2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA)

Computational network energy efficiency evaluation, low-carbon scheduling, and intelligent optimization system

Yushen Wang
Department of Key Accounts
China Mobile Group Jiangsu Co. Ltd
Nanjing, China
wangyushen@js.chinamobile.com

***Abstract*—In order to solve the problems of low energy efficiency operation of computing network infrastructure in the development process of industry applications, this project takes the optimization and matching of computing network business energy consumption as the concept, and takes "resource dynamic perception, intelligent cloud edge integrated scheduling, green energy conservation" as the core. It provides AI energy consumption pre evaluation, low-carbon scheduling of computing network resources, dynamic monitoring of energy consumption, post evaluation optimization and other capabilities for applications, and achieves strategies such as optimal deployment of energy consumption nodes, Building a network energy optimization system with continuous evolution of circulation and application migration capabilities, and continuously providing efficient, low-cost, and low energy integrated network services, is of great significance for innovation and empowering various industry applications to reduce costs and increase efficiency.**

Keywords: Key words—AI energy consumption evaluation model; Network resource scheduling mechanism; Energy consumption monitoring closed-loop system

I. INTRODUCTION

Data is an important factor of production in today's society. As the "central brain" for receiving, processing, storing, and forwarding data streams, data centers have become one of the most important information infrastructure in the era of digital economy. However,at the same time, data centers are also major energy consumers. On one side, there is an urgent need for the rapid development of the digital economy, and on the other side, there is a strong concern for energy consumption. The high-quality development of data centers has attracted much attention.

The Action Plan for Green and Low-carbon Development of the Information and Communication Industry (2022-2025), issued by the Ministry of Industry and Information Technology, proposes to speed up the construction of national green data centers, guide enterprises to build green and intensive data centers, and speed up the resource integration and energy-saving transformation of "old and small dispersed" data centers[1]. By 2025, the power utilization efficiency (PUE) of newly built large and ultra large data centers nationwide will be reduced to below 1.3. Data centers are high energy consuming industries, with electricity costs accounting for 60% -70% of the total cost. The development of computing power networks should always aim for green and low-carbon development.

At present, the energy efficiency optimization of data centers mainly focuses on energy-saving design and environmental tuning for hardware and other infrastructure, including reducing infrastructure power, dynamic voltage automatic regulation, low-power chips, etc. There are few methods for evaluating energy consumption of various software, including application systems. Moreover, insufficient consideration is given to the impact of different task types on resource scheduling strategies, and there is a lack of relevant technologies that combine application software energy consumption with resource scheduling of hardware, network, and other infrastructure, and continuously optimize and evolve[2].

Therefore, this project aims to explore a comprehensive system for energy efficiency evaluation, low-carbon scheduling, and intelligent optimization of computing power networks. Based on the energy consumption evaluation model corresponding to the classification and matching of computing network tasks, before application deployment, the AI algorithm is used to evaluate the energy consumption of applications deployed at different nodes, and the optimal energy consumption nodes are selected for application deployment and distribution[3]. After application deployment, various energy optimization measures, including dynamic environment adjustment and application migration, are automatically generated through dynamic monitoring and energy consumption post evaluation. This is of great significance for supporting the green, intensive, and efficient development of computing power networks, as well as supporting cost reduction and efficiency enhancement of various industry applications[4].

II. PROJECT OBJECTIVE

This project focuses on issues such as low energy efficiency operation of computing infrastructure, and constructs a computing network energy efficiency evaluation, low-carbon scheduling, and intelligent optimization system to provide customers with low-carbon and low-cost integrated computing network services. The system's energy consumption level benefits are converted into cost advantages, and innovative information technology methods are used to promote the achievement of China's dual carbon strategy goals. By building dynamic monitoring, effective evaluation, and intelligent analysis capabilities for computing network energy consumption, we provide efficient and flexible resource scheduling and energy optimization strategies, intelligently

979-8-3315-0976-7/25/$31.00 ©2025 IEEE

recommend intelligent solutions such as energy consumption and cost priority to users, meet the differentiated needs of various businesses, create efficient, low-carbon, and low-cost integrated computing network services, and continue to promote the green and intensive development of computing networks, Innovative in improving the overall "green and low-carbon" level of application systems and computing power networks[5].

The achievements of this project can be widely adapted to various application scenarios, serve various vertical industries such as power, steel, and chemical industry, continuously improve resource utilization, reduce overall energy consumption, help customers further achieve the goal of "cost reduction, quality improvement, and efficiency increase", support the acceleration of industrial Digital transformation, release the core value of computing network for the overall digital economy development and industrial digital intelligence, and help to improve the overall level of social energy efficiency, Improve the green and low-carbon operation capacity of new information infrastructure, and accelerate the realization of "carbon peaking and carbon neutrality".

III. System Function and Scheme

A. *System function*

The computing power network energy efficiency evaluation, low-carbon scheduling, and intelligent optimization system is centered on "resource dynamic perception, intelligent cloud edge integrated scheduling, green energy conservation, and security and credibility". Based on the computing power network resource perception and multi factor collaborative optimization scheduling, application oriented energy consumption evaluation, energy consumption measurement, node selection, energy consumption optimization, and other methods are used to achieve energy consumption evaluation for applications The ability to measure energy consumption and schedule software based on energy consumption plans will continuously improve the green, intensive, and efficient development level of computing power networks[6]. The functional framework is shown in Figure 1:

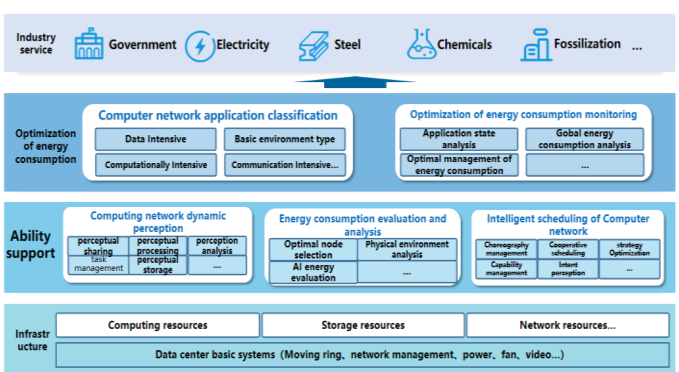

Figure 1 System functional framework

Calculate the application energy consumption and environmental energy consumption generated by deploying applications on different servers and rack units on a node by node basis, and select the most energy-efficient node for application deployment[7]. The data flow diagram is shown in Figure 2:

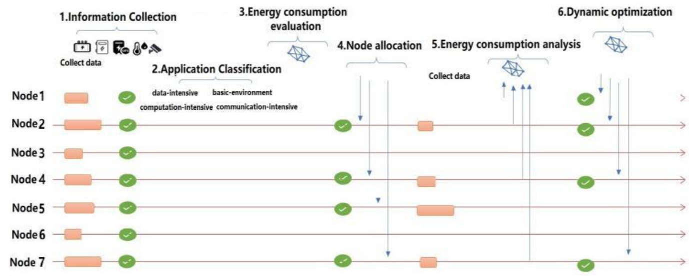

Figure 2 System Data Flow Diagram

The main functions of the system include:

- Basic data access

Provide interface adaptation and collection tasks for dynamic environment data, device data, application data, etc. In order to meet the requirements of high scalability and reusability, a plug-in approach is adopted, and different adapters are constructed to meet the data access needs in different scenarios.

- Platform Capability Support

Unified access, storage, analysis, management, and sharing of the state of all domain elements such as computing network resources, services, and applications, in order to provide data and scheduling event triggering for orchestration and scheduling; As well as practical business scenarios, it provides unified arrangement, management, scheduling and execution of various computing network resources, services, applications, etc., and decouples, abstracts, and encapsulates capabilities, rules, policies, and solutions to form a fine-grained and highly available computing network service capability with universality and flexibility.

- Energy consumption optimization operation

Classify network applications, match one or more energy consumption evaluation methods, and generate one or more energy consumption evaluation standards; Obtain software energy consumption test data based on one or more energy consumption evaluation methods; Compare the software energy consumption test data with the energy consumption evaluation standard, get the software energy consumption test model, and generate software deployment optimization suggestions; Real time or periodic monitoring of the application and performance parameters during software operation, generating energy consumption measurement results and optimization plans, continuously improving the green and low-carbon level of the computing network.

B. *Key technical solutions*

The key technologies of the system include: dynamic perception of computing networks and intelligent scheduling of computing networks. Computational network dynamic perception mainly provides unified access, storage, analysis, management, and sharing of the state of global elements such as resources, services, and applications, thereby providing data and scheduling event triggering for the orchestration center.

Computer network intelligent scheduling is mainly responsible for unified arrangement and management of resources, businesses, services, applications, etc. It schedules and executes according to the needs of business scenarios, and decouples, abstracts, and encapsulates capabilities, rules, policies, and solutions to form a universal and flexible basic service[8]. The key system capabilities are shown in Figure3:

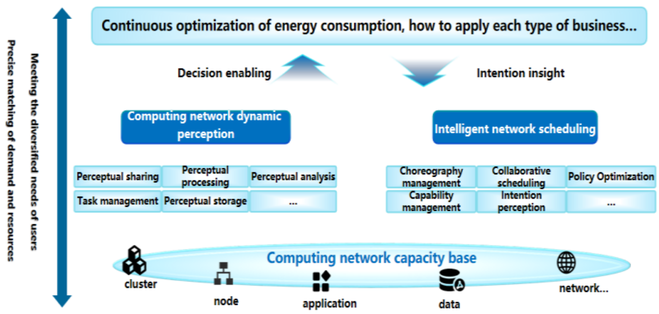

Figure 3 Key System Capabilities

Among them, the core of dynamic perception of computing network is energy consumption assessment and analysis, while the core of intelligent scheduling of computing network is energy consumption assessment and optimization. Relevant contents are described as follows:

- Evaluation and analysis of energy consumption

Classify applications, match one or more energy consumption assessment methods, and generate one or more energy consumption assessment standards; Obtain energy consumption test data of the software based on one or more energy consumption evaluation methods; By comparing the software energy consumption test data and energy consumption evaluation standards, the energy consumption test model of the software can be obtained, and software deployment optimization suggestions can be generated according to the energy consumption test model; Real-time or periodic monitoring is conducted on the application and performance parameters of the software during operation, and energy consumption measurement data is generated according to the monitoring data.

1) AI energy consumption evaluation

The AI energy consumption evaluation model mainly includes: application energy consumption evaluation and hardware energy consumption evaluation.

Application energy consumption evaluation: Based on application hardware deployment time, and analysis of application classification results, obtain the application load under different categories and predict the calculation time occupied by deployment in different locations based on AI algorithm; Hardware energy consumption evaluation: Develop a hardware oriented energy consumption analysis that evaluates the energy consumption of applications deployed on different server and rack units using AI algorithms through the hardware application energy consumption ternary relationship group, and draws a graph of energy consumption patterns for hardware load.AI energy consumption evaluation optimization is shown in Figure 4:

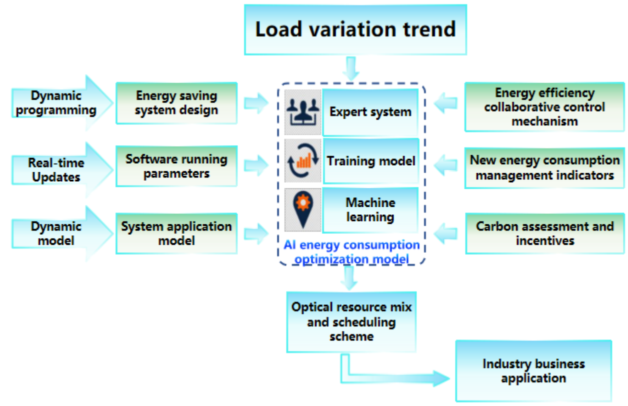

Figure 4 AI Energy Consumption Evaluation Optimization

The specific energy consumption evaluation algorithm is as follows:

By collecting load information that is most closely related to system energy consumption, mainly including energy consuming components such as CPU, memory, storage, and network. Due to the differences in the contribution of various resource factors obtained from sampling to energy consumption, the following m regression model containing dependent variables is established:

$$\begin{cases} y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_m x_{im} + \varepsilon_i \\ E(\varepsilon_i) = 0, \ \mathrm{Var}(\varepsilon_i) = \sigma^2 \end{cases}, \quad i = 1, \ldots, n \tag{1}$$

In formula (1), y_i is the observed real-time energy consumption, $\beta_0, \beta_1, \beta_2, \ldots, \beta_m$ is the regression coefficient, ε_i is the unobservable random error, $x_{i1}, x_{i2}, \ldots, x_{im}$ is the observed value of the usage rate of each component, that is, the regression factor; $\beta_0, \beta_1, \beta_2, \ldots, \beta_m$ Reflects the y contribution of factors x_i to the dependent variable.

Further, incorporating data related to energy consumption, such as system utilization, into a multiple linear regression model, and using the least squares method to estimate the regression coefficients $\beta_0, \beta_1, \beta_2, \ldots, \beta_m$, the following energy consumption calculation formula is obtained in formula 2:

$$\begin{aligned} Q(\beta_0, \beta_1, \beta_2, \ldots, \beta_m) &= \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \ldots - \beta_m x_{im})^2 \\ &= (y - \beta_0 - X\beta)^T (y - \beta_0 - X\beta) \end{aligned} \tag{2}$$

Based on the analysis of application classification results, the application loads under different categories are obtained, and the M-element regression equation is formed in formula 3:

$$y' = \beta_0' + \beta_1' x_1 + \beta_2' x_2 + \ldots + \beta_m' x_m \tag{3}$$

Based on the above algorithms, an application energy consumption model is established based on the utilization rate of system resources, and reasonable screening of energy consumption factors is achieved to reduce model complexity while ensuring accuracy. Comparison between prediction results of AI energy consumption evaluation model and actual measurement is shown in Figure 5:

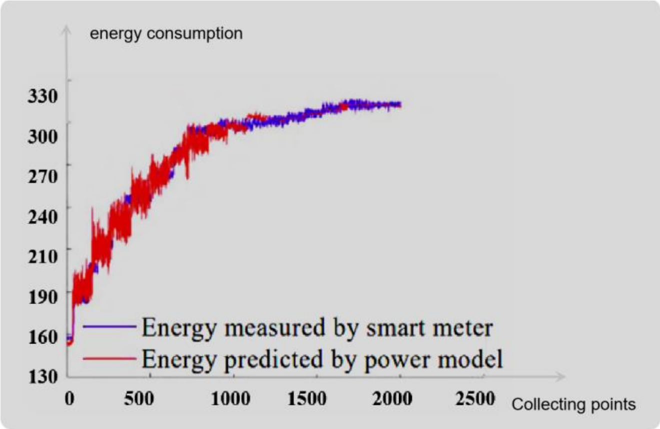

Figure 5 Comparison between prediction results of AI energy consumption evaluation model and actual measurement

2) Physical environment analysis

Physical environment analysis mainly evaluates the environmental changes that will occur when applications are deployed in different locations, resulting in additional energy consumption.

Collect dynamic environment data before application deployment, analyze whether the current dynamic environment conditions meet the requirements of the group's specifications for the physical environment; Evaluate the heat generated by devices and changes in ambient temperature when applications are deployed on hardware of different servers and rack units based on AI algorithms; Analyze the additional energy consumption required to adjust the physical environment to meet regulatory requirements based on the changes in the environment after the application deployment is evaluated.

3) Optimal node selection

Based on the aforementioned AI energy consumption and environmental energy consumption, analyze the energy consumption cost of deploying applications at each node, and select the most energy-efficient node for application deployment; Calculate the application energy consumption and environmental energy consumption generated by deploying applications on different servers and rack units based on nodes, and obtain the total energy consumption of the application deployment; Sort the total energy consumption from low to high, and select the node with the lowest energy consumption as the final node for application deployment[9].

- Evaluation and optimization of energy consumption

Energy consumption evaluation optimization is shown in Figure 6. In the periodic operation or completion of the software, generate the software energy consumption measurement model, report and optimization scheme, and optimize the software energy consumption evaluation methods, standards and computing resources energy consumption evaluation model.

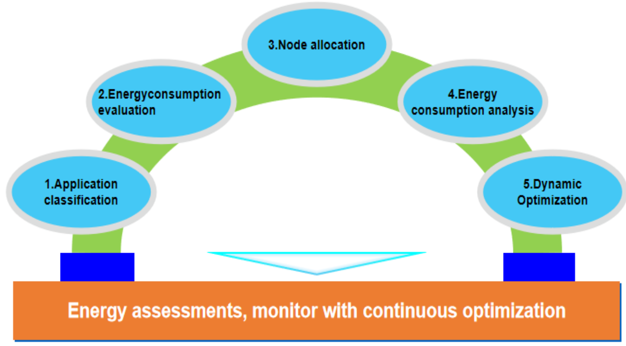

Figure 6 Energy consumption evaluation optimization

4) Classification of Computer Network Applications

Classify the software, match one or more energy consumption assessment methods, and generate one or more energy consumption assessment criteria:

Data-intensive application: a large number of independent data analysis and processing jobs are distributed in different nodes of the loosely coupled computer cluster system, and the amount of data to be processed is huge and rapidly changing, often distributed and heterogeneous.

Computational-intensive applications: Refers to tasks with high CPU and GPU resource consumption and high computational load, with the pursuit of computing speed as the main indicator, such as high-definition decoding of videos.

Communication-intensive applications: applications that require high transmission speed, such as symmetric encryption, firewalls, network virtualization, etc.

I/O intensive applications: With high CPU efficiency and low occupancy, most of the time is spent waiting for I/O operations, such as web applications.

Information such as the amount of application data processing, CPU and GPU resource consumption, data traffic and I/O operation collected is input into the application classification model; According to the pre-trained model, the deployed applications are classified according to the input information.

5) Energy consumption monitoring optimization

Generate software energy consumption measurement models, reports, and optimization plans during or after periodic software operation, and optimize software energy consumption evaluation methods, standards, and computational resource energy consumption evaluation models.

a) Application Status Analysis

Application status analysis includes the operational status, runtime, transferability, and timeliness of all deployed applications in the current data center.

Operation status analysis: Analyze based on the characteristics of different types of applications, monitor the operation status of each application, and analyze the expected duration of equipment occupation;

Application transferability analysis: Analyze whether each application has transferability and whether there is room for scheduling and optimization;

Application timeliness analysis: Analyze the requirements of each application for timeliness to determine whether to schedule the execution time of the application;

Application analysis results: Generate comprehensive analysis result data for each application based on the above application operation status data, application migration data, and application timeliness data.

b) Global energy consumption analysis

Based on application state analysis, obtain comprehensive analysis results of each application, and conduct global energy consumption analysis in conjunction with the current environment.

Dynamic loop optimization analysis: Based on comprehensive analysis results of different applications, dynamically generate dynamic loop data adjustment strategies and measures.

Application optimization analysis: Overall analysis of whether applications can be postponed for execution when the data center is idle, and whether transferable applications can be migrated to nodes with better energy consumption for task execution.

Energy consumption optimization analysis: comprehensively analyze the current load situation of various equipment applications, application energy consumption, hardware energy consumption, and other factors, analyze the optimization space of each factor, and obtain the comprehensive optimization results of energy consumption.

c) Energy consumption optimization management

Comprehensively evaluate the impact of different dynamic loop operation instructions after execution, and generate corresponding dynamic loop scheduling execution sequence and execution time strategies; Comprehensively evaluate the impact of application migration, application delay execution, and other operation instructions after execution, and generate corresponding application scheduling execution sequence and execution time strategies.

C. Business process

This topic mainly provides efficient, flexible and intelligent energy consumption optimization strategies by building systematic capabilities of dynamic monitoring, effective evaluation and intelligent analysis of energy consumption of computing network, which helps to continuously reduce the energy consumption operation cost of computing network, and enables various upper layer computing network application industries and business scenarios.

The specific business process is shown in Figure 7:

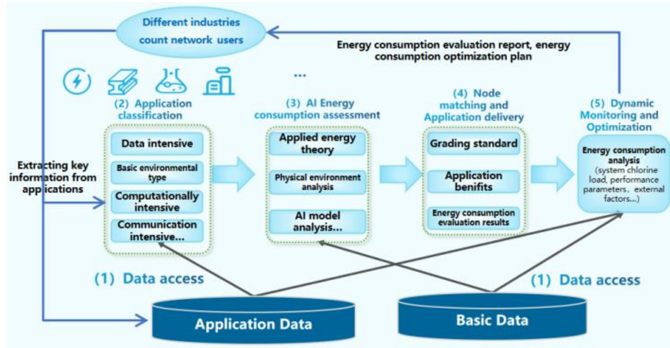

Figure 7 Business process

IV. PROJECT APPLICATION

By establishing the overall solution of energy efficiency assessment, low-carbon scheduling and intelligent optimization of computing power network, this project provides the whole process service with the mode of "platform + service" to continuously improve the energy consumption level of various industries. It is expected to be popularized in the electric power, steel, chemical and other industries, and continue to expand the market coverage. It is expected that the contract amount will reach 3 billion. To realize the double optimization of economic benefits and environmental benefits.

The project achievements have been applied and implemented in 10+ vertical industries such as electric power, steel and chemical industry, supporting the acceleration of industrial digital transformation, and continuously improving the efficiency, intelligence and lean level of computing network services. In terms of smart factories, it has effectively supported thousands of "city-province-national" smart 14 factories to improve the green and low carbon level, further realizing cost reduction and efficiency increase. For example, Inner Mongolia Autonomous Region online energy monitoring project, access to more than 700 key energy consumption units, support Inner Mongolia Institute of Metrology to build the first national carbon metering center; Xinjiang Uygur Autonomous Region online energy consumption monitoring and carbon metering project, access to more than 600 key energy consumption units.

At the same time, the project helps different types of computing network customers to further achieve the goal of "cost reduction, quality improvement and efficiency". While meeting the same business needs, the project also guarantees the QoS and energy efficiency demands of the business, and achieves a 30% ~ 40% increase in computing efficiency level and a >50% reduction in overall costs (including electricity expenditure, resource input, operation and maintenance costs, labor costs, etc.). Reduce carbon emissions by 40% ~ 50%, and realize the surplus carbon quota obtained from emission reduction in the carbon trading market to obtain emission reduction income, and continue to promote the intensive, efficient and green development of the computing network. Take Lianyungang Petrochemical Data Center (6500 standard cabinet) as an example, through energy consumption optimization, electricity saving 68.42 million kWh/ year, reduce carbon emissions of 65,600 tons/year; By reducing the cost of electricity, resource investment, operation and maintenance costs, etc., it helps reduce the overall expense of the enterprise by 5.1 million yuan/year.

This project also has the following social value:

- Promotion of market employment

Through green and low-carbon transformation, the output effect and factor substitution effect of enterprises can be improved, thereby increasing the employment rate by 5% to 10%[10]. Among them, the output effect increases the total number of employed people, while the factor substitution effect is manifested as the substitution of different employment positions; At the same time, technological and model innovation will expand enterprise output, promote factor substitution, and thereby strengthen its role in promoting employment.

- Promotion of industrial development

By improving the resource utilization rate by 30%~40%, reducing the overall energy consumption, improving the balance between cost, efficiency and quality, promoting the high-quality development of computing power network in vertical industries and social scene services, and helping to improve the overall social energy efficiency level, improve the green and low-carbon operation capacity of new information infrastructure, so as to accelerate the realization of "carbon peak, carbon neutrality".

V. Conclusion

This project establishes an overall solution for the energy efficiency assessment, low-carbon scheduling and intelligent optimization of computing power networks. It provides full-process services in a "platform + service" model, which can help enterprises reduce carbon emissions by 40% to 50% and overall energy consumption costs by more than 50%, assisting enterprises in further reducing costs and increasing efficiency. Through promotion and application in industries such as power, steel, and chemical engineering, it continuously offers efficient, low-carbon and low-cost integrated computing and network services, accelerating the realization of the "dual carbon" goals.

Acknowledgment

Thank you to Inspur Communication Information System Co., Ltd. for its strong support in the implementation process of this project, which has played a positive role in the functional implementation of the system.

References

[1] Zhang Shu, Application of Artificial Intelligence in hospital supply chain Management in the context of carbon neutrality, Medical Information, No. 6, 2023, pp. 41-45

[2] Xie Haizhi, Wang Huarong, Research on energy saving and Emission reduction of 5G base Station based on intelligent power saving Technology, Communication Power Technology, No. 010, 2021, pp. 31-35

[3] Avnet, Smart Energy Metering Technology will help achieve carbon neutrality, Software and 15Integrated Circuits, No. 5,2021, pp. 16-17

[4] Sun Guotong, Tang Yan. Research on Green and Secure Upgrade Technology Solution for Core Data Center [J]. Guangdong Communication Technology, 2024, 44(5), pp.2-6

[5] Zhang Xiaoling. "Energy Saving Optimization Analysis of Green Data Center." East China Science and Technology (Comprehensive) 000.004 (2021), pp.1-2

[6] Gong Weiwei, Li Shuangshuang, Li Jiajia, et al. Low-carbon green practice of building Communication basic services [J]. Shandong Communication Technology, 2022, 42(3), pp.44-47

[7] Xing Qiangqiang Research on Grid based Complaint Warning and Scheduling System Based on AI Intelligent Delimitation [J] Cybersecurity and Informatization, 2024 (7)

[8] Chen Luyang Research on Intelligent Dispatch System for Fire Communication Based on AI Algorithm [J] Internet Weekly, 2024 (6): pp. 36-38

[9] Yang Mingxuan, Hong Xuehai, Tang Hongwei Artificial intelligence computing power scheduling based on task resource demand prediction [J] High tech Communications, 2024, 34 (5): pp. 475-485

[10] Wu Guangshuai, Shi Wenjuan, Zhang Wei, etc The Application of AI Technology in Power Dispatching Automation [J] Digital Boutique World, 2023 (4)