

# Supply Chain Resource Integration Method Under 4PL Mode Based On Associated Data

Huayu Chu<sup>1\*</sup>, Lichong Cui<sup>1</sup>, Yuejia Li<sup>1</sup>, Lei Su<sup>1</sup>, Yanyang Fu<sup>1</sup>, and Yuxiang Wang<sup>2</sup>

<sup>1</sup>State Grid Hebei Procurement Company, Shijiazhuang, Hebei, 050000, China

<sup>2</sup>State Grid Baoding Electric Power Supply Company, Baoding, Hebei, 071000, China

\*Corresponding author. E-mail: chuileizifu@163.com

Received: Apr 24, 2023; Accepted: Sept. 17, 2023

---

The relevance of power supply chain resources is a key factor affecting the hit rate of resource integration. Therefore, a 4PL mode based supply chain resource integration method based on associated data is proposed. Firstly, analyze the functions of the power supply chain under the 4PL mode, mainly including intelligent procurement, digital logistics, panoramic quality control, supply chain collaboration, and operational compliance. Secondly, according to the basic characteristics of the 4PL model, which is intensive, valued, standardized, and internationalized, the Apriori association rule algorithm is used for association data mining of power supply chain resources. Finally, based on the results of power supply chain association data mining, the improved ant colony segmentation algorithm is used to divide the power supply chain knowledge base into modules, and the mapping principle is used to complete the integration of supply chain resources. The experimental results show that the proposed integration method realizes effective knowledge base mapping in the process of resource integration, which can improve the utilization of power supply chain resources and reduce energy consumption, with the integration hit rate reaching 99.04%.

**Keywords:** Associated data; 4PL mode; Power supply chain; Resource integration

©The Author(s). This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are cited.

[http://dx.doi.org/10.6180/jase.202408\\_27\(8\).0003](http://dx.doi.org/10.6180/jase.202408_27(8).0003)

---

## 1. Introduction

Supply chain resource integration refers to the construction, organization, and coordination of a supply chain system to form a highly efficient operational strategic system of individuals that are related and separated from each other inside and outside the system [1]. Due to the variety and complexity of supply chain resources, how to better improve customer service level through resource integration, solve the contradiction between the overall income of the integrated system and the individual income of resources, and balance the contradiction between the integrated income and risk has become an important issue in resource integration [2]. The fourth party logistics (4PL) is a supply chain integrator that can fully integrate, deploy and manage the resources, capabilities and technologies of the

supply chain system, and provide comprehensive solutions for it. Its core advantage is that it can give play to the flexibility in the allocation of high-quality resources in the supply chain, and provide customers with personalized and diversified comprehensive solutions [3]. 4PL mode is a highly integrated and coordinated supply chain management logistics mode. How to provide reasonable and effective supply chain resource integration scheme through 4PL, fully benefit and reduce the risks of the cooperative members participating in the integration under the premise of realizing the strategic integration goal is a complex and systemic problem. At present, mature integration theory and quantitative methods have not been formed to guide the operation practice. This is a bottleneck that limits the expansion of 4PL applications [4-6].

The key to realize the integration of supply chain re-

sources under the 4PL mode is how to systematically integrate the individual resources corresponding to a certain production/service activity (a specific activity or a collection of certain activities within a period of time), and realize the optimization and value-added of the supply chain system value chain on the premise of satisfying the satisfaction level of individual value appreciation [7]. To achieve this process, it is necessary to first clarify which factors are used to grasp the integration suitability of individual resources, so as to reasonably evaluate and select high-quality resources. These factors are referred to as the dominant factors for integration.

Power supply chain is the process of converting all kinds of energy into electricity and transmitting it to the end user through the transmission network. It is an indispensable infrastructure in modern society. Power supply is the basis for maintaining the normal operation of various industries and life, and it has the characteristics of high correlation, diversity, large-scale transportation and real-time. All links in the power supply chain are closely related, including energy production, transmission, transformer, distribution and other links, and must coordinate operations to ensure reliable power supply. It involves a variety of energy forms, such as fossil fuels, nuclear, wind, solar, etc. There are differences in the production and conversion processes of these energy sources, which increases the complexity of the electricity supply chain. Electricity needs to be transmitted from the power plant to the end user via transmission lines. Transmission networks need to cover a vast geographical area, involve large-scale transmission line construction and transportation problems, and involve complex engineering and logistics arrangements. The power supply chain needs to be adjusted and managed in real time to meet the power demand of different time periods and different regions. It can be seen that the complexity of power supply chain is mainly reflected in the aspects of multi-participation, uncertainty, resource allocation and risk management. There are multiple parties involved in the electricity supply chain, including energy producers, grid operators, power supply companies and end users. The interest relationship, cooperation mechanism and policy regulations among the various participants will have an impact on the operation of the supply chain and increase the complexity of management. Electricity demand has a certain degree of uncertainty, the demand in different periods and regions is different, and is affected by weather, season, industrial structure and other factors. This uncertainty of demand poses challenges to the planning, scheduling and real-time operation of the supply chain. The power supply chain involves the production and transformation of all kinds of energy, which

requires rational resource allocation and energy selection. The comprehensive consideration of different energy supply characteristics, environmental impact, economic cost and other factors increases the complexity of supply chain decision-making. The power supply chain faces many risks, such as energy supply disruptions, transmission line failures, and poor transportation. Risks need to be identified, assessed and managed to ensure a reliable supply of electricity. Supply chain resource integration is a complex and critical link, which involves the cooperative operation, optimal allocation and efficient utilization of resources in the supply chain. However, in the traditional 4PL mode, due to the lack of effective mechanism of information exchange between resources, it often leads to the problems of resource mismatch, repeated investment and low efficiency. Therefore, an innovative approach is urgently needed to improve the integration level of supply chain resources in order to achieve higher benefits and competitive advantages.

As an effective data analysis technique, relational data mining can find the correlation and potential rules among resources. In this paper, a novel supply chain resource integration method based on associated data mining method and fourth party logistics model is proposed. By using the Apriori association rule algorithm to mine the power supply chain resources, we can find the correlation between resources, and based on the mining results, we use the improved ant colony segmentation algorithm to divide the power supply chain knowledge base into modules, so as to achieve a more efficient integration of supply chain resources.

## 2. Theoretical background

Yu et al. [8] proposes a resource integration method based on deep reinforcement learning. This method uses Zhang quantitative state representation, deterministic action output, convolutional neural networks, and weighted reward mechanisms to construct an end-to-end decision-making model for migrating strategies from data center states to virtual machines. An automated state generation mechanism and a reverse gradient constraint mechanism are designed to improve the deep deterministic strategy gradient algorithm, completing the classification and integration of data. Zhang et al. [9] proposes a resource integration method based on kernel limit learning machines. To adapt to the data flow environment, a sliding window mechanism is used to divide the data flow into data blocks, and an integrated classification model of  $k$  kernel limit learning machines is constructed on the first  $k$  data blocks. Considering the class tag correlation, the Apriori algorithm is used to obtain the association rules between the tags of each data

block, and the confidence of the co-occurrence tags in the association rules is introduced into the integration process based on the integration model. The MUENFLForest model is introduced to detect whether the new data block has concept drift. The loss function is set for the classifier to update the integrated model to adapt to the concept drift problem, and the data integration and classification processing is completed. Zhao et al. [10] proposes a resource integration method based on RELM, which uses wavelet packet transform to de-noise the data according to the noise contained in the data. By training the number of hidden layer nodes in a regularized limit learning machine (RELM), a RELM based classifier is effectively selected. Particle swarm optimization (PSO) algorithm is used to optimize the weight of the RELM-based classifier to achieve the weighted integrated classification of time series data.

Although the above methods can complete the integration of resource data, the effect of applying the above methods to the integration of power supply chain resources is not very ideal. This is because there are many types and large amounts of data in the power supply chain under the 4PL mode, which leads to some deficiencies in the accuracy and efficiency of resource integration. Therefore, a method of supply chain resource integration based on the associated data is proposed under the 4PL mode. Compared with other supply chains, the electricity supply chain is energy dependent and needs to ensure a reliable, stable and sustainable energy supply. Secondly, it also needs large-scale operation. The power supply chain involves a huge production, transmission and distribution network, including power plants, substations, transmission lines, distribution systems and other links, which makes the power supply chain large in scale and high in operation complexity. Third, the power supply chain has high timeliness and continuity requirements. Any disruption or delay in the supply of electricity can have a serious impact on social life and the economy, so the power supply chain needs to ensure the timeliness, continuity and reliability of supply to cope with emergencies and changes in demand.

### 3. Analysis of supply chain function under 4PL mode

According to the requirements of the Material Department of the State Grid Headquarters, the construction of supply chain operation centers in each province should be in accordance with the construction content uniformly designated by the headquarters, and complete the construction of the relevant functions of the supply chain operation centers in the five major sectors of intelligent procurement, digital logistics, panoramic quality control, supply chain coordina-

tion, and operation compliance, with the following details:

(1) Basic configuration function development: Complete the development of integrated configuration and maintenance functions such as the correspondence between the county company and the prefectural and municipal company, the random inspection bill of materials, the bill of materials to be subject to on-site supervision, the list of suppliers of collective enterprises, the list of key projects, the list of dishonest suppliers, the list of private enterprises, and the reasonable control range of peer benchmarking indicators.

(2) Intelligent procurement: According to the unified construction requirements of the headquarters, complete annual demand execution statistics, demand plan business overview, batch demand plan approval alert, demand plan accuracy alert, demand cycle rationality monitoring, project budget monitoring alert, power transmission and transformation project agreement inventory order execution overall and detailed query, procurement timeliness analysis, agreement inventory execution analysis, batch procurement plan analysis, the whole process monitoring of procurement project implementation progress, the whole process monitoring of online material planning, the calculation of reasonable supply cycle, the review of procurement data, the query of material reference price, the analysis of procurement situation, the query of the overall and detailed situation of centralized scale procurement, the analysis of material characteristics, the analysis of abnormal occurrence rate of key business nodes, the statistics of bid winning details, the statistics of procurement situation, development of procurement process data statistics (package division statistics, clarification and modification statistics, rejection statistics, flow bid statistics, supplier bid rejection rate, supplier bid success rate, and supplier bid package number success rate).

(3) Digital logistics sector: According to the unified construction requirements of the headquarters, complete contract signing and change statistics, agreement inventory signing statistics, contract signing indicator statistics and early warning, supply and arrival statistics, long-term non acceptance warning for goods handover, arrival acceptance warning, material supply plan completion rate indicator statistics and early warning, arrival payment monitoring, inventory turnover rate monitoring, total material resource statistics, warehousing resource statistics, voucher offset warning, early warning of overdue inventory materials, monitoring of emergency events, monitoring of implementation proportion of agreed inventory, monitoring of balance rate of agreed inventory, disposal statistics of waste and used materials, monitoring of inventory status of scrap

materials, monitoring of inventory status of scrap materials, distribution statistics of key material orders, preparation of overall production scheduling plan, early warning of operation order processing, early warning of quality guarantee document processing, quality guarantee deposit return monitoring, operation payment monitoring, advance payment monitoring, development of the function of monitoring the payment for goods and the accuracy of the agreed inventory.

(4) Panoramic quality control plate: According to the unified construction requirements of the headquarters, complete the implementation information of the provincial company's manufacturing supervision tasks, analysis of the one-time pass rate of factory tests, statistical analysis of quality problems found during manufacturing supervision, statistics and early warning of production progress overruns, application information of differentiated sampling inspection strategies, analysis of annual supplier sampling coverage, analysis of annual sampling inspection quota compliance rate, analysis of the relationship between sampling inspection items and quality issues, analysis of sampling qualification rate, and analysis of grading of sampling nonconformities, development of quality penalty payment amount information function.

(5) Operation supervision section: According to the unified construction requirements of the headquarters, the timely rate of completion of the plan declaration, the completion rate of the material procurement plan, the monitoring of the materials in the first-level procurement directory of the provincial company procurement headquarters, the monitoring of the deviation between the supplier's bid price and the reference price of the procurement application, the monitoring of less than three bidders' illegal bid opening, the monitoring of the inconsistency between the bidding results and the results recommended by the evaluation committee, and the monitoring of too frequent clarification before the bid opening, monitoring the winning situation of bidders with excessively high prices, monitoring the winning situation of bidders with excessively low prices, monitoring and analysis of purchase orders with contract amounts exceeding the bidding limit for generally mandatory bidding procurement and without going through the bidding process, monitoring and analysis of fewer than three effective bidders in the detailed evaluation stage, monitoring and analysis of the timeliness of objection handling, monitoring and analysis of the behavior of buyers evading private procurement by e-commerce, monitoring and analysis of the behavior of buyers negotiating with e-commerce to exchange goods privately, monitoring and analysis of the large deviation between the planned

delivery date and the actual delivery date, monitoring and analysis of the delivery date adjustment times greater than 2 times, monitoring and analysis of the suppliers with low performance score winning the bid many times, monitoring and analysis of the high freezing rate of bid evaluation experts, statistics of the basic information of experts, monitoring of the non-compliance of bid evaluation experts warehousing, and monitoring and analysis of the compliance of the attendance of bid evaluation experts.

(6) Supply chain collaboration: According to the unified construction requirements of the headquarters, complete contract performance business volume, drawing interaction progress statistics, material contract arrival settlement rate, supply plan completion rate, supply plan completeness rate, delivery notice completeness rate, agreement inventory arrival accuracy rate, deployment list query, key project material supply chain visualization, business chain early warning, annual demand plan and demand matching early warning, development of classified early warning function for the supply of single category key materials.

#### **4. Supply chain resource integration method under 4PL mode based on associated data**

##### **4.1. Resource association data mining of power supply chain in 4PL mode**

Looking at the development process of logistics, it can be found that there have been periodic breakthroughs in logistics theory, and the organizational structure of logistics in electric power enterprises has also undergone changes over the same period. These changes have also been affected and driven by changes in the external environment, such as the emergence and promotion of new theoretical technologies represented by communication technology and information technology, providing a technical prerequisite for breakthroughs and innovation in logistics, and promoting this process. The transformation of logistics models is basically consistent with the development of logistics theory and logistics organizational structure. This transformation is driven by the impact of the external environment, and in turn adapts to and promotes the progress of the external environment. It is an inevitable manifestation of the evolution of complex logistics systems driven by both internal and external factors [11].

With the further evolution of the reform, compared with the past, 4PL has wider coverage, more service contents, higher requirements and greater value-added. Among them, its biggest advantage is to ensure that the products can meet the needs of logistics customers "faster, better and cheaper". Its basic characteristics are mainly manifested in

the following four aspects:

(1) Intensification. 4PL integrates professional capabilities such as technology, management, consulting and logistics services, and provides a comprehensive set of supply chain logistics solutions by integrating relevant logistics resources to effectively meet the diversified and complex needs of the demanders.

(2) Value. 4PL aims to realize the planning and implementation of logistics activities in the whole supply chain, reduce the operating costs through the improvement of resource utilization, and create greater value for customers in the whole supply chain.

(3) Normalization. The emergence and development of 4PL helps to promote the unification of logistics technical indicators and quality standards, promote the standardization and standardization process of the entire logistics industry, and complete the standardization of logistics management procedures and practices.

(4) Internationalization. 4PL emerged under the general trend of economic globalization, and its own internationalization is inevitable. It is mainly manifested in the internationalization of logistics market, service demand, logistics system, supply chain management, and corporate culture.

Based on the above analysis, the operation principle of power supply chain logistics under the 4PL mode is proposed, as shown in Fig. 1. In order to integrate resources from all parties and improve industry efficiency, the difference between power supply chain logistics is that it does not require specific fourth party logistics companies to focus on the entire supply chain. Customers only need to find third-party logistics companies based on their current logistics needs, while third-party logistics companies can form resource sharing through carpooling, leasing, and other forms to promote cooperation between enterprises. Instead of using the information platform provided by third-party logistics enterprises, it provides an industry's 4PL-based information platform. Suppliers, electric power enterprises, vendors, customers and third-party logistics enterprises are all users of the platform. Its purpose is to well collect the information and resources of third-party logistics enterprises, establish the relationship between enterprises in the industry, and also allow enterprises with self-operated logistics to provide idle resources or carrier logistics, it enables enterprises to realize the integration, utilization and sharing of power supply chain resources on the basis of keeping private information of enterprises, and effectively reduce logistics costs [12–14].

After completing the analysis of the logistics operation principle of the power supply chain under the 4PL mode, the association rule algorithm is used to mine the power

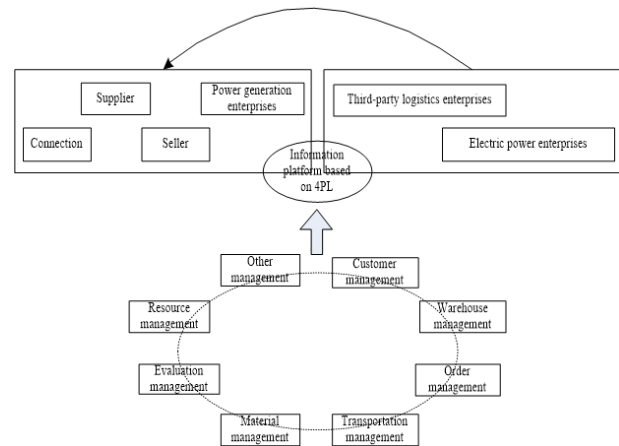


Fig. 1. Logistics operation principle of power supply chain under 4PL mode

supply chain resources.

In data mining, the association rule algorithm is one of the more common research methods. The association rule was originally designed to discover the rules of the relationship between different goods. Through the analysis of a large amount of data in the mining data, we can find customers' shopping psychological needs, scientifically place goods, reasonably select inventory and purchase, thus saving costs and improving sales. The focus of Apriori's association rule algorithm is to find simultaneous events by retrieving all data in the database, with the aim of finding reliable and effective rules [15–17]. This algorithm is called the most classical of association rules. It mainly generates new candidate item sets through continuous connection and pruning, and then calculates confidence and support to complete the mining of association data. The main association mining process is as follows:

(1) Minimum support and minimum confidence: Support refers to the probability that item sets A and B occur simultaneously in the power supply chain database:

$$\text{Support}(A \rightarrow B) = P(A \cup B) \quad (1)$$

Confidence refers to the probability of occurrence of item set B when item set A occurs in the power supply chain database:

$$\text{Confidence}(A \rightarrow B) = P(B | A) \quad (2)$$

Support is an important parameter to evaluate the validity of data mining results. When the support is high, it means that the results often appear, which is an effective result; When the degree of support is low, it means that the result is only accidental and does not have research value. From a practical point of view, it is unnecessary to pay

attention to the data mining results with low support, and electric power enterprises do not need to make changes for these rare phenomena. Therefore, the goal of minimum support is to eliminate irrelevant association rules.

Confidence is derived from association rules and has relative reliability. Assume that the data mining result is  $A \rightarrow B$  (80%), that is, the conditional probability of event  $B$  is 80%. The higher the confidence, the higher the reliability of the rule. The minimum support level is a threshold defined by the user of an association rule algorithm, which is the lowest confidence level of an association rule result. When the result of the association rule algorithm meets the threshold of minimum support and minimum confidence, the result is called strong association rule.

For the result of the occurrence of association rules, it represents the degree of support and confidence for the occurrence of both events under the rule. However, such inference does not necessarily have a causal relationship, so it is very important to analyze the occurrence result. As an analyst, it is important to have a store of knowledge about the causal and outcome attributes of data in causal analysis, and these causal analyses are likely to involve associations that emerge over time.

(2) The so-called itemset refers to the collection of several items. When there are  $K$  items in an itemset, it is called  $K$  itemset.

(3) Support count: In the power supply chain database, the number of occurrences of the itemset is called the support count or absolute support. If the support count of the itemset exceeds the originally defined threshold, then  $K$  is called the frequent itemset and recorded as  $LK$ .

When the total support count of item set, the support count of item set  $A$  and the support count of item set  $A \cup B$  are known, the support and confidence of association rule  $A \rightarrow B$  can be calculated, and the calculation formula is as follows:

$$\text{Support}(A \rightarrow B) = \frac{\text{Support\_count}(A \cup B)}{\text{Total\_count}(A)} \quad (3)$$

In the formula,  $\text{Support\_count}(A \cup B)$  represents the number of events that occur simultaneously in itemset  $A$ ,  $B$ , and  $\text{Total\_count}(A)$  represents the number of all events.

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A)} \quad (4)$$

In the formula,  $\text{Support}(A \cup B)$  represents the support count for itemset  $A \cup B$ , and  $\text{Support}(A)$  represents the support count for itemset  $A$ .

Therefore, to determine whether the association rule  $A \rightarrow B$  or  $B \rightarrow A$  is a strong association rule, you only

need to know the total support count of the itemset, the support count of the itemset  $A$ , the support count of the itemset  $B$ , and the support count of the itemset  $A \cup B$ , and then calculate whether the support and confidence of the association rule meet the threshold through Eq. (3) and Eq. (4) to determine whether the association is a strong association rule.

The main calculation idea of Apriori algorithm is to find the maximum frequent item set in the power supply chain database, compare the obtained maximum frequent item set with the preset minimum confidence and support thresholds, and generate strong association rules.

Finding out the maximum frequent itemsets in the data set is the core idea of Apriori association of association rules. The maximum frequent itemsets and minimum confidence obtained by calculation can be used to determine whether it is a strong association rule.

There are two steps to implement the Apriori association rule algorithm:

(1) Connection. The function of the connection is to generate the  $K$  item set. For the preset minimum support, by deleting the item set less than the minimum support for the candidate set  $D1$  with the number of items of 1, we can get one frequent item set  $L1$ . Then  $L1$  can generate two candidate sets  $D2$  by connecting with itself, keep the item set that meets the minimum support in  $D2$ , get two frequent item sets  $L2$ , and then connect  $L2$  with  $L1$  to generate three candidate sets  $D3$ , keep the item set that meets the minimum support in  $D3$ , get three frequent item sets  $L3$ , and so on, the frequent itemset  $LK$  will be obtained.

(2) Cut the branches. Next to the connection is pruning. The function of pruning is to eliminate the unsatisfied item set by comparing the minimum support, so as to reduce the scope of search space for the next operation. Because all non-empty subsets of a frequent itemset are also frequent itemsets, the connection of  $LK - 1$  and  $LK$  to generate  $CK$  is also a frequent itemset [18].

The process of Apriori association data mining is shown in Fig. 2.

## 4.2. Integration of power supply chain resources

### 4.2.1. Supply chain resource data segmentation

After the associated data mining of power supply chain resources is completed, the improved ant colony clustering and segmentation algorithm is used to divide the power supply chain knowledge base into modules, so as to complete the integration of power supply chain resources. Ant colony algorithm has the characteristics of self-organization, scalability, robustness and ease of use. In order to get better results, the crossover process of genetic

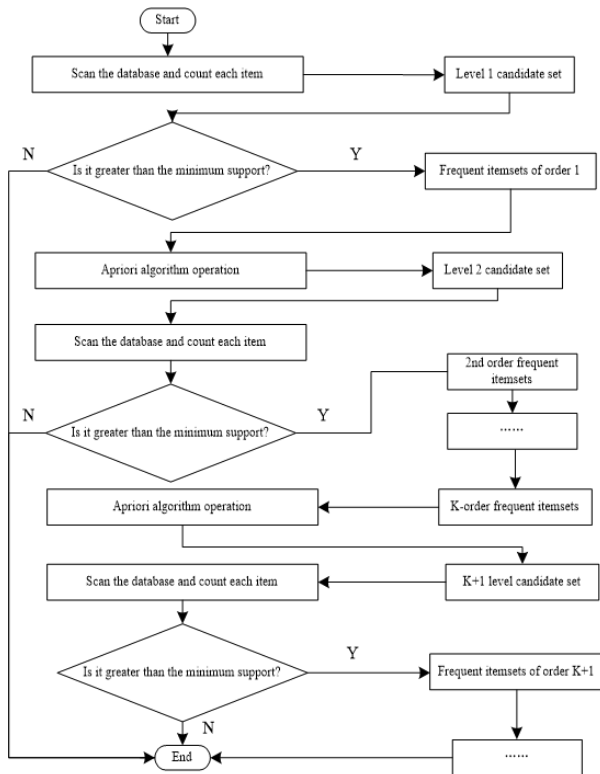


Fig. 2. Apriori association rule mining algorithm flow

algorithm is used to improve ant colony algorithm. The uniform crossover algorithm can easily generate multiple random solutions, expand the search range of ants, and thus enlarge the solution space, which can effectively avoid the premature phenomenon of ant colony algorithm. This method is applied to the division of knowledge base module of electric power supply chain, and the probability and concept similarity of each knowledge block are calculated by using similarity calculation method. By calculating the similarity between the knowledge blocks, the knowledge concept pairs about the similarity exceeding the threshold value are selected as the mapping results, and a mapping method is adopted to connect multiple knowledge bases to a network, divide the knowledge base into modules, generate knowledge blocks, and complete the integration of supply chain resources.

At present, the research on large-scale supply chain knowledge base segmentation is not mature. Most segmentation methods need to input a large number of important parameters, such as the number of blocks, which is not conducive to the automation of method implementation. How to effectively divide the supply chain knowledge base into knowledge block sets has become the main research direction of researchers. Therefore, this study proposes an improved ant colony clustering segmentation algorithm to

modularize the supply chain knowledge base. This method combines the characteristics of self-organization, scalability, robustness and ease of use of ant colony, so as to efficiently realize the division of power supply chain knowledge base modules. Ant colony algorithm is a kind of bionic optimization algorithm, which seeks the optimal solution by simulating the behavior of ants searching for food. The application of ant colony algorithm to the modular supply chain knowledge base has the following advantages: 1) distributed computing: The ant colony algorithm searches through the parallel work of multiple individuals (ants), and multiple possible solutions can be considered at the same time, improving the efficiency of the algorithm and the scope of search. 2) Self-organization: individuals (ants) of the ant colony algorithm can form an overall cooperative behavior through local information and mutual communication and cooperation. In the modular supply chain knowledge base, ant colony algorithm can automatically adjust the search strategy according to the characteristics of the knowledge base, and realize the self-organization and optimization of knowledge. 3) Adaptability: Ant colony algorithm can adapt to different environments and problems, and has strong robustness. For the modularization of supply chain knowledge base, ant colony algorithm can search based on different objectives and constraints and adapt to different needs. 4) Global search ability: The ant colony algorithm can conduct global search in the solution space through the guidance of pheromones to avoid falling into the local optimal solution. For the modular supply chain knowledge base, we can find a better way to organize knowledge and improve the overall efficiency and performance.

Italian mathematician Marco Dorigo made long-term observation, it was found that ants secrete a pheromone on the food source and the path to and from the nest, forming a pheromone track. Ants can sense the direction and intensity of pheromones when choosing the moving path, and guide themselves to choose the path. The more ants a path passes by, the higher its pheromone concentration, and the higher the probability that the ants will choose the path in the future. At the same time, the pheromone will also volatilize with the passage of time. It can be seen that the whole ant colony shows a positive feedback of information [19, 20].

Ant colony clustering algorithm based on foraging behavior regards the data object to be processed as different ants in the colony, and the cluster center can be regarded as the food source, and simulates the foraging behavior of ants by clustering process. Based on the traditional ant colony algorithm, this study made the following improvements to the ant colony clustering algorithm:

(1) The crossover process in genetic algorithm is introduced to uniformly crossover the optimal solution  $b_s$  obtained by preserving evolutionary dependencies and the better solution searched by  $m$  ants in each generation to obtain two new gene string individuals and two new clustering results. The objective function values of these two new individuals are calculated, and if they are better than those of two parents, they are retained as  $b_s$ . The uniform crossover algorithm can easily generate many random solutions, expand the search range of ants, thus expanding the solution space, and effectively avoid the premature phenomenon of ant colony algorithm.

(2) The pheromone update method of the maximum and minimum ant colony algorithm is used to update the pheromone, and the random disturbance factor is introduced to avoid the interference of the ants receiving the wrong information from concentrating on the wrong path, effectively avoiding the local optimization.

(3) The pheromone update rules are as follows:

$$\tau(i, j) = \rho\tau(i, j) + \Delta\tau(i, j) + \Delta\tau(i, j)^* \tag{5}$$

$$\begin{cases} \tau(i, j) = \begin{cases} \frac{Q}{l_m}, T_i \in C_j \\ 0, \text{ other} \end{cases} \\ \Delta\tau(i, j)^* = \begin{cases} \theta \frac{Q}{l_m}, T_i \in C_j \\ 0, \text{ other} \end{cases} \end{cases} \tag{6}$$

In the above formula,  $\Delta\tau(i, j)$  represents the increment of pheromone,  $l_m$  represents the sum of the distance between the sample and the cluster center in the optimal class result,  $\rho$  represents the degree of pheromone volatilization,  $\theta$  represents the global impact factor of the adjusted suboptimal solution,  $T_i$  represents the power supply chain resource sample set, and  $C_j$  represents the class collection.

Through the above improvement of ant colony algorithm, an improved ant colony clustering algorithm is proposed and applied to the division of power supply chain knowledge base module. The sample set of the supply chain knowledge base is defined as  $Z = \{T_i | T_i = (t_{i1}, t_{i2}, \dots, t_{id}) i = 1, 2, \dots, n\}$ , and the sample has  $z$  pattern classifications. The ant colony clustering problem is to find a partition  $C = \{C_1, C_2, \dots, C_z\}$  that satisfies  $Z = \bigcup_{i=1}^z C_i; C_i \neq \emptyset; i = 1, 2, \dots, n; C_i \cap C_j; i, j = 1, 2, \dots, z; i \neq j$ . Make the sum of intra-class dispersion minimum, that is,  $F = \min \sum_{j=1}^z \sum_{T_i \in C_j} d(T_i, m_j)$ , where  $d(T_i, m_j)$  represents the Euclidean distance from class  $C_j$  to the center  $m_j$  of the class. Therefore,  $F$  is the ant colony clustering objective function.

The algorithm is described as follows:

Step 1: Initialize. Set the values of each parameter  $\alpha, \beta, \rho, Q$ , the number of ants  $m$ , the maximum evolu-

tion algebra  $N_c$ , and initialize the pheromone  $\tau(i, j) (i = 1, 2, \dots, n, j = 1, 2, \dots, z)$ .

Step 2: Initialize  $z$  cluster centers. Randomly select  $z$  pattern sample points as the initial clustering center.

Step 3: Calculate the distance  $d(i, j)$  from the pattern sample to each cluster center and the heuristic function  $\eta(i, j)$ , where  $\eta(i, j) = 1/d(i, j)$ .

Step 4: Ants randomly select a pattern sample  $T_i$  as the starting point, calculate the probability  $P_{ij} (j = 1, 2, \dots, z)$  of each cluster center selected by the sample according to the probability formula  $P_{ij} = \frac{\tau^{\alpha}(i, j)\eta^{\beta}(i, j)}{\sum_{s \in S} \tau^{\alpha}(i, j)\eta^{\beta}(s, j)}$ , and determine the cluster  $r (1 < r < z)$  to which it belongs according to the probability  $P_{ij}$ . The pattern sample  $T_i$  is merged into  $C_r$ . Ants then select another pattern sample, repeat the above process, and merge it into the corresponding class. This cycle continues until Ant  $\alpha$  searches all the sample spaces and forms a solution.

Step 5: After forming a clustering result, the values of each component of the cluster center  $m_j$  are the average values of each attribute of each pattern sample in this category  $C_j$ . The formula is:  $m_j = \frac{1}{C_j} \sum_{T_i \in C_j} T_i$ , and the value of the objective function  $F$  is calculated.

Step 6: If  $m$  ants construct complete solutions, go to Step 7; Otherwise, go to step 4.

Step 7: Compare the objective function values obtained by  $m$  ants and save their optimal values and solutions.

Step 8: Uniformly cross the global optimal solution obtained and the current generation optimal solution obtained in Step 7. If the new solution is superior to its parent generation, save it; Repeat this step several times. After reaching the maximum number of iterations, stop and output the final result, which is the optimal solution.

The overall flow of this process is shown in Fig. 3.

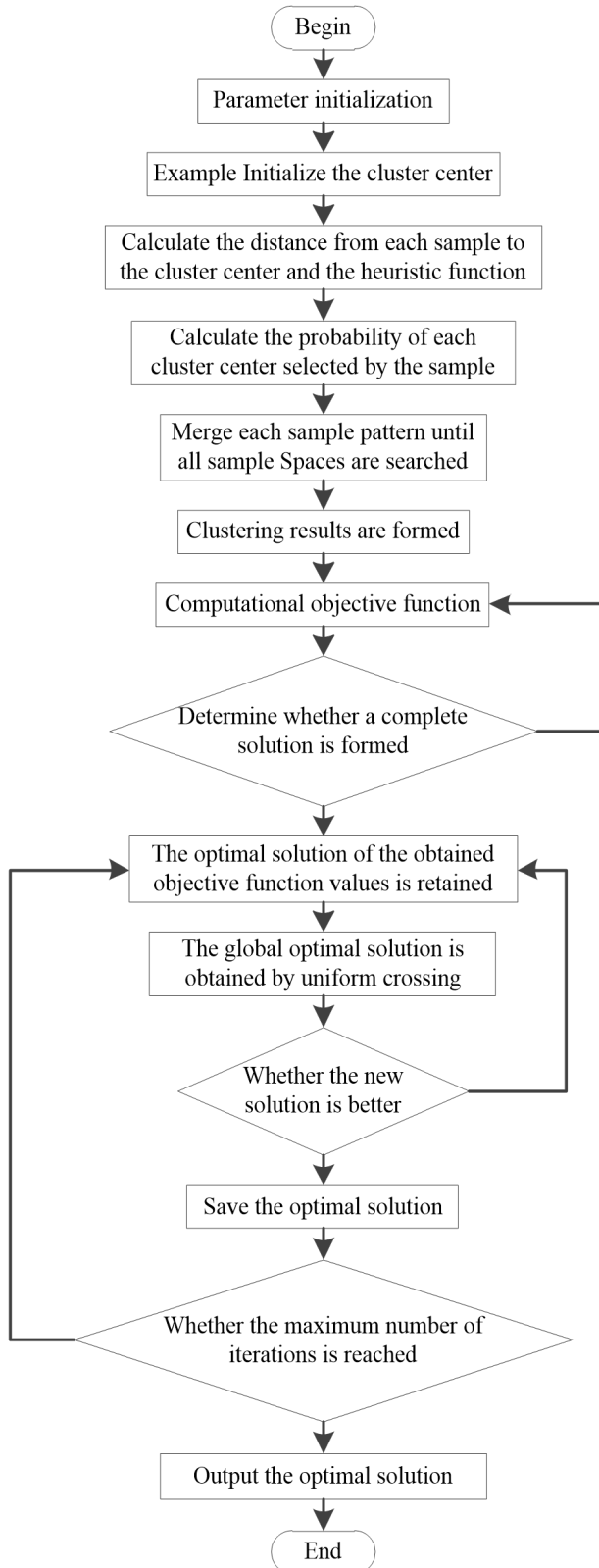
#### 4.2.2. Mapping Based Supply Chain Resource Integration

After dividing the power supply chain resources into knowledge block sets, the similarity calculation method is used to calculate the probability and concept similarity in each knowledge block, and then the knowledge concept pairs whose similarity exceeds the similarity threshold are stored in the candidate set of knowledge concept pairs.

Assume two concepts  $W_1$  and  $W_2$ , if  $W_1$  has  $n$  "semantic"  $W_{11}, W_{12}, \dots, W_{1n}$  and  $W_2$  has  $m$  "semantic"  $W_{21}, W_{22}, \dots, W_{2m}$ , the similarity between  $W_1$  and  $W_2$  is the maximum value of the similarity of each semantic, namely:

$$\text{sim } W(W_1, W_2) = \max_{\substack{i=1,2,\dots,n \\ j=1,2,\dots,m}} \text{sim } WS(S_{1i}, S_{2j}) \tag{7}$$

In the formula,  $\text{sim } W(W_1, W_2)$  represents the similarity of two conceptual languages,  $\text{sim } WS(S_{1i}, S_{2j})$  represents



**Fig. 3.** Division steps of power supply chain knowledge base module based on improved ant colony clustering algorithm

the similarity of two semantic primitives, which are atoms used to describe and define concepts. The calculation of semantic similarity can be obtained from empirical formula:

$$\text{simWP}(p_1, p_2) = \frac{\ell}{d + \ell} \quad (8)$$

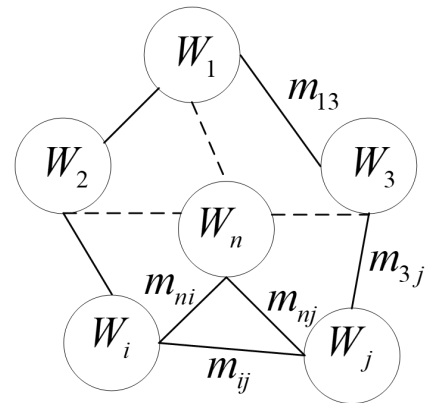
In the formula,  $d$  represents the semantic distance between  $p_1$ , and  $p_2$ , and  $\ell$  represents a variable parameter.

The scale of the knowledge base is  $n$ , and it is divided into  $n^{2/3}$  after clustering. The time complexity of generating candidate sets of knowledge concept pairs is:

$$O(n^{2/3} \times n^{2/3}) = O(n^{4/3}) \quad (9)$$

Through the calculation of similarity between knowledge blocks, about  $n^{2/3}$  pairs of knowledge concepts whose similarity exceeds the threshold are selected as the mapping result.

In order to improve the location efficiency of power supply chain resources, a mapping method is used to connect multiple knowledge bases into a network, as shown in Fig. 4.



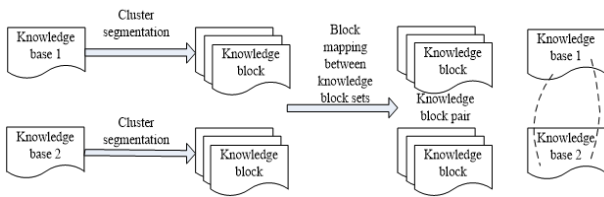
**Fig. 4.** Mapping between knowledge blocks

The nodes of the network represent knowledge blocks. The edges between nodes indicate that there is a mapping between two concepts, and the identifiers on the edges indicate that there are similar sets of all concept pairs between the two knowledge blocks.

The process of power supply chain resource integration method is shown in Fig. 5.

Assume the knowledge base  $\Sigma_1, \Sigma_2, \dots, \Sigma_n$ , and perform the following operations on them:

According to the improved ant colony clustering segmentation algorithm, the knowledge block  $W_1, W_2, \dots, W_n$  corresponding to  $\Sigma_1, \Sigma_2, \dots, \Sigma_n$  is obtained by dividing the knowledge base into modules and generating knowledge blocks.



**Fig. 5.** Power supply chain resource integration process

The similarity calculation method is used to calculate the semantic similarity between any knowledge blocks, establish the mapping relationship, and obtain the set of candidate concept pairs of knowledge blocks.

The mapping map between knowledge blocks is generated based on the candidate concepts of knowledge blocks, and the candidate concepts are used as the identification of the corresponding edges of the set.

The candidate concept pairs are stored in the coordinator by using the insertion sorting method. If all the knowledge blocks are judged, the integration algorithm is terminated; Otherwise, calculate the semantic similarity again.

## 5. Experimental verification

In order to verify the application effect of the proposed supply chain resource integration method based on correlated data in the 4PL mode, a comparative test experiment was conducted.

This experiment involves power enterprises, mainly power groups in Northeast China, which include many thermal power plants and dozens of power grid companies. They need a large amount of coal as the main raw material, and the market demand mainly changes with the season. Thermal power plants purchase coal from nationwide sources, while grid companies purchase electricity from thermal power plants and sell it to industrial enterprises and individual users.

In the experiment, 5 suppliers and 5 distributors were selected as resource nodes of the power supply chain to participate in resource integration. Assuming that the capacity data of the power plant is 500 MW, the maximum transmission capacity of the transmission line is 1000 MW, and the load of the distribution node is 300 MW, the results of mining the associated data by applying the method in this paper show that there is a positive correlation between the capacity of the power plant and the transmission capacity of the transmission line, and the correlation coefficient is 0.8, indicating that there is a strong positive correlation between them. The average prices of the two products are shown in Table 1.

**Table 1.** Average price of power supply chain products

Month	Unit cost/yuan	Unit price/yuan	Sales price/yuan
1	43.22	50.29	60.35
2	51.62	60.23	72.28
3	41.81	48.78	58.53
4	40.55	47.31	56.77
5	45.37	52.93	63.52
6	37.74	44.03	52.83
7	37.23	43.43	52.12
8	13.81	17.57	19.33
9	43.41	50.65	60.78
10	40.23	46.67	56.00
11	48.99	57.15	68.58
12	52.03	60.70	72.84

According to the collected power supply chain resource data, taking the power supply chain resource utilization rate, integration response time, energy consumption and average integration hit rate as indicators, the method in this paper is compared with the method in Yu et al. [8] and the method in Zhang et al. [9], and the research results of these two methods have good practical application effects. Comparing the method in this paper with these two methods can better demonstrate the advantages of this research method. In this experiment, the parameters of the ant colony algorithm are set as the total iteration times of 1000, the number of ants is 50, the volatilization rate of pheromone is 0.1, the initial concentration of pheromone is 1.0, the weight of Heuristic Information is 0.5, the crossing method is uniform crossing method, and the random crossing time range is as follows: 10 s to 30 s.

### 5.1. Resource utilization rate of power supply chain

The goal of carrying out power supply chain resource integration is to improve the utilization rate of power supply chain resources. Therefore, taking the utilization rate of power supply chain resources as an indicator, the performance of the proposed resource integration method is verified, and the method in this paper is compared with the two traditional methods. The results of power supply chain resource utilization ratio of the three integration methods are shown in Fig. 6.

Looking at Fig. 6, it can be seen that the integrated power supply chain resource utilization rate of this method is significantly higher than that of the methods in Yu et al. [8] and Zhang et al. [9], and the power supply chain resource utilization rate of this method does not decrease with the increase in the number of tasks, but can always maintain a relatively stable level of resource utilization.

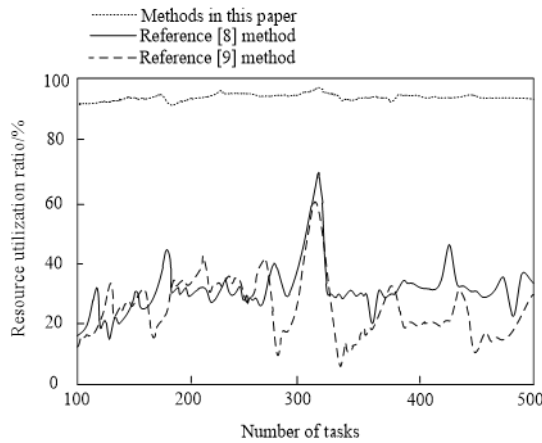


Fig. 6. Resource utilization rate of power supply chain

5.2. Integrated response time

When an integration method receives an integration request, it will respond within a certain period of time and begin the integration operation. The time from receiving the command to starting the integration operation is called the integration response time. The shorter the response time, the stronger the performance of the method. The method in this article is also compared and validated with the methods in Yu et al. [8] and Zhang et al. [9]. The integrated response time comparison results of the three methods are shown in Fig. 7.

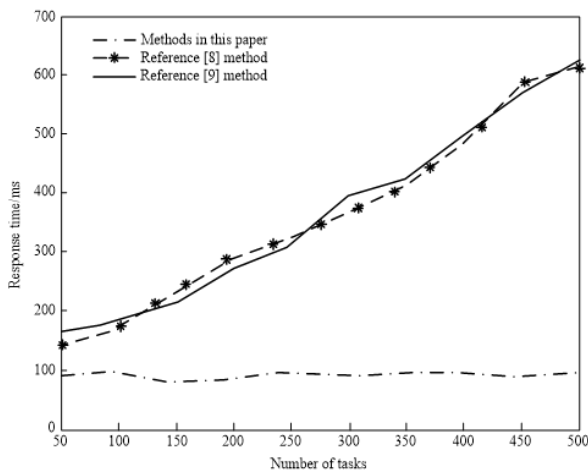


Fig. 7. Integrated response time comparison results

As can be seen from Fig. 7, with the increase of the number of tasks, the response time of the methods in Yu et al. [8] and Zhang et al. [9] has been continuously improved, and the response time of the two literature comparison methods has reached 600 ms. However, the response time of the method in this paper does not change significantly

with the increase of the number of tasks. The response time of the method in this paper is basically stable at 100 ms. Therefore, it shows that the integration response time of the method in this article does not change with the increase of the number of tasks, and has strong response performance.

5.3. Energy consumption

During the process of integrating power supply chain resources, a portion of network energy is consumed. If the energy consumption is too large, it will reduce the stability of the supply chain network operation. Therefore, taking energy consumption as an indicator, this method is compared with two traditional methods to verify. The energy consumption comparison results of the three methods are shown in Fig. 8.

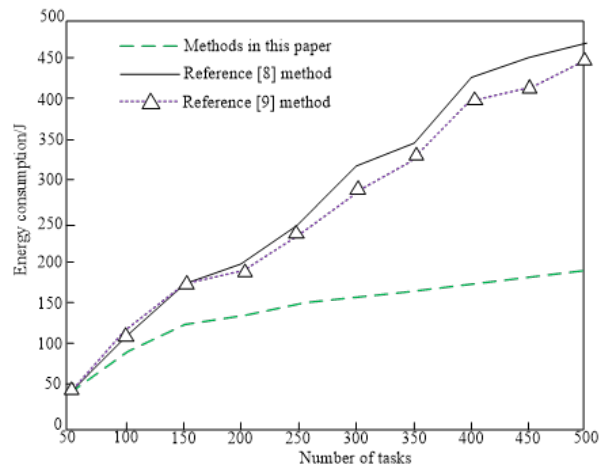


Fig. 8. Energy consumption results

The energy consumption result in Fig. 8 shows that as the number of tasks increases, the energy consumed by the three methods for resource integration shows an increasing trend, consistent with general principles. However, with the same number of tasks, the energy consumption of the methods in Yu et al. [8] and Zhang et al. [9] is higher than that of this method. For example, when the number of tasks is 500, the energy consumption of the methods in Yu et al. [8] and Zhang et al. [9] is 450 J, while the energy consumption of this method is 200 J. Therefore, this method can reduce the energy consumption of supply chain resource integration.

5.4. Average integration hit rate

The integration hit rate can reflect the resource integration performance of different methods. The higher the integration hit rate, the more accurate and effective integration of the same type of resource data can be achieved. The

**Table 2.** Average integration hit rate results

Number of tasks	Average integration hit rate/%		
	Methods in this paper	Yu et al. [8] method	Zhang et al. [9] method
100	98.87	75.64	82.67
200	98.65	79.13	81.71
300	97.92	80.55	76.98
400	99.04	78.42	73.68
500	98.34	80.31	80.75

average integration hit rate results of the three methods are shown in Table 2.

As can be seen from Table 2, the integration hit rate of the method used in this article for integrating power supply chain resources is higher than that of the two literature comparison methods. The integration hit rate of the method used in this article has always been higher than 97%, with a maximum of 99.04%, while the maximum integration hit rate of the two literature comparison methods has not reached 90%. Therefore, it shows that the supply chain resources integrated by this method are more comprehensive and accurate, and that the resource integration performance of this method is better.

## 6. Conclusion

The amount of power supply chain resources under 4PL mode is relatively large, and the types are various, which increases the difficulty of using power supply chain resources. Therefore, it is necessary to integrate power supply chain resources, and propose a method of integrating power supply chain resources under 4PL mode based on associated data. The performance of the method is verified from both theoretical and experimental aspects. When integrating power supply chain resources, the method has a high resource utilization rate and integration hit rate. Specifically, compared with the method based on deep reinforcement learning, the resource utilization rate of the text method in the supply chain is significantly improved, and the resource utilization rate of this method is always above 90%; Compared with the method based on kernel limit learning machine, the average integration hit rate of the method in this paper is significantly improved, up to 99.04%. Therefore, the proposed resource integration method based on associated data can better meet the requirements of power supply chain resource integration.

## 7. Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## References

- [1] K. Matsui, (2021) "Power structure and profitability in a three-echelon supply chain facing stochastic demand" **Computers Industrial Engineering** 160: 107246. DOI: [10.1016/j.cie.2021.107246](https://doi.org/10.1016/j.cie.2021.107246).
- [2] J. Anderson, (2021) "ISO-NE updates power generators ahead of winter amid supply chain constraints" **Platts Energy Trader** 9(4): 1–30.
- [3] M. Huang, L. Dong, H. Kuang, Z.-Z. Jiang, L. H. Lee, and X. Wang, (2021) "Supply chain network design considering customer psychological behavior-a 4PL perspective" **Computers & Industrial Engineering** 159: 107484. DOI: [10.1016/j.cie.2021.107484](https://doi.org/10.1016/j.cie.2021.107484).
- [4] H. Wang, M. Huang, W. Ip, and X. Wang, (2021) "Network design for maximizing service satisfaction of suppliers and customers under limited budget for industry innovator fourth-party logistics" **Computers & Industrial Engineering** 158: 107404. DOI: [10.1016/j.cie.2021.107404](https://doi.org/10.1016/j.cie.2021.107404).
- [5] T. Saikouk, N. Fattam, G. Angappa, and A. Hamdi, (2021) "The interplay between inter-personal and inter-organizational relationships in coordinating supply chain activities" **The International Journal of Logistics Management** 32(3): 898–917. DOI: [10.1108/IJLM-11-2020-0443](https://doi.org/10.1108/IJLM-11-2020-0443).
- [6] W. A. Jauhari, I. N. Pujawan, M. Suef, and I. D. Wangsa, (2021) "Sustainable Electrical Energy Supply Chain System With Hybrid Power Generation: An Inventory Approach" **IEEE Access** 9: 102207–102224. DOI: [10.1109/ACCESS.2021.3098543](https://doi.org/10.1109/ACCESS.2021.3098543).
- [7] C. Melo, R. Moita, and S. Sunao, (2021) "Passing through the supply chain: Implications for market power" **International Journal of Industrial Organization** 79: 102796. DOI: [10.1016/j.ijindorg.2021.102796](https://doi.org/10.1016/j.ijindorg.2021.102796).
- [8] X. Yu, Z. Li, S. Sun, G. Zhang, Z. Diao, and G. Xie, (2021) "Adaptive Virtual Machine Consolidation Method Based on Deep Reinforcement Learning" **Journal of Computer Research and Development**:

- [9] H. Zhang, P. Li, and X. Hu, (2022) "Multi-label Data Stream Ensemble Classification Approach Based on Kernel Extreme Learning Machine" **Shuju Caiji Yu Chuli/Journal of Data Acquisition and Processing** 37: 183–193. DOI: [10.16337/j.10044A9037.2022.01.016](https://doi.org/10.16337/j.10044A9037.2022.01.016).
- [10] L. Zhao, Z. Chen, L. Ding, and B. Song, (2022) "A weighted ensemble classification method for time series data based on regularized extreme learning machine" **Computer Engineering & Science** 44: 545–553.
- [11] B. Shu, F. Pei, K. Zheng, and M. Yu, (2021) "LIRP optimization of cold chain logistics in satellite warehouse mode of supermarket chains" **Journal of Intelligent & Fuzzy Systems** 41(4): 4825–4839. DOI: [10.3233/JIFS-189968](https://doi.org/10.3233/JIFS-189968).
- [12] G. C. de Bittencourt, R. D. S. Chagas, V. A. Silva, I. G. P. Vianna, R. P. Longhi, P. C. Ribas, and V. J. M. Ferreira Filho, (2021) "A solution framework for the integrated problem of cargo assignment, fleet sizing, and delivery planning in offshore logistics" **Computers & Industrial Engineering** 161: 107653. DOI: [10.1016/j.cie.2021.107653](https://doi.org/10.1016/j.cie.2021.107653).
- [13] W. Gao, (2021) "Intelligent Prediction Algorithm of Cross-Border E-Commerce Logistics Cost Based on Cloud Computing" **Scientific Programming** 2021: 1–10. DOI: [10.1155/2021/7038294](https://doi.org/10.1155/2021/7038294).
- [14] C. Han and Q. Wang, (2021) "Research on commercial logistics inventory forecasting system based on neural network" **Neural Computing and Applications** 33: 691–706. DOI: [10.1007/s00521-020-05090-4](https://doi.org/10.1007/s00521-020-05090-4).
- [15] H.-B. Wang and Y.-J. Gao, (2021) "Research on parallelization of Apriori algorithm in association rule mining" **Procedia Computer Science** 183: 641–647. DOI: [10.1016/j.procs.2021.02.109](https://doi.org/10.1016/j.procs.2021.02.109).
- [16] J. Guan, G. Shi, and H. Chen, (2022) "Research on Network Vulnerability Mining Based on Apriori Risk Data Analysis" **Computer Simulation** 39(1): 343–347.
- [17] J. Gao, Y. Zhao, L. Li, and S. Deebhijarn, (2021) "Feasibility Analysis Model of Transformation from Real Economy to Virtual Economy Based on Association Rule Algorithm" **Mobile Information Systems** 2021: 1–9. DOI: [10.1155/2021/3479532](https://doi.org/10.1155/2021/3479532).
- [18] A. Saxena and V. Rajpoot. "A comparative analysis of association rule mining algorithms". In: *IOP Conference Series: Materials Science and Engineering*. 1099. 1. IOP Publishing. 2021, 012032. DOI: [10.1088/1757-899X/1099/1/012032](https://doi.org/10.1088/1757-899X/1099/1/012032).
- [19] B. K. Dewangan, A. Jain, R. N. Shukla, and T. Choudhury, (2022) "An ensemble of bacterial foraging, genetic, ant colony and particle swarm approach EB-GAP: a load balancing approach in cloud computing" **Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)** 15(5): 693–699. DOI: [10.2174/2666255813666201218161955](https://doi.org/10.2174/2666255813666201218161955).
- [20] M. He, Z. Wei, X. Wu, and Y. Peng, (2021) "An adaptive variable neighborhood search ant colony algorithm for vehicle routing problem with soft time windows" **IEEE Access** 9: 21258–21266. DOI: [10.1109/ACCESS.2021.3056067](https://doi.org/10.1109/ACCESS.2021.3056067).