ obeys Poisson distribution with the parameter $\lambda t$. $\Lambda$ is known as the intensity of Poisson stream.

Therefore, as far as parameter analysis of Poisson distribution is concerned, it is a one-parameter discrete distribution whose distribution mean $\lambda$ and distribution variance $\sigma^2$ are uniform. From the distribution pattern point of view, Poisson distribution generally shows asymmetry morphological characteristics. When $\lambda$ is not big, Poisson distribution shape is skewed. However, when $\lambda$ is gradually increased, Poisson distribution curve rapidly shows a symmetrical trend, gradually approaching to normal distribution, as shown in Figure 3-3.

According to the analysis, if the value of $\lambda$ is greater than 20, we can use Poisson distribution area to approximate normal distribution. At this time, the Poisson distribution can naturally be treated as a normal distribution; when the value $\lambda$ exceeds 50, we can regard Poisson distribution as consistent with normal distribution. Figure 3.3 shows the morphology of Poisson distribution curve when the value of $\lambda$ is 1, 3, 6, 10. It can be seen that for a smaller value of the population mean $\lambda$, skew characteristic is more obvious, and as $\lambda$ increases, the crest moves to the right and the distribution tends to be symmetrical, as shown in Figure 1.
Chi-square Fitting Test

Record $A_i = \{X(t) = i\}, i = 0, 1, \ldots, n$, then in the $k$ observations, $A_0, A_1, A_2, \ldots, A_n$ respectively show $k_0, k_1, k_2, \ldots, k_n$ frequencies, and:

$$\sum_{j=1}^{n} k_j = k$$

The test hypothesis $X(t) \sim p(\lambda)$ can be converted to:
Where:

Finally, is tested using distribution:

When is established, it approximately obeys distribution with n-1 degree of freedom.

The rejection region of is:

If , then , and

Then:

3.3 Preprocessing algorithm

Approximation with poisson distribution: In the process of thinking selection statistics, if is recorded as the probability that event A appears in a test, it is related to the total number of tests (a set of statistical values), if \( \lim_{n \to \infty} np^n = \lambda > 0 \), then:

If \( \lambda_n = np^n \), then \( \lim_{n \to \infty} \lambda_n = \lambda \), and \( p_n = \frac{\lambda_n}{n} \)

Then:

\[
b(k; n, p_n) = \binom{n}{k} \left( \frac{\lambda_n}{n} \right)^k \left( 1 - \frac{\lambda_n}{n} \right)^{n-k}
\]
That is: positive integer \( k \), \( n \in (n-1)(n-k+1) \ldots \frac{k!}{n!} \cdot \left( 1 - \frac{\lambda}{n} \right)^{n-k}
\]

\[
\frac{\lambda^k}{k!} \left( 1 - \frac{1}{n} \right) \ldots \left( 1 - \frac{k-1}{n} \right) \cdot \left( 1 - \frac{\lambda}{n} \right)^{n-k}
\]

\[
\rightarrow \frac{1}{k!} \lambda^k e^{-\lambda} (n \rightarrow \infty)
\]

That is: \( \forall \) positive integer \( k \geq 0 \),

\[
\lim_{n \rightarrow \infty} b(k, n, P_n) = \frac{\lambda^k}{k!} e^{-\lambda}
\]

That is, for an event stream selected from a course content, if it meets:

1) \( p \) is small enough;
2) The number of selections is sufficient;

Then, we use Poisson stream with intensity \( \lambda = np \) for approximation.

Certainly, we can also deduce:

If \( X \) obeys Poisson distribution, i.e.:

\[
P_X (x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}
\]

A normal distribution can be used to approximate a Poisson distribution.

Algorithm thought: In the aspect of the choice of teaching contents, it is clear from principle that we have inferred that the above conditions are met for the event stream, which means that we can use a Poisson stream in description.

From the previous section, it can be theoretically determined that the content selection event obeys Poisson distribution. For several types of knowledge and thinking and consolidated statistics data, recurrence is possible as follows:

\[
P(X_1 + X_2 = n)
\]

\[
= \sum_{k=0}^{n} P(X_1 = k) \cdot P(X_2 + X_1 = n | X_1 = k)
\]

\[
= \sum_{k=0}^{n} P(X_1 = k) \cdot P(X_2 = n-k | X_1 = k)
\]
That is, the content options should match Poisson distribution after consolidation.

**Algorithm Description:** According to the above analysis, this paper preprocesses each type of statistical data using classification statistical algorithm based on chi-square test.

Algorithm: Classification statistics algorithm based on chi-square test

**Input:** Statistical data

**Output:** Statistical distribution characteristics

**Parameters:** Threshold 1 (used to determine whether to merge data) \( \text{th1} \), threshold 2 (used to determine whether to perform data rejection) \( \text{th2} \), time window \( t \)

**Method:**
- Statistical data;
- \( t=100 \);
- \( \text{seg}[i]=0 \); //record the minimum statistical interval
- \( \text{mean}[i]=0 \); //record the detection mean
- \( i=0; \)
- for each crime_type
- \( i++; \)
- if statistical data>\( \text{th1} \)
- for \( (j=1; j<30; j=j+14) \)
  - Chi-square testing(Poisson); //chi-square test of Poisson distribution
  - \( \text{seg}[i]=t; \) //record the minimum interval that satisfies the distribution test
  - \( \text{mean}[i][j]=\text{mean}; \) //record the minimum interval that satisfies the distribution test
- end for
- else
  - if statistical data>\( \text{th2} \)
Merged into a new category; //group rare events into one category
else
  Reject this type of event; //reject very rare events
endif
endif
efor
i++; ;
for (;t>30; t-=14)
  Chi-square testing(Poisson); //Chi-square test of Poisson distribution for merging events
  seg[i]=t; //
endfor
segmin=100;
for each seg[i]
  if seg[i]<segmin
    segmin=seg[i] ; //minimum distribution interval
  endif
proc find_means;
endfor
seglen=100-segmin ; //interval span
return;
procedure find_means;
Input: mean vector
Output: optimal statistical interval
Method:
avr=0;
len[];
for each pass_seg [i] //only for upcheck interval
  avr=seg[i]/pass_seg; //averaging
  len[i]=find max_j; //find the interval value closest to avr according to the interval from small to large
endfor
classification len; //cluster len to find the largest similar interval in len
return opti_len;

4 Design of Database Teaching Model based on Thinking Training

According to the above analysis, establishment and training of scientific computational thinking is the inherent requirement for computer teaching in the current era. Thinking training model must be introduced into course teaching and continuously improved. It should be said that the current teaching contents mainly stress continuity of knowledge and progressiveness of knowledge thinking. Mathematical thinking and computational thinking with higher abstract requirements
are still not systematic, which are generally only reflected in several knowledge points.

Considering the needs of students' thinking training, we set up a database teaching model based on computational thinking training. This database model takes computational thinking as the core and combines database teaching contents through multiple levels of thinking ability including abstraction, automation, recursion, separation, heuristic reasoning, data protection and so on. As a result, the thinking training can be carried out subtly in the process of knowledge learning, thereby achieving the purpose of thinking improvement.

In the teaching process, it is necessary to give full play to interactive learning between teachers and students, actively combine theory with application, scientifically unify technology with application, thus finally realizing the construction goal of computational thinking, as shown in Figure 2.

Figure 2 shows the database teaching system and computational thinking training model we designed. The model can be divided into two areas: the knowledge system and thinking training system. The core of knowledge system is the basis and carrier of thinking training.

Thinking training relies on knowledge acquisition and learning. The periphery is a thinking training system. It can be seen that thinking training is comprehensive training, rather than single training that isolates computational thinking. Considering that the level of human thinking is progressive and interrelated, it is impossible for a kind of thinking to accomplish in an action through a course in accordance with our assumptions.

Therefore, we combine a variety of thinking and multiple abilities in training, not only promoting improvement of students' thinking ability in the process, but also directing students to feel and experience the role of thinking in the process of knowledge learning, and guiding students to independently, spontaneously, conscientiously train their thinking ability.

Considering that the receptors are engineering students, we also consider the basic requirements of “excellent engineer training” in the design model, and strive to integrate students' thinking ability and engineering ability, thus building a solid foundation for the improvement of students' comprehensive quality.

In addition, considering that it is impossible to completely build a mature thinking model through a course, we have left room for expansion when designing the model. The model can be carried out as an independent part in the teaching implementation of a certain course, and at the same time, can be expanded into a large system, which can be comprehensively implemented in relevant professional courses to achieve the effect of macro training.
Fig. 2. Database teaching model based on computational thinking
The model periphery is designed with three basic abilities of students, including engineering ability, that is, the ability to solve practical problems with knowledge system; knowledge ability, that is, students’ ability to quickly understand and grasp learned knowledge through learning; comprehensive ability, that is, the ability to simultaneously improve thinking in the process of knowledge learning. Of course, there is intersection between these three abilities. The three should cross each other and promote each other, demonstrating a spiral trend in the actual implementation process.
5 Summary

This paper analyzes the basic methods of thinking training in the database teaching system, and proposes combination of thinking training and knowledge learning in specific course implementation. In the course teaching, knowledge acquisition and thinking training should be organically combined to achieve the curriculum design goals. On this basis, considering students' needs of thinking training, a database teaching model is designed based on computational thinking training. With computational thinking as the core, close integration is made with database teaching content through multiple layers of thinking ability such as abstraction, automation, recursion, separation, heuristic reasoning, data protection, etc. As a result, thinking training is carried out in the process of knowledge learning, which positively combines theory and application to achieve the goal of computational thinking construction.

This paper designs database teaching system and computational thinking training model. Consisting of knowledge system and thinking training system, it comprehensively trains various thinking and multiple capabilities to promote the improvement of students' thinking ability, directs students to experience the role of thinking in the process of knowledge learning, and guides them to independently, spontaneously and conscientiously train their thinking ability.

This paper first analyzes the teaching of database courses, and puts forward learning clues from the aspects of basic database knowledge and abstract training, relational system and computational thinking, database design and computational thinking ability. On this basis, this paper describes characteristics of computational thinking in terms of decomposition and reduction, abstraction, transformation and simplification, classification and division, reverse thinking and bold assumptions. Moreover, this paper designs pre-processing method before implementing the training model, merges similar items and then designs the processing scheme.

6 Acknowledgement

This project is supported by the project of “Excellent Talent of Jinchu ” Synergistic Education Plan of Hubei Provincial Department of Education (2017 Information Security Specialty, 2018 Network Security and Law Enforcement Specialty), the Teaching Research Project of Hubei University of Police (JYXM2018A06), 2019 Scientific Research Project of Hubei University of Police.

7 References


8 Authors

Xue-peng Huang, born in Hanchuan City, Hubei Province, Lecturer of Information& Network Center, Hubei University of Police, his research interest is computer application and information security.

Jing Leng, Corresponding author, born in Wuhan City, Hubei Province, Lecturer at the Department of Information Technology, Hubei University of Police, and her research interest is computer application and information security. Address: No. 86 Jiefang Avenue, Hubei, Hubei University of Police, Wuhan, China, Postcode: 430034.