Eng Int Syst (2023) 2: 165–178

Engineering
Intelligent Systems

Optimal Management of Collaborative Goals in Power Enterprises Based on Data Mining Algorithm

Shibo Yang[1,a], Shaobo Liu[2,*], Shuiting Du[3,b], Zuohu Chen[1,c] and Zhenfen Zhang[1,d]

[1] *Gansu Tongxing Intelligent Technology Development Co., LTD., Lanzhou 730050, Gansu, China*
[2] *State Grid Gansu Electric Power Company Internet Division, Lanzhou 730050, Gansu, China*
[3] *State Grid Gansu Electric Power Company Information Communication Company, Lanzhou 730050, Gansu, China*

Data mining, otherwise referred to as discovery of information and knowledge in databases, extracts potentially useful data and expertise from a vast amount of imperfect, sporadic, noisy, fuzzy, and random material that is concealed and hitherto unknown to people. With the continuous development of big data, the power industry and enterprises generate a large amount of data in various forms such as text, video, audio, and images, that indicate enterprise development trends, enable problems to be discovered, and facilitate better decision-making. At present, the operating data of many power companies are fragmented and cannot form a unified business model. Moreover, the efficiency of data retrieval is not high, the storage environment is not unified, and the utilization rate is low. The stable operation of power enterprises has a great impact on people's daily life and social production. As reported in this paper, a study was conducted of the optimal management of power enterprises' collaborative objectives based on data mining algorithms, the current organization of power enterprises was analyzed in depth, and the overall structure of enterprise collaborative management was examined. On this basis, according to the characteristics of multi-objective optimization and large-scale problems, an improved co-evolution method was proposed and used to solve these two types of problems. Research showed that in order to adapt to current developments, the new thermal power production capacity in the February-March period from 2013 to 2019 has been increasing, but new thermal power capacity has been decreasing over the same period since 2020, reaching 5.61 GW in 2022.

Keywords: multi-objective optimization, co-evolutionary algorithms, power companies, data mining

1. INTRODUCTION

With the expansion of the power grid, the scale of the power grid also increases, and the interconnection of large power grids has become an inevitable trend of China's power development. The power grid is an important part of power transmission and of the power system. There are two networks in the power system: transmission and distribution networks. Accurate and efficient power grid planning can ensure the stability and economy of the power grid, which in turn directly affect its security and financial stability. When planning and establishing a mathematical model for large-scale power systems, it is important to give careful consideration to line expansion, network loss and overload costs during normal operation, among other factors. By transforming the distribution network, the overload problem of the power grid can be effectively solved, the voltage quality of the power grid can be improved, the loss of the power

*Address for correspondence: State Grid Gansu Electric Power Company Internet Division, Lanzhou 730050, Gansu, China, Email: huzi37812@163.com, [a]Email: z-j505@163.com, [b]18093197079@163.com, [c]czh8256359@163.com, [d]lxt6856@163.com

grid can be reduced, and the economy of the power grid can be improved. Because all of these issues are interrelated, the reorganization of a distribution network for the purpose of reducing power grid loss is a nonlinear combinatorial optimization problem, and the number of optimal parameters being changed is large. If an exhaustive method is used for the search, a "combinatorial explosion" will occur, resulting in an unreasonable amount of computation. Due to the complexity and scale of power systems, traditional mathematical statistics and heuristic algorithms have difficulty solving large-scale and complex power system planning problems. Therefore, it is very important to study the optimal management of collaborative objectives in power enterprises.

The collaborative objective optimization problem presents in many fields. Chen et al. (2018) conducted in-depth research on oil drilling and the transportation system, and summarized the design and scheduling problems as a kind of multi-vehicle routing problem with a stringent time window. Under the condition that each FPSO can provide services within a certain time window, aimed at the minimumminimizing total transportation cost, an optimal fleet design scheme and ship type scheme (ship type and ship type number) were established, and each route was optimized. The Liezi-generation method was applied to obtain accurate calculations, and a simulation test was carried out. The research results showed that this method can better solve the collaborative optimization problem associated with tanker design and scheduling [1]. In order to define the task-coupling model in the development process, WANG et al. (2019) established a weighted digraph based on information correlation, and demonstrated the corresponding relationship between the weighted digraph model and the numerical design structure matrix model of coupling tasks. The task-coupling model was quantified and the interaction matrix of the task package was established. A multi-objective task-decoupling method based on improved genetic algorithm was proposed to decouple the task-coupling model, and the decoupling of the task package was transformed into a multi-objective optimization problem. Then the improved genetic algorithm was used to solve the interaction matrix of coupling tasks. Finally, taking the task package decoupling of the joint development of radar-phased array antenna as an example, the effectiveness of this decomposition method was proved [2]. Pearce et al. (2018) presented an optimizing frame to generate task allocations and timetables for teams of human-robots with the goal of improving time and ergonomics, and demonstrated its use in six real-world fabrication projects now being undertaken by humans. Using a stimulus index method to indicate human physical stress levels, a solution was created that ranked the order of priorities for every target. The resulting timeline provided engineers with insights enabling them to choose the appropriate level of compromise and integrate the robot in a way that best suited the requirements of individual processes [3]. Li and Wang (2017) proposed a new multi-closed-loop space deployable mechanism. Based on the cyclic constraint equation, an analysis was carried out of the deployable mechanism's position. Then, under the premise of a given range of aperture and opening angle specifications, a mathematical model for the optimization of the size of the deployable mechanism was established, where the design variable is the length of the rod. Through the collaborative optimization method, the optimization problem was divided into two levels: one system level is used to coordinate the inconsistency of coupling variables in each subsystem, and three parallel subsystems are used to minimize the difference between the optimization scheme of each subsystem and the system level objective. The size parameters of the deployable mechanism can be optimized by the established co-optimization model. The simulation results showed that when the expansion ratio, the driven stroke and the minimum transmission angle were 15:1, 198mm and 22°, respectively, the unfolding process was smooth and there was no dead point [4]. But the scope of application was rather limited.

The reduction of costs associated with packaging is critical for all types of companies and institutions. Usually, different packages are designed for individual products, which leads to less cost-effective pricing of packaging systems. Zhao et al. (2017) developed a data-mining model with triple agglomerative clustering algorithms to modularize package systems by reducing the diversity of package sizes. The three algorithms are k-means clustering, aggregated hierarchical clustering, and self-organizing feature maps. Package models with similar shapes and sizes were automatically clustered and replaced by a package model with a size to fit all. The results showed that the packaging system using the clustering hierarchical clustering algorithm is more cost-effective in this case than the packaging system using the other two clustering algorithms and the packaging system without modularization [5]. The main purpose of data mining is to efficiently deal with a wide range of data, to extract actionable series, and to obtain valuable business and other insights. Data mining is the only component of Knowledge Discovery (KDD) activities within databases. Success and better decision-making often depends on the speed with which data are identified and analyzed. The information derived from data mining can be used to determine the most effective actions to be taken in future, to streamline operations, and improve a company's future prospects. Anandakumar and Arulmurugan (2019) proposed certain algorithms suitable for large datasets. These algorithms explained some of the structures and techniques that have been applied to address big data issues. In addition, the common strengths and limitations of these algorithms were evaluated. Thus, data mining researchers can either be instructors or trailblazers by selecting or developing algorithms that can be used to deal with identified challenges [6]. Subeesh et al. (2018) compared and evaluated the properties of data mining algorithms for the testing of signatures. The three most frequently used data mining algorithms (DMAs) (reporting odds ratio (ROR), reporting ratio (PRR), and information component (IC)) were chosen and applied retrospectively in the FDA Adverse Event Reporting System in a database to examine a portfolio of five identified drug events. Various data mining algorithms were compared for their sensitivity and early-stage detection capabilities. Of these three data mining algorithms, information composition had the highest sensitivity (100%), followed by reporting odds ratio (60%) and proportional reporting rate (40%) [7]. These methods offered some references for our investigation, but this study has not yet achieved recognition due to the short length of the study and the small-scale sample size.

This current study was conducted to analyze data in terms of human resource management optimization and capacity supply in power enterprises. Imperfect incentive and assessment systems were responsible for 48% of the human resource management problems in electric power enterprises. Innovative countermeasures to improve the performance and compensation management mechanism is one of the most concerning issues as these mechanisms cause 62% of problems. From February to March 2022, the average utilization hours of thermal power equipment was 821 hours, and the new thermal power production capacity reached 5.61 million kilowatts in 2022. New energy hydropower, wind power and solar power produced 18, 26 and 44 ten thousand kilowatts more than for the same period during the previous year.

2. OPTIMAL MANAGEMENT METHOD OF POWER ENTERPRISE COLLABORATIVE TARGET BASED ON DATA MINING ALGORITHM

At present, the power industry is facing an increasing number of challenges and pressures on a global scale. The gap between the growth in demand and the supply of energy is widening, the pressure on environmental protection and land resources is increasing, customers have greater requirements for power supply reliability and service levels, and new technologies such as sensors, communication devices, and IT are constantly being introduced [8]. These new trends have prompted the power industry to enter an important transition period, that is, from the simple focus on safe and reliable operation mode in the past, to a new sustainable development model that pays more attention to customer satisfaction, comprehensive benefits, energy conservation and environmental protection. In order to cope with these pressures and challenges, developed countries in Europe and the United States began their research and development of smart grids in 2005 and 2004, respectively [9]. Developed countries in Europe and the United States established the "Smart Grid International Alliance" organization to promote the development of the power industry, and conduct research on new technologies, methodologies, operating models, and evaluation models involved in smart grids, and have carried out a large number of practical applications [10]. Worldwide, the smart grid has become a new trend in the generation of electric power [11].

In order to support the construction of smart grid, in terms of informatization, and based on the main network, it is necessary to connect the equipment, systems, customers, employees, etc. [12]. "On-demand" access, use, and analysis make the operation and management of the entire power grid enterprise more efficient, automated and optimized. However, the construction of a smart grid faces three major problems: large amounts of data, integration and coordination, and continuous change [13].

(1) The amount of data is large. With the continuous development of smart grid, more and more sensors are deployed in the grid to monitor its operations; the large-scale application of sensors necessitates new requirements for enterprise informatization work, namely: how to construct a complete information system structure and effectively manage the huge data generated therefrom; how to synthesize all aspects of data to establish a complete and comprehensive view that reflects the operating conditions of the power grid; how to convert a large amount of data into useful information, which can be used to improve and improve the efficiency of enterprises; and how to provide users with a unified data expression and unified information publishing platform [14–15].

(2) Integration/coordination. Smart grid is related to the operation of the entire enterprise, and requires information exchange and process integration between multiple business departments and multiple application systems. Hence, the IT system of power companies needs to be able to effectively support the exchange and integration of data across business and application systems [16]. In this regard, the smart grid has higher requirements compared to those of traditional systems due to several factors: the scope of information exchange and integration is broader, and the interaction between a few independent application systems has risen to the enterprise level; the content of information interaction and related data are more abundant, including real-time data and management information; the exchange and integration of information has become more complex [17–18]. According to the general idea underpinning the construction of China's State Grid Corporation informatization, in terms of information integration and collaboration, business collaboration and integration based on human and financial management, asset life-cycle management, customer all-round management, and energy full-process management have been established [19].

(3) Continuous change. The smart grid itself is always evolving, which means that power companies are constantly undergoing changes in terms of organization, business model, business process, and technical means [20]. This requires IT systems to be able to adapt and support future changes, thereby supporting the sustainable development of smart grids.

Power companies tend to have several power engineering and infrastructure projects being undertaken simultaneously, such as substation construction, transmission infrastructure construction, etc. In order for the company to improve its operating strength, it must continuously invest in various power projects, and carry out ongoing maintenance [21]. Excellent project management can effectively improve the work efficiency and quality of the project team, reduce operating expenses such as maintenance costs and equipment renewal costs, reduce asset depreciation and loss, and improve corporate competitiveness in the increasingly competitive power industry [22]. The schematic diagram of the project management system is shown in Figure 1.

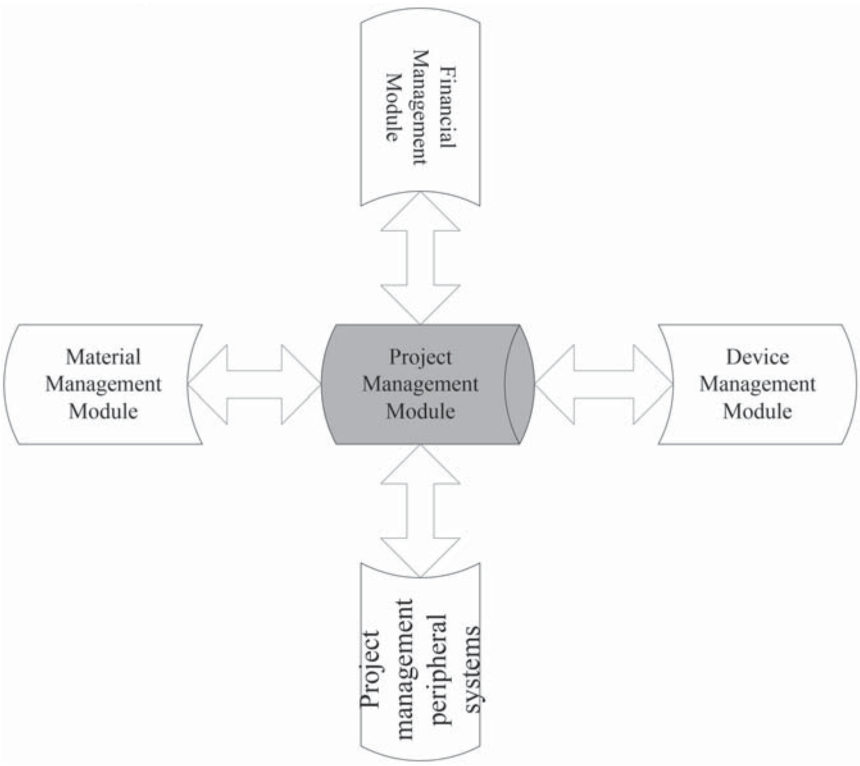

Figure 1 Schematic diagram of the project management system.

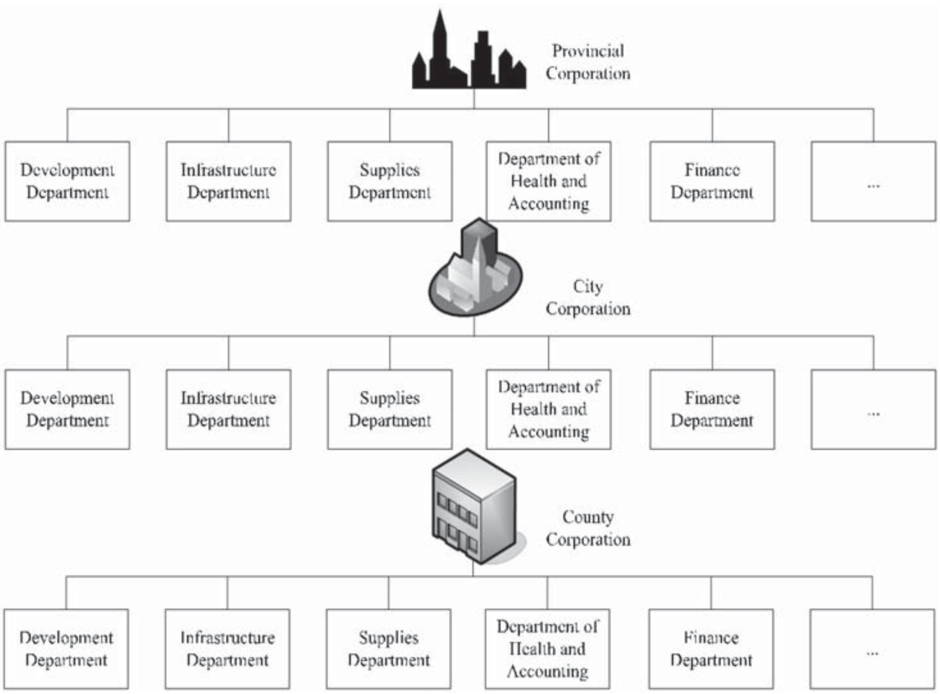

Figure 2 Schematic diagram of the current state of the power organization.

Electricity has established a flat, market-oriented organizational framework. In the framework, the management of a construction project can include the following departments: including: production technology, project management (development planning, infrastructure, marketing, etc.), finance, logistics procurement, and other departments, all of which take some responsibility for the business flow of the project [23]. Its organizational structure is shown in Figure 2.

According to the theory of synergy, the collaborative management of logistics node enterprises in the power generation supply chain means that in order to achieve the common goal of the equipment logistics activities of power generators, collaboration and cooperation among the node enterprises in the supply chain are interdependent and indispensable [24–25]. Based on this, the overall structure of collaborative management of logistics node enterprises in the

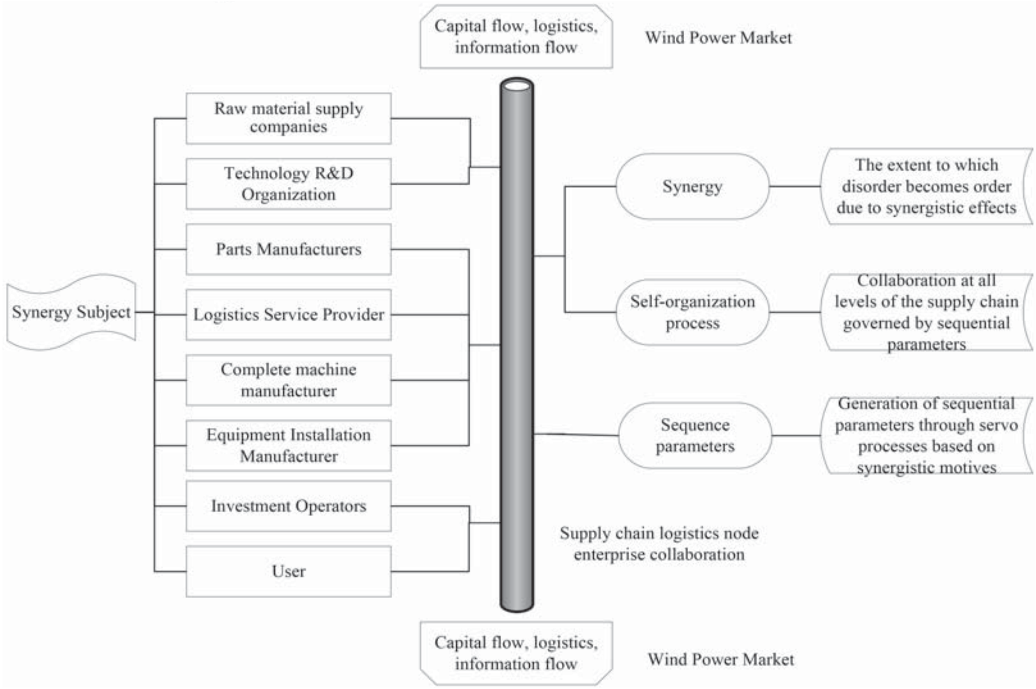

Figure 3 The overall structure of collaborative management of logistics node enterprises in the power generation supply chain.

power generation supply chain can be constructed, as shown in Figure 3.

The overall structure of the collaborative management of supply chain logistics nodes is depicted in Figure 3, and can be described in terms of three levels: micro, medium and macro. At the micro level, the collaborative management structure focuses on the relationship between supply companies and users, and its purpose is to ensure a good connection between upstream and downstream node companies in the supply chain when they carry out logistics activities [26]. At the medium level, the synergy of supply chain logistics can provide a reliable guarantee of the smooth flow of logistics, information flow and capital flow in the power generation supply chain. Essentially, the key goal of meeting users' needs guides the collaboration. At the macro level, the synergy of the power generation supply chain can organically unify all the node enterprises in the supply chain, and achieve the strategic goal of maximizing the overall benefit of the power generation supply chain through mutual cooperation [27]. Taking the wind power supply chain as an example, the main entities comprising this supply chain logistics node enterprise collaboration are the users and the companies specializing in raw material supply, parts supply, logistics service, complete machine manufacturing, wind power equipment installation, and wind power investment and operation companies. Various types of collaborative entities also contain a lot of specific content, as shown in Table 1.

Safe, reliable, economical and stable power supply quality of power grid is the core objective of power grid dispatching and management. An uninterrupted, high-quality power supply is crucial to today's social and economic development. A power enterprise is a whole; it is not a simple business combination: it has common interests and strategic purposes. From the perspective of the main body of collaborative implementation, in order to achieve the overall operating benefits and strategic goals of a power enterprise, there must be cooperation between the company's headquarters and its various departments, as well as between the departments.

A large-scale optimization problem is characterized by multi-dimensional and complex relationships between components. Therefore, it is of great theoretical and practical significance to study high-efficiency algorithms with simple computation, low storage requirements, and high efficiency in solving large-scale problems. Large-scale nonlinear optimization problem is:

$$\text{Minimize: } f(x) S.t. x \in \Omega \subseteq R^n \tag{1}$$

Among them, f is a real-valued function with definition R^n above, and the function F is called the objective function of the problem, and the set Ω is the feasible domain of the problem, and the points in the feasible domain are feasible points.

As the size of the problem increases, the solution space and time requirements increase simultaneously, and the nature of the problem also changes as the size increases. But not all problems will increase in computational complexity as the dimension increases. In some high-dimensional problems, the variables are independent and unrelated to each other, and they can be solved by decomposing them into multiple sub-parts with general algorithms. Problems of this type are called decomposable problems. The decomposable large-scale problem is defined as:

$$x_1 \cup x_2 \cup \ldots \cup x_p = x \tag{2}$$

$$\arg\min_x f(x) = (\arg\ \min_{x_1} f(x_1), \ldots, (\arg\ \min_{x_m} f(x_m)) \tag{3}$$

A multi-objective optimization problem can generally be expressed using n decision variables, P objective functions and K constraints. The optimization objective is expressed as:

Table 1 The main body of the wind power supply chain logistics node enterprise collaboration.

Collaborative Subjects	**Specific content**	**Collaboration Subjects**	**Specific content**
Raw material suppliers	Light wood supply companies, bearing supply companies, slip ring supply companies, lightning protection equipment supply companies, sensor supply companies, etc.	Logistics service enterprises	Supply logistics service enterprises, production logistics service enterprises, sales logistics service enterprises, third-party logistics service logistics service enterprises, etc.
Parts supply enterprises	Tower supply enterprises, blade supply enterprises, pitch yaw system supply enterprises, gearbox supply enterprises, generator supply enterprises, hub supply enterprises, etc.	Complete machine manufacturing enterprises	Wind turbine manufacturing enterprises
Equipment installation enterprises	Wind power equipment installation enterprises	Wind power investment and operation enterprises	Wind power investment enterprises, wind power operation enterprises
Users	Power grid enterprises, end users	–	–

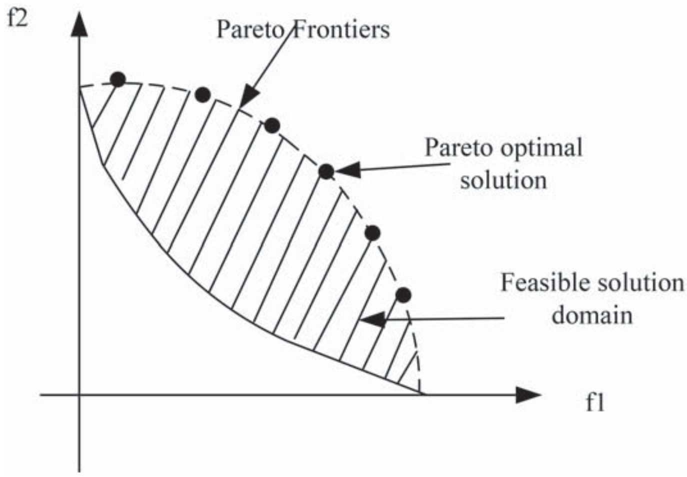

Figure 4 Pareto optimal solution and Pareto frontier.

$$\max\ y = f(X) = \left[f_1(X), f_2(X), \ldots, f_P(X)\right] \quad (4)$$

$$g_i(X) \le 0(i-1, 2, \ldots, m) \quad (5)$$

$$h_i(X) = 0(i-1, 2, \ldots, q) \quad (6)$$

In Formula (4), Formula (5) and Formula (6), $X = (x_1, x_2, \ldots, x_n) \in D$ is the decision vector; $y = (f_1, f_2, \ldots, f_n) \in Y$ is the target vector; D is the decision space formed by the decision vectors; Y is the objective space formed by the objective vectors.

In the case of conflicting objective functions, there is no optimal solution method that can simultaneously achieve the optimization of all objective functions. In this case, the concept of a non-inferior solution or Pareto optimal solution can be used, which is defined as follows.

A solution X^* is called a non-inferior solution or Pareto optimal solution if and only if there is no feasible solution for x

$$f_i(X) \ge f_i(X^*), i = 1, 2, \ldots, m \quad (7)$$

and the strict inequality sign holds for at least one sequence number i. Obviously, in the continuous case, the set of all non-inferior solutions is actually a curved (line) surface, often called the Pareto front (Figure 4).

The dotted line represents the front end of Pareto, with D as the reference point, and the two points B and C are located in the lower left corner of D, then there are B $\prec$ D and C $\prec$ D. In the same way, there are B $\prec$ C, but D and E are not comparable, indicating that D is the dominant region. The optimal boundaries of the two optimization objective functions constitute a curve, and the optimal boundaries of the three optimization objectives constitute a curved surface. The optimal boundaries of the four optimization objective functions constitute a hypersurface.

In recent years, rough set theory has been applied successfully in artificial intelligence, data mining and other fields, and has attracted increasing attention from researchers. Rough set theory has the following characteristics: different from the methods of statistics and fuzzy set theory used to deal with inaccurate data, it is through the observation and measurement of incomplete data, through the analysis and reasoning of incomplete data, that inherent relationships and characteristics are discovered, and useful patterns can be extracted. Rough set theory is based on a definite mathematical expression and is completely dependent on data. The analysis process is objective, can effectively eliminate redundant information, simplify the calculation process, and give a minimal and easy-to-understand knowledge representation, which overcomes the disadvantages of traditional uncertainty information processing methods that need to rely on additional information or prior knowledge when processing data, making it difficult to process massive amounts of information.

For algorithmic considerations, object information was given in the form of a data table, where each row corresponds to a different object, and each column corresponds to a different attribute under consideration. Each cell in the table represented an evaluation of the object placed in the row using the attributes of the corresponding column.

A data table is a quad

$$S = \langle T, Q, V, f \rangle \tag{8}$$

where T is a finite set of objects,

$$Q = \left\{q_1, q_2, \ldots, q_p\right\} \tag{9}$$

is a finite set of attributes, V_q is the domain of attribute q, $V = \cup_{q\in Q} V_q$, and $f \quad T \times Q \rightarrow V$ is a complete function that satisfies $f(x, q) \in V_q$ for each of $q \in Q$ and $x \in T$, called the information function.

Each object x in T is described by a vector

$$Des_Q(x) = \left[f(x, q_1), f(x, q_2), \ldots, f(x, q_p)\right] \tag{10}$$

which is called the description of x given by the evaluation of the properties in Q, which represents the available information about x.

Each attribute of subset W is associated with an indistinguishable relation on U, denoted as I_h, which is defined as:

$$I_h = \{(x, y) \in T \times T : f(x, q) = f(y, q), \forall q \in H\} \tag{11}$$

The set of those relations I_h is written as $T|I_h$, and the equality class that includes an element $x \in T$ is recorded as $I_h(x)$. The efficiency class of the relative I_h is called the H-basic set or the particles of knowledge encoded by H.

Let W be a data table, X be a non-empty subset of T, and $\emptyset \neq H \subseteq Q$. A set X can be described by two ordinary sets, called the H-down approximation and H-up approximation of X in W, denoted by $\underline{H}(X)$ and $\overline{H}(X)$, respectively, and defined as:

$$\underline{H}(X) = \{x \in T \quad I_h(x) \subseteq X\} \tag{12}$$

$$\overline{H}(X) = \cup_{x\in X} I_h(x) \tag{13}$$

The sets $X \subseteq T$ with the same H-lower approximation and H-upper approximation are called H-rough sets. The elements in $\underline{H}(X)$ are all and only objects $x \in T$; $\overline{H}(X)$ that belong to the equivalence class produced by the indistinguishable relation I_h and contained in X; elements in $\overline{H}(X)$ are all and only objects $x \in T$ that belong to the equivalence class produced by the indistinguishable relation I_h and that contains at least one element x belonging to X. In other words, $\underline{H}(X)$ is the largest combination of H-basic sets containing X, and $\overline{H}(X)$ is the smallest union of H-basic sets containing X.

The H-boundary of X in W is denoted as $Cn_h(X)$ and is defined as:

$$Cn_h(X) = \overline{H}(X) - \underline{H}(X) \tag{14}$$

Rough approximation is a general term used to denote H-lower approximation and H-upper approximation operations on one set or a combination of sets. The rough approximation obeys the following basic laws.

Inclusion:

$$\underline{H}(X) \subseteq X \subseteq \overline{H}(X) \tag{15}$$

Complementarity:

$$\underline{H}(X) \subseteq X - \overline{H}(T - X) \tag{16}$$

Thus, if some object x belongs to $\underline{H}(X)$, it was definitely contained in X, and if x belongs to $\underline{H}(X)$, it was only possibly contained in X. $Cn_h(X)$ gave a suspicious region of X: with the knowledge contained in H, it was impossible to determine whether an element in it was contained in the set X.

In order to be a feasible solution for multi-objective optimization problems, a multi-level neighborhood topology structure suitable for multi-particle swarm PSO algorithm was improved, and the concepts of cooperative particles and virtual centers were introduced, as shown in Figure 5.

In the neighborhood topology shown in Figure 5, the particles (indicated by the black solid origin in Figure 5) in a single population (indicated by the dotted circle in Figure 5) adopted a wheel-type neighborhood topology; between multiple populations, the fully connected star-shaped neighborhood topology is used for the exchange of information between the populations. Unlike the definition of particle individuals in the basic PSO algorithm, the particles at the individual level defined in this paper are called 'cooperative particles', and the common reference individuals in the wheel topology are called 'virtual centers'.

The calculation and evaluation of the fitness of the virtual center guide the corresponding cooperative particles to update the speed and position information. In this way, the problem-solving mechanism of the multi-particle swarm optimization algorithm is formed by a process of repeated iteration and information exchange. In this paper, the arithmetic mean of the coordinate particle position vectors is taken as the position vector of the virtual center, and the calculation expression is:

$$x_i^{'s} = \frac{\sum_{j=1}^{m} x_i^s}{m} \tag{17}$$

In Formula (17), x_i^s represents the position information of the j-th cooperative particle in the s-th iteration; $x_i^{'s}$ represents the position information of the i-th virtual center comprising the above-mentioned cooperative particles; m is the number of cooperative particles.

In the fully-connected neighborhood topology at the population level, the fitness function value of each virtual center was normalized and used as the weight indicator of

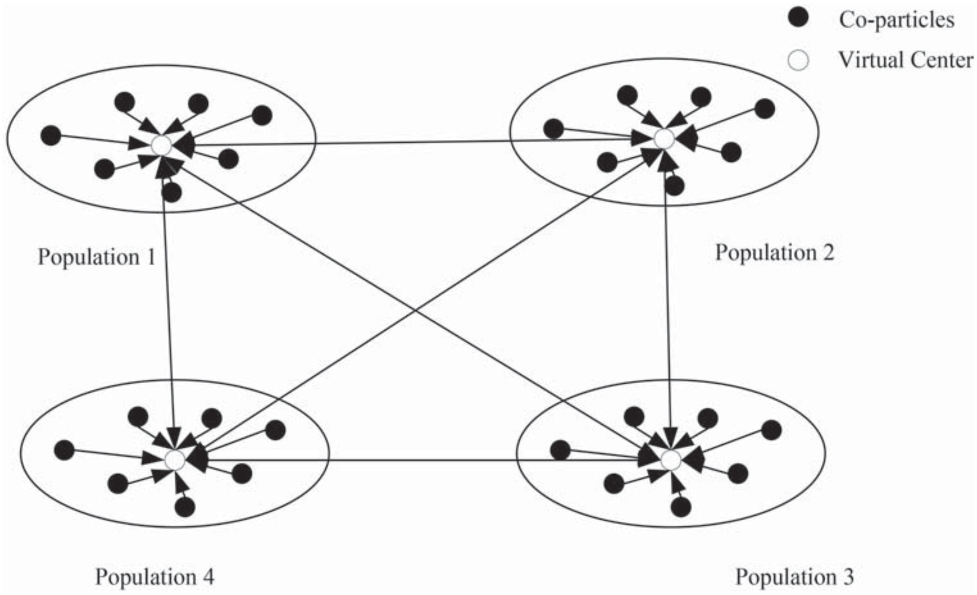

Figure 5 Multi-particle swarm PSO coordinated co-evolutionary algorithm domain topology.

the information interaction between populations to realize the information interaction between the population levels. Assuming that the minimum fitness function value is optimal, the weight index of the i-th virtual center affected by the j-th virtual center is:

$$w_{ij} = \frac{f(x_i^{'s})/f(x_j^{'s})}{\sum_{l=1}^{n} \frac{f(x_i^{'s})}{f(x_l^{'s})} - 1}, i \neq j \tag{18}$$

Then the update calculation of the c-th collaborative particle velocity and position corresponding to the i-th virtual center can be modified as:

$$\begin{aligned} v_{ic}^{k+1} &= \omega \cdot v_{ic}^{k} + r_1 \cdot z_1 \cdot (p_{ic}^{k} - x_{ic}^{k}) \\ &\quad + \sum_{j=1}^{n} w_{ij} \cdot (r_2 \cdot z_2(x_j^{'s} - x_i^{'s})), (i \neq j) \end{aligned} \tag{19}$$

$$x_{ic}^{k+1} = x_{ic}^{k} + v_{ic}^{k+1} \tag{20}$$

where p_{ic}^{k} represents the optimal position of the c-th cooperating particle individual corresponding to the i-th virtual center in the k-th iteration process.

3. EXPERIMENT DESIGN OF COLLABORATIVE TARGET OPTIMIZATION MANAGEMENT OF POWER ENTERPRISES

Human resource management is an important issue in current enterprise management. The competition between enterprises is actually a competition between talents. The effective use of data can effectively improve the speed of enterprise operations. By managing the unstructured data in the database, the level of enterprise information management can be improved. Statistics show that at present, the amount of data generated by most enterprises in the form of files, pictures, videos, emails, and other data, is doubling every year, and 80% is unstructured data stored in various business systems and computers. The general structured data only accounts for 20% of the total data volume. These invisible resources require enterprises to actively manage and mine them for valuable information for potential use. How all the unstructured information is being managed is an important indicator of the level of enterprise informatization and the extent to which stored data is being utilized. The search engines used by companies must have the capability to quickly and safely find all information related to enterprise operations and activities, and should enable users to access the data according to their needs, including all kinds of unstructured data. Companies' search engines can carry out targeted quantitative collection, with greater efficiency and wider retrieval range. At the same time, personalized retrieval and push can be carried out according to the consumer's interests and hobbies, so as to improve the accuracy of retrieved data. After classifying, clustering, and semantic analysis of unstructured data, it builds functions such as huge unstructured data organization, analytical processing, and secure storage.

At present, many companies use keyword search, cross-filtering, language methods, labeling methods, etc. to improve and enhance the access and processing capabilities of unstructured data, which basically require external manual intervention. At present, the more mature method is to use content understanding and concept matching technology to improve the automatic analysis and processing of unstructured data. This enables internal and external data sources between various heterogeneous systems to be integrated. Enterprise search engines can usually analyse and process both structured and unstructured information by means of the functions shown in Figure 6.

In the power companies and the power industry in general, data is of great value as an invisible asset. Unstructured data

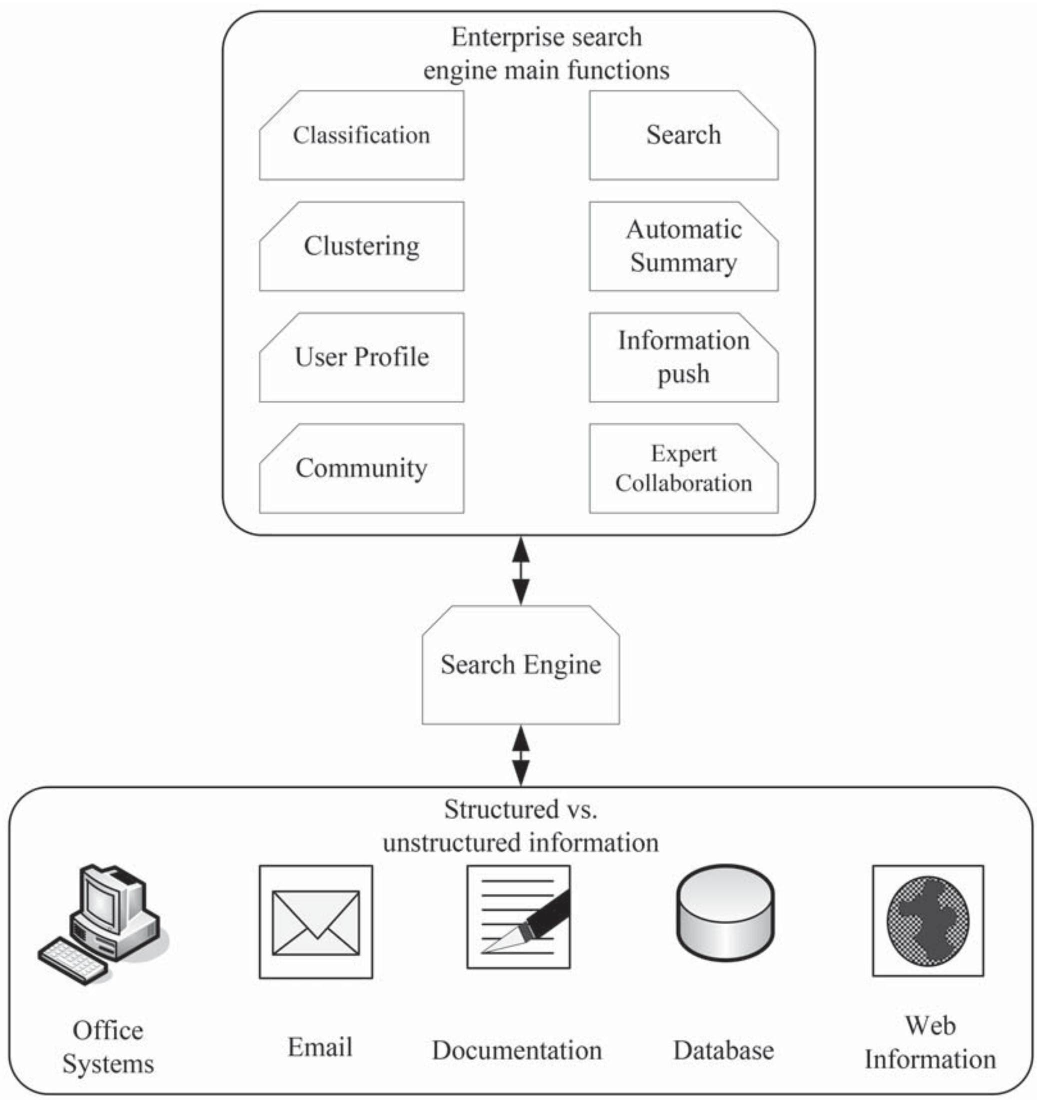

Figure 6 Enterprise search engine features.

Table 2 Experimental results of XML document storage.

XML documents	Document size	Loading time
Doc1.XML	512K	43s
Doc2.XML	1.5M	80s
Doc3.XML	10.4M	896s

generally consists of text files, audio and video files, XML files, pictures, and web pages. It is a data type that cannot be represented by a two-dimensional table structure. Table 2 shows the experimental results of XML document storage. A company's management of unstructured data involves data collection, comprehensive management, secure storage, and content publishing. Efficient data management can strengthen a company's competitiveness in the market.

The power industry, and power companies and their departments have many different business application systems, and each system has its own storage database, which is distributed on different servers. The platform needs to manage the unstructured data within the enterprise as a whole, optimize the data storage strategy, integrate storage resources, and improve the utilization rate; and optimize the system configuration to ensure the safe storage of data. Due to the limited experimental conditions, the experimental platform cluster in this paper consisted of 3 computers, and 4 nodes were built. The configuration is shown in Table 3.

4. POWER ENTERPRISE MANAGEMENT DATA

In order to improve the core competitiveness of power enterprises and better adapt to the development trend of the power industry in the future, it is necessary to establish a sound personnel management mechanism. This can be done by fully exploring the potential of human resources, improving

Table 3 Experimental environment configuration.

Node Type	CPU	RAM	Hard Disk
NameNode, DataNode	4 Core Intel Core 2 Q8200 at 2.33GHz	4G	500G
DataNode	Dual-core Intel Celeron E3300 at 2.00GHz	3G	250G
DataNode	Dual-core Intel Core 2 T7250 at 2.50GHz	1G	120G

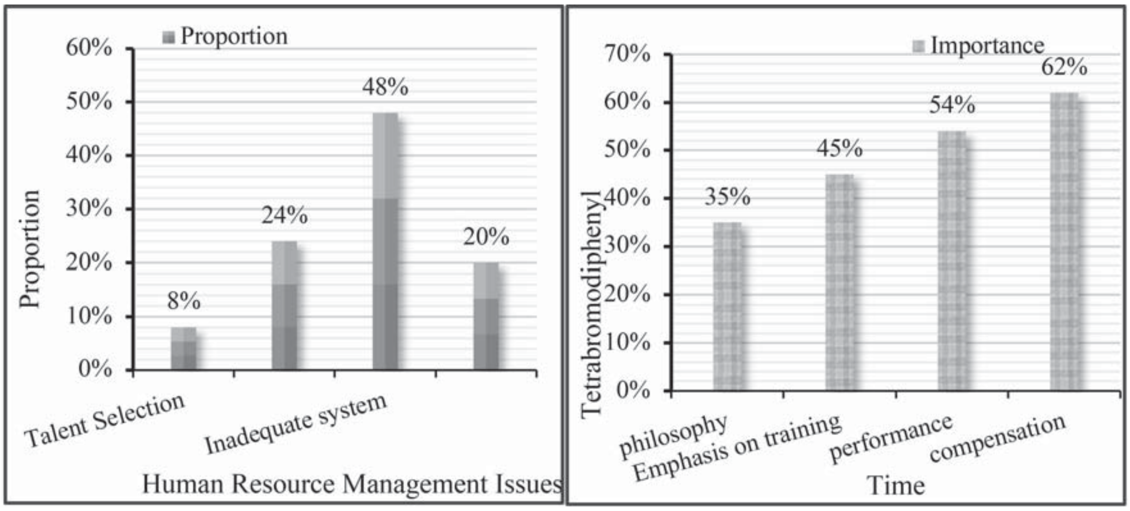

Figure 7 Current situation and innovative suggestions of human resource management in electric power enterprises.

the comprehensive utilization level of human resources, and promoting the orderly development of various management tasks. Combined with the staffing situation of electric power enterprises, a more scientific and special team planning strategy is formulated, and the internal member structure is scientifically allocated to give full play to the leading role, which is the continuous improvement of the internal management level of the enterprise. The status quo and innovative suggestions for human resource management in electric power enterprises are shown in Figure 7.

With time, the competition between companies will become a rivalry involving talents and technology. Hence, companies should pay more attention to human resources management particularly during restructuring or making changes in line with current trends. At present, there are many problems in the human resources management of Chinese electric power enterprises, which are evident in the outdated human resources management methods, insufficient innovation, poor recruitment decisions, inadequate human resource development, a lack of incentives, assessment systems, and outmoded concepts. The most serious reason is that the incentive and assessment system is not perfect, accounting for 48% of human resource problems. The lack of innovative countermeasures to improve the performance and compensation management mechanism is one of the most concerning issues, accounting for 62% of issues.

Compared with other projects, electric power engineering has the characteristics of extensive design disciplines and strong professional content. In order to improve the fundamental quality and efficiency of electric power projects, it is necessary to strengthen the management and control of the design link to ensure that the design scheme can guide the orderly development of the project. The continuous optimization and innovation of the management mechanism of a power company can help it to better cope with a major transition period, standardize the implementation behavior of the company's internal departments, and enable the enterprise to operate efficiently and achieve common development goals. At the same time, an effective management mechanism can also help enterprises to better adapt to the current development law imposed on the electricity market, enhance their credibility, and expand their development space. Under the new normal economic situation, the demand for electricity in factories is declining, but the expansion of power plants has not stopped, which has caused the power industry to have excess capacity, which is no less than that of the coal industry. Effective collaborative management methods can fundamentally improve the production and operation level of electric power enterprises. It is necessary to strengthen the cooperation between various departments during the actual production and operation period, and timely find and address the problems existing in the actual design work. However, currently, some power companies do not recognize the importance of collaborative management work, the distribution of management powers is unscientific, and it is difficult to translate management objectives and management tasks into practical work, which poses a huge threat to the future development of enterprises. Figures 8, 9, and 10 show the power generation capacity of various energy sources from February to March in a specific location since 2013.

By 2022, the thermal power plants will have an average operating time of 821 hours, 76 hours more than for the same

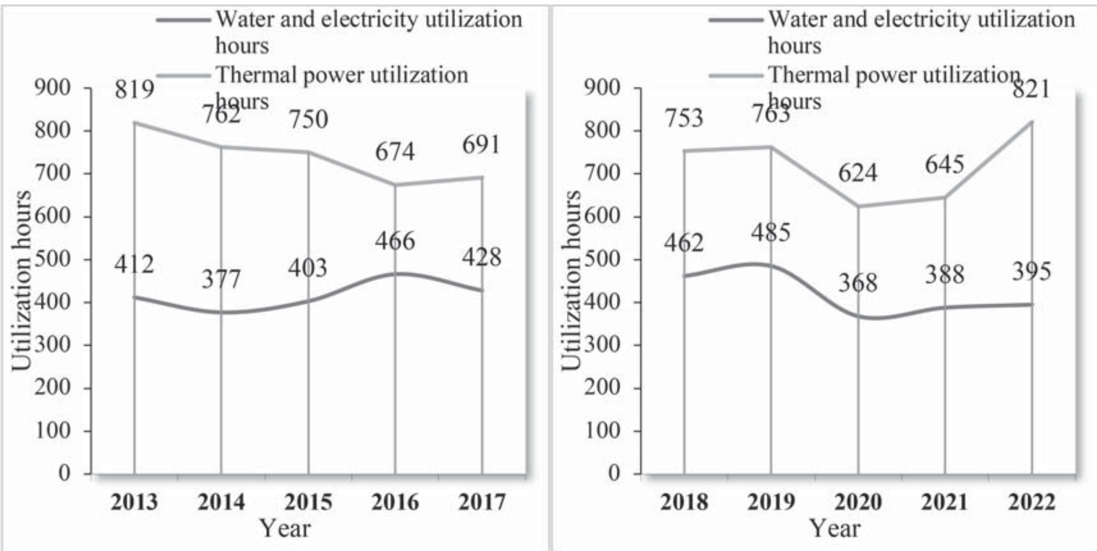

Figure 8 Utilization hours from February to March in the calendar year since 2013.

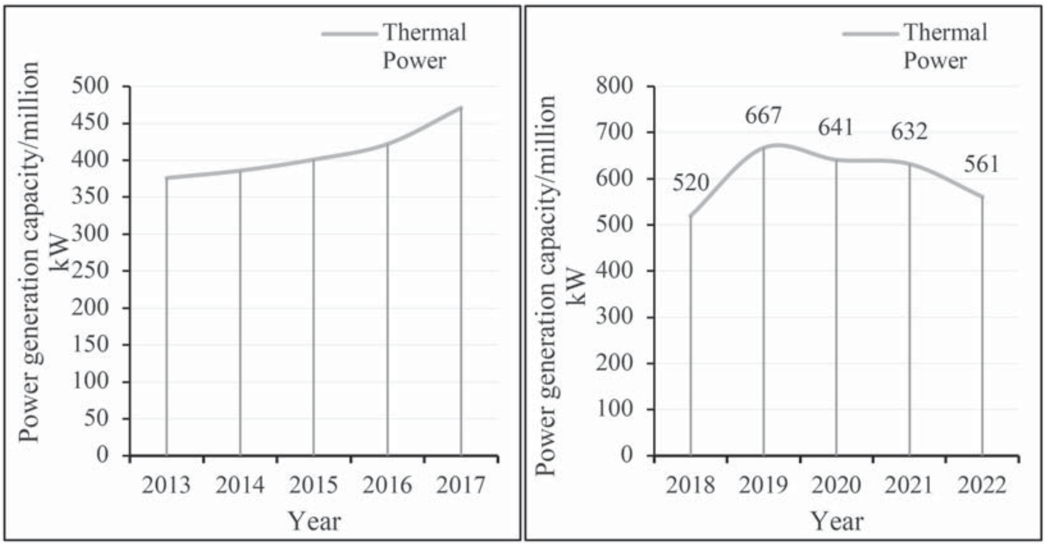

Figure 9 New thermal power production capacity.

period the previous year. The coal-fired generation equipment will be operated for an average of 852 hours, 80 hours more than for the same period the previous year; and the gas-fired generation equipment will be used for an average of 423 hours, 38 hours more than for the same period the previous year.

China's power is mainly coal-fired, and China has the largest coal mine resources and the conditions to develop thermal power. New thermal power capacity for the February–March period been increasing from 2013 to 2019, and reached 5.61 GW in 2022. However, due to the depletion of fossil fuel supplies, countries are trying to reduce the generation of thermal power, and are making serious attempts to develop hydropower, nuclear power, and high-efficiency solar power generation.

From February to March 2022, wind power production was 5.82 million kilowatts, and solar power was 11 million kilowatts. Hydropower, wind power, and solar power produced 180,000, 260,000, and 440,000 kilowatts more than the same period last year, respectively. Nuclear power produced 610,000 kilowatts less than for the same period the previous year.

At present, the focus of power generation energy has changed from thermal power generation to other clean energy power generation, which has alleviated the problems of environmental pollution and shortage of fossil fuels. However, the problem of overcapacity in the power industry has also attracted much attention. The main position of thermal power generation has not been greatly weakened, while new energy power generation has expanded rapidly. Based on this, controlling investment in new power plant projects and developing energy storage equipment is an effective mitigation channel.

5. CONCLUSION

This paper introduced the definition and research status of a large-scale optimization problem and a multi-objective

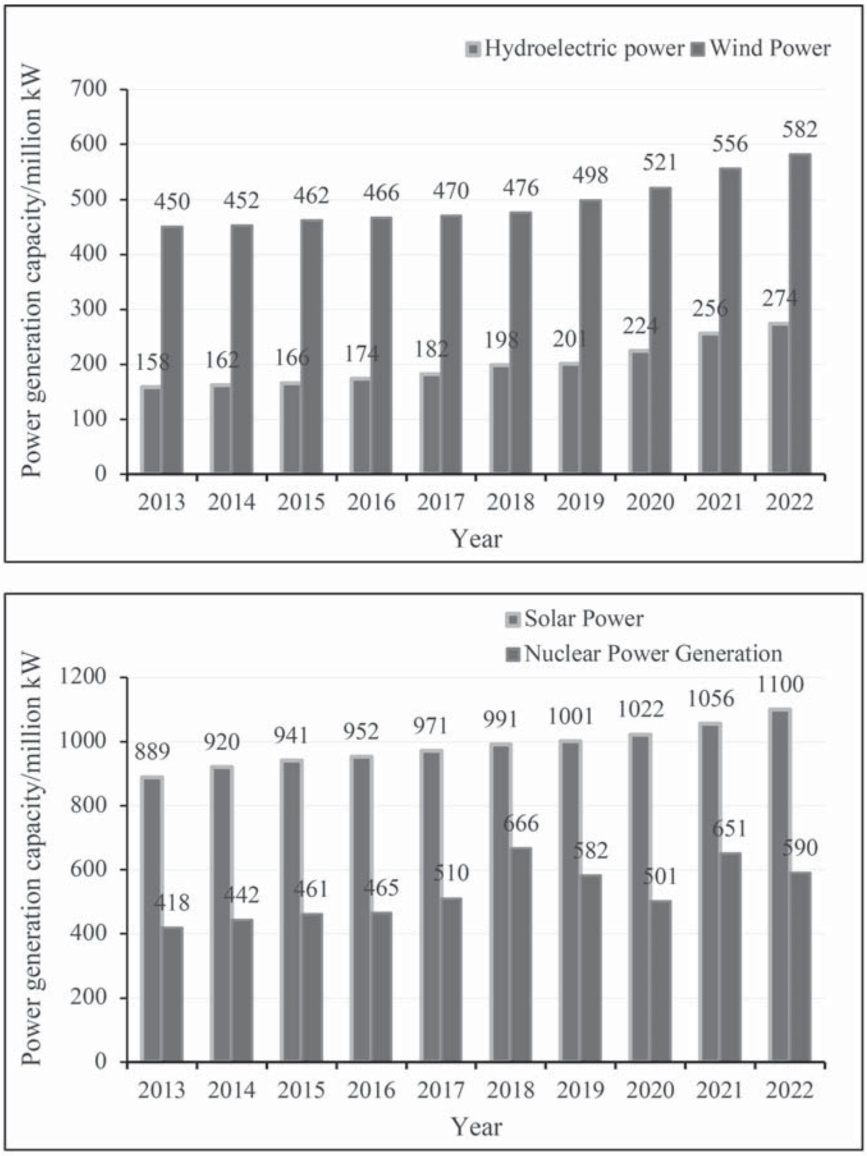

Figure 10 Power generation production capacity of clean energy, February-March 2013–2022.

optimization problem, and designed effective improved algorithms to address both problems. The development of an enterprise is largely determined by the management of its human resources. Electric power enterprises must strengthen their own management, allocate human resources scientifically and rationally, adapt to current development, and strengthen human resource management. The development goals established by the enterprise are the concrete manifestation of the realization of the strategy, and the strategy is the premise of the goal setting. In recent years, power plants have continued to expand, and the power industry in some regions has experienced excess capacity. The development of energy storage equipment is an effective solution. There is still room for improvement in the non-inferior solution selection algorithm for multi-objective optimization problems. Although the algorithm proposed in this paper obtained a more evenly-distributed solution set, it also caused higher time complexity. How to design a fast and effective selection method is still worthy of further exploration, which can be explored in combination with the characteristics of other clustering algorithms.

REFERENCES

1. Chen Z. G., Xin X., Chen K. A. collaborative optimization model for design and scheduling of multi-ship tanker fleet considering hard time windows. Dalian Haishi Daxue Xuebao/Journal of Dalian Maritime University, 2018, 44(3): 49–55.
2. Wang X., Liao W., Guo Y., Wang F., Pan Z. and Liu, D. Modeling and Decoupling of Coupling Tasks in Collaborative Development Process of Complicated Electronic Products. Transactions of Nanjing University of Aeronautics and Astronautics, 2019, v.36(05):180–190.
3. Pearce M., Mutlu B., Shah J. Optimizing Makespan and Ergonomics in Integrating Collaborative Robots in Manufacturing Processes. Automation Science and Engineering, IEEE Transactions on, 2018, 15(4):1772–1784.
4. Li T., Wang P. Dimension collaborative optimization design of a multi-closed-loop space deployable mechanism. Xi'an

Dianzi Keji Daxue Xuebao/Journal of Xidian University, 2017, 44(3):19–24.
5. Zhao C., Johnsson M., He M. Data Mining with Clustering Algorithms to Reduce Packaging Costs: A Case Study. Packaging Technology and Science, 2017, 30(5):173–193.
6. Anandakumar H., Arulmurugan R. Computational Mining Algorithms for Future Technologies. Journal of Advanced Research in Dynamical and Control Systems, 2019, 11(7): 817–822.
7. Subeesh V., Maheswari E, Saraswathy. G R.A Comparative Study of Data Mining Algorithms used for Signal Detection in FDA AERS Database. Journal of Young Pharmacists, 2018, 10(4):444–449.
8. Xu B., Tang Y., Zhu Y. Bilateral Collaborative Optimization for Cloud Manufacturing Service. Computers, Materials and Continua, 2020, 64(3):2031–2042.
9. Han X., Mandal S., Pattipati K. R. An Optimization-Based Distributed Planning Algorithm: A Blackboard-Based Collaborative Framework. IEEE Transactions on Systems Man & Cybernetics Systems, 2017, 44(6):673–686.
10. Navod Neranjan Thilakarathne, Rohan Samarasinghe, Mohan Krishna Kagita, Surekha Lanka, Hussain Ahmad. Smart Grid: A Survey of Architectural Elements, Machine Learning and Deep Learning Applications and Future Directions, Journal of Intelligent Systems and Internet of Things, 2021, Vol. 3, No. 1, pp: 32–42.
11. Rojas R. A., Garcia M., Gualtieri L. Combining safety and speed in collaborative assembly systems–An approach to time optimal trajectories for collaborative robots. Procedia CIRP,2021,97(1):308–312.
12. Guo, K. (2020). Research on location selection model of distribution network with constrained line constraints based on genetic algorithm. Neural Computing and Applications,32(6), 1679–1689.
13. Ye-Biao, Xu, Ya-Jun. On the Collaborative Game of Low-skill Manufactures Based on Value-creation. IOP Conference Series: Earth and Environmental Science,2019,237(5):52042–52042.
14. Rs A., Ab B., Sm A. The Pediatric Oncology East and Mediterranean(POEM)group–A regional collaborative platform for childhood cancer healthcare professionals. Pediatric Hematology Oncology Journal, 2020, 5(1):3–6.
15. L. Fu, Z. Jingli. Network Traffic Monitoring and Real-Time Risk Warning Based on Static Baseline Algorithm. Engineering Intelligent System, 2021, vol 29 no 3, pp. 183–189.
16. I. Butun, P. Österberg and H. Song. Security of the Internet of Things: Vulnerabilities, Attacks, and Countermeasures, in IEEE Communications Surveys & Tutorials, vol. 22, no. 1, pp. 616–644, First quarter 2020, doi: 10.1109/COMST.2019.2953364.
17. Zheng Z., Sun H., Zhang G. Multiple kernel locality-constrained collaborative representation-based discriminant projection for face recognition. Neurocomputing, 2018, 318(NOV.27): 65–74.
18. Kong Xiangyu, Zhang Xiaopeng, Zhang Xuanyong, et al. Adaptive Dynamic State Estimation of Distribution Network Based on Interacting Multiple Model. IEEE Transactions on Sustainable Energy, Apr 2022 13(2): 643–652.
19. Zhanghui X., Dao G., Jinsong Z. On the Influence Mechanism of the Under-Chassis Equipment Eccentric on the Vehicle Ride Quality of High-speed EMU Train. Journal of Mechanical Engineering, 2017, 53(6):131–137.
20. Liang X., Xiu B., Fan C. The aerospace cooperative planning model for maritime moving target continuation observation. Xitong Gongcheng Lilun yu Shijian/System Engineering Theory and Practice, 2018, 38(1):229–240.
21. Paradarami T. K., Bastian N. D., Wightman J. L. A hybrid recommender system using artificial neural networks. Expert Systems with Applications, 2017, 83(OCT.):300–313.
22. Liu G., Qi N., Chen J. Enhancing clustering stability in VANET: A spectral clustering based approach.China Communications, 2020, 17(4):140–151.
23. Yan H., Ye Q., Zhang. T. L1-Norm GEPSVM Classifier Based on an Effective Iterative Algorithm for Classification. Neural Processing Letters, 2018, 48(1):273–298.
24. Alonso-Mora J., Baker S., Rus D. L. Multi-robot formation control and object transport in dynamic environments via constrained optimization. International Journal of Robotics Research, 2017, 36(9):1000–1021.
25. Y. Bingjie, Y. Huifeng, S. Chenjun, Z. Zhi, F. Jinghang. Research on Network Security Collaborative Defense Technology Based on Swarm Intelligence and Big Data Network Security. Engineering Intelligent Systems, 2021, vol. 29 no.6, pp. 379–386.
26. Yang Z., Jin Y., Hao K. A. Bio-Inspired Self-Learning Coevolutionary Dynamic Multi-Objective Optimization Algorithm for Internet of Things Services. IEEE Transactions on Evolutionary Computation, 2019, 23(4):675–688.
27. Shin C., Cho Y., Park J. An Energy Consumption Prediction Model for Smart Factory Using Data Mining Algorithms. KIPS Transactions on Software and Data Engineering, 2020, 9(5):153–160.

Shibo Yang (1981.04-), male, (Han nationality), native place: Wuwei city, Gansu Province. Education background: Master degree. Professional title: senior engineer. Research direction: electric power information technology
E-mail:z-j505@163.com

Shaobo Liu (1988.03-), male, (Han nationality), native place: Pingliang City, Gansu Province. Education background: Master degree. Professional title: senior engineer. Research direction: electric power information technology
E-mail:huzi37812@163.com

Shuiting Du (June 1994 -), female, (Han nationality), native place: Lanzhou, Gansu Province. Education: Master degree. Professional title: assistant engineer. Research direction: electric power information technology
E-mail:18093197079@163.com

Zuohu Chen (1989.12-), male (Han nationality), native place: Baiyin City, Gansu Province. Educational background: Bachelor degree. Professional title: junior engineer. Research direction: electric power informatization
E-mail:czh8256359@163.com

Zhenfen Zhang (1988.08-), female, Han nationality, native place: Dingxi city, Gansu Province. education background: Master degree. Professional title: intermediate engineer. research direction: electric power engineering technology
E-mail: lxt6856@163.com