[1,*]Zuowei Chi
[2]Bicheng Tan
[3]Yunfei Chu
[4]Daye Wang

"Comprehensive Integration" Method Based on the Bayesian Network and the Big Data Using in the Construction of New Energy System

Journal of Electrical Systems

***Abstract:* -** With the continuous application of the big data technology, "Comprehensive Integration" theory is an approach to management based on Rooted theory, which has been successfully applied to the management of large enterprises and research organizations of all kinds. State Grid Corporation of China in the management process of the construction of a new energy system, with a strong innovative, large-scale investment, many participating units, urgent needs and other characteristics of its management capabilities put forward higher requirements. In this paper, a system integration principle based on the Bayesian probabilistic approach and the big data is proposed to realize the engineering application of the theory of "Comprehensive Integration". By splitting the management system of the new energy system construction, the unit composition of the system can be interpreted to form a complete system and node interaction Bayesian network structure. The algorithm proposed in this paper is implemented in Netica, a commercial statistical analysis software, and its effectiveness is verified, which provides an important theoretical basis for the engineering practice of the theory of "Comprehensive Integration". Finally, the effectiveness of the Dynamic probabilistic network method was quantitatively verified in a lithium battery pack energy storage management example.

Keywords: "Comprehensive Integration" Theory, Rooted Theory, State Grid Corporation of China, Construction of a New Energy System, Bayesian Probabilistic Approach, Big Data.

I. INTRODUCTION

The integration of big data technology with the national power grid enables intelligent monitoring, predictive maintenance, and optimization of energy loads, promoting the integration of renewable energy and enhancing the efficiency and reliability of the power system. The construction of the new energy system, compared with other power system projects, is characterized by strong innovativeness, large-scale investment, many participating units and urgent needs. In this process, the management model of power grid enterprises faces many challenges, not only to adapt to technological innovation, environmental protection challenges and new market competition mode, but also to meet the needs of different stakeholders [1]. For State Grid Corporation of China, the key factor in the construction of a new energy system is how to formulate and implement a clear and reliable strategic policy and management model based on the actual situation, in order to build a large platform for cross-border cooperation, to achieve the purpose of proactively servicing the overall pattern, and to fully tap the potential of various elements [2, 3]. In the actual implementation process, not only need to organize and coordinate many departments within the company, at the same time also need to plan with many universities, research institutes on new energy development, consumption and other scientific theories of innovation and R & D, to build with the industrial system and the backbone of the enterprise depth of cooperation matrix, which puts forward higher requirements on the management capacity of the State Grid Corporation of China, the innovation of the scientific management model is imperative [4, 5].

The theory of "Comprehensive Integration" is a multifaceted management theory that includes various methods and tools to help companies consider and analyze the interrelationships between various factors in order to achieve comprehensive and coordinated management [6-8]. If the theory and method of "Comprehensive Integration" can be incorporated into the management system of State Grid Corporation of China. and the innovation of management method can be realized for the planning and construction project of new energy system, and more efficient cooperation and operation can be realized among various departments of the company, universities, research institutes and key enterprises, so as to promote the realization of a dynamic balance among the external environment, internal conditions and the capability of the enterprise, and then it will be possible to provide reliable

[1] State Grid Jilin Electric Power Co., LTD, Changchun, Jilin, China
[2] State Grid Jilin Electric Power Co., LTD, Changchun, Jilin, China
[3] State Grid Jilin Electric Power Co., LTD. Economic and technological research institute, Changchun, Jilin, China
[4] Jilin Power Trading Center Co., LTD, Changchun, Jilin, China
*Corresponding author: Zuowei Chi

guarantee for the subsequent development of new energy on a larger scale, higher level of energy consumption, more diversified use and higher safety coefficient [9].

Rooted theory is an important means of realizing the theory of "Comprehensive Integration", which is a method of using empirical material to form theoretical results through bottom-up refinement and generalization. Rooted theory requires the support of empirical materials, on the basis of which concepts and categories can be extracted and theories can be formed. The research method of Rooted theory has been widely applied in many fields such as psychology, sociology, anthropology and management science. Rooted theory adheres to the ontology of constructivism and the epistemology of relativism, and is committed to the development of new theories and the development of new insights and understandings of phenomena, with the research process emphasizing the interaction of data collection and information analysis.

II. Based on Dynamic Probabilistic Networks to Achieve "Comprehensive Integration" Theory and Methodology

Dynamic probabilistic networks, with their unique uncertainty knowledge representation and rich probabilistic expressive ability, have significant features in describing nonlinearity, temporality, evolution, and uncertainty, and have been gradually applied to the fields of classification, decision-making and prediction, information recovery, and expert systems [10]. The time series in a dynamic probabilistic network is discrete and is represented by a time slice, on each time slice is a probabilistic network, and there are associative dependencies between neighboring time slices with different nodes. Since there are many difficulties in learning, modelling and reasoning about dynamic probabilistic networks for real-world problems, in recent years a number of experts and scholars have worked on simplifying them, given some constraints, in order to make them more widely used [11].

Mathematical modelling using dynamic probabilistic network models is a viable way of modelling and analyzing complex systems, as well as modelling and analyzing uncertainty and risk in the context of the "Comprehensive Integration" theoretical approach [12].

Dynamic Bayesian Network (DBN) is a probabilistic graphical model that can be used to represent stochastic processes over time [13, 14]. When modelling a Dynamic Probabilistic Network, the system needs to be decomposed into a series of nodes and the correlations between them established, and then the conditional probability distributions between them, as well as their evolution in time, are represented using the Bayesian Network approach.

Mathematical modelling using dynamic probabilistic network models requires consideration of a variety of factors, such as the dynamics, uncertainty, complexity and diversity of the system. At the same time, adequate observation and data collection of the system is required to obtain key parameters and characteristics of the system for parameter estimation and validation of the model.

The final state of the management of the State Grid Corporation of China can be regarded as the formation of many elements with the overall function of the system, and the principle of system integration is to associate these originally independent and heterogeneous elements together, and gradually integrate them in stages to form a system with a holistic function. The system fusion principle formula can be expressed as:

$$\langle G,\Theta\rangle = F\left[\left(S_1,\cdots,S_m\right),\left(R_1,\cdots,R_n\right)\right] \tag{1}$$

Where ($S_1,\cdots,S_m$) denotes the system components, ($R_1,\cdots,R_n$) denotes the relationship between the system components, F denotes the system fusion principle function, G denotes the overall system structure, and Θ denotes the overall system objective parameters.

In order to consider the correlation between the stages, a dynamic probabilistic network can be used to model that system fusion principle. In this case, a Bayesian probabilistic approach can be used for system fusion.

Based on the Bayesian network idea, this paper describes the Bayesian network of system fusion by $B=\langle G,\Theta\rangle$, where G is a directed graph whose nodes of directed arcs correspond to each system element $X_1,\cdots,X_{\text{n}}$, and the directed arcs themselves represent correlations between the elements. Let $\pi(X_i)$ be a parent node of X_i and $A(X_i)$ be a non-parent node of $\pi(X_i)$. Then, for all X_i , X_i and its parents are independent of $A(X_i)$ .The structural parameter Θ represents a set of parameters describing the network, which for each element X_i , is $\Theta_{X_i|\pi(X_i)} = P(X_i \mid \pi(X_i))$. At this point, the joint probability density function of the whole system can be expressed as:

$$P_B(X) = \prod_{i=1}^{n} P(X_i | \pi(X_i)) = \prod_{i=1}^{n} \theta_{X_i|\pi(X_i)} \tag{2}$$

System integration is to also find the Bayesian network $B = \langle G, \Theta \rangle$ to achieve the goal of maximizing holistic functionality.

The micro-complexity and macro-integrity of the system are constantly changing in the above principle of system integration. In order to describe this phenomenon, this paper applies the most general method of describing complexity — the information entropy method — to illustrate the relationship between micro-complexity and macro-integrity of the system and the changes in different stages. The reason for the use of this method is not only that it reflects both informational and relational complexity, but also that it describes the relationship between complexity and the goal of engineering integrity. Let the microstructure of the system at a certain stage be $G[S,R]$, and the probability of achieving the engineering goal under this structure be $P(\Theta | G)$, at this point, the information entropy of the system can be expressed as:

$$E(\Theta|G) = -\sum_{(S_i,R_j)\in G} \sum_{\theta_k \in \Theta} P(\Theta, \quad) \log P(\Theta|G) \tag{3}$$

III. Application of Dynamic Probabilistic Network Method in State Grid Corporation of China

For the State Grid Corporation of China, the construction of a dynamic probabilistic network model can realize the application of the “Comprehensive Integration” theoretical approach, which can be carried out according to the following steps:

(1) Definition of system scope and objectives: Firstly, it is necessary to define the system scope and objectives of the State Grid Electricity Company Limited enterprise, including the various departments and business areas involved, as well as the objectives and targets to be optimized.

(2) Identification of key nodes: Depending on the scope and objectives of the system, key nodes, i.e., the parts or variables in the system, and the correlations between them, need to be identified. Key nodes can be various energy types, grid construction, policy environment, etc., and variables can be production capacity, consumer demand, policies and regulations. All key nodes and variables need to be considered and their interrelationships and impacts identified.

(3) Data collection: Sufficient data need to be collected to model the dynamic probabilistic network, including key parameters such as conditional probability distributions between nodes and transfer probabilities. In addition, uncertainty and risk factors need to be considered and risk and decision analyses need to be performed.

(4) Modelling: Based on the collected data and modelling requirements, a dynamic probabilistic network model can be built using relevant software. Based on probabilistic and statistical methods, this model can be used to predict future trends and possible risks to support decision-making.

(5) Model validation and optimization: After the model has been built, it needs to be validated and optimized, including considerations of accuracy, stability, interpretability and reliability of the model, and it also needs to be updated and amended to adapt to changing environments and requirements.

(6) Application models: Based on the results of the models, applications such as decision support, optimization analysis, risk assessment, etc., can be carried out in order to realize the application of "Comprehensive Integration" theory and methods

IV. Auxiliary Model Construction Based on Netica Commercial Simulation Software

In the case of management system integration of State Grid Corporation of China, the entity hardware system, management element virtual system and organizational soft system included in the system are assembled into a whole system, and based on the mapping relationship between the system constituent elements and the Bayesian network, the system integrality and the above three systems are transformed into the root nodes and intermediate nodes of the Bayesian; at the same time, the constituent functions of the sub-systems are split and transformed into the Bayesian network of the nodes of each intermediate process, based on the above way, the whole system is split, the unit composition of the system can be interpreted, and through the intrinsic connection of each unit node, the associated nodes will be connected in series with directed arcs to form a complete system and node interaction of the Bayesian network structure, as shown in Figure 1.

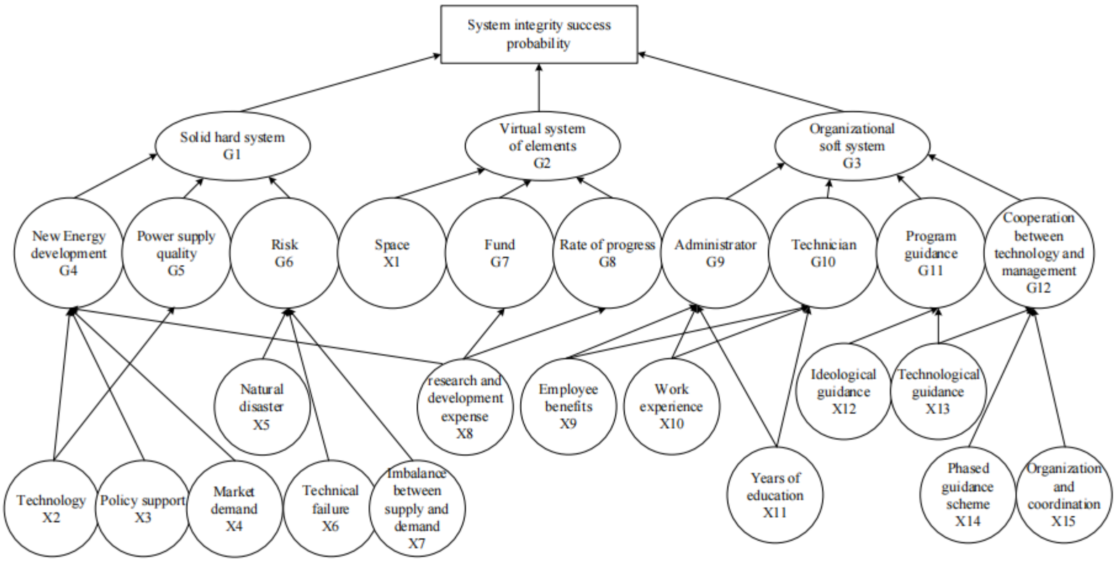


Figure 1: Bayesian Network Topology Structure of the Stable Probability of System Integrity

Determination of the root node prior probability: Bayesian network constructed in this paper has two states, corresponding to “Y” and “N”. The probability of an event with a probability value of “Y” occurring has been calculated on the basis of actual situation in the case and historical experience, and the probability of an event with a probability value of “N” not occurring is certain. It can be obtained by the theorem that the event of each node in the Bayesian network according to the sum of all state probabilities is 1, so that the probability value of each event state is “N”. The prior probability of each node is shown in Table 1.

Table 1: The Prior Probability of the Root Node of the Bayesian Network

Node number	$P(X_i=1)$	$P(X_i=0)$	Node number	$P(X_i=1)$	$P(X_i=0)$
X_1	0.64	0.36	X_9	0.402	0.598
X_2	0.906	0.094	X_{10}	0.929	0.071
X_3	0.956	0.044	X_{11}	0.81	0.19
X_4	0.838	0.162	X_{12}	0.9	0.1
X_5	0.048	0.952	X_{13}	0.817	0.183
X_6	0.065	0.935	X_{14}	0.95	0.05
X_7	0.1	0.9	X_{15}	0.598	0.402
X_8	0.894	0.106			

Determination of non-root nodes conditional probability table: the non-root node probability value is fundamentally different from the root node, mainly reflected in the source of the value, the data for the root node basically comes from aggregating the probability of occurrence of real situations, while the non-root node probability is conditional probability, which is required to be calculated by conditional probability function of root node connected to itself. Then based on the layer-by-layer mapping of the conditional probability function, the non-root node probability value is finally obtained. For the convenience of the following analysis, nodes G_{11} and G_6 are used to represent the relationship of “with” “or”, and the calculated conditional probability table is shown in Table 2 and Table 3, with the conditional probability relationship of node G_{10} as the representative, and its probability distribution is shown in Table 4, and similarly, conditional probability distribution table of other non-root nodes can be deduced.

Probabilistic inference is performed using the Bayesian network analysis software Netica, which is a software that can efficiently calculate the model parameters and clearly display the relationship graphs in network modelling analysis, and can satisfy the probabilistic calculation needs in Bayesian network analysis. Therefore, this software is used in this section to demonstrate the automated probabilistic solving process. Based on the constructed Bayesian network graph, Figure 1, the known root node probabilities and conditional probability formulas of non-root nodes are entered into the software, and then the occurrence probabilities of the parent nodes are solved using the Netica software. The final result is shown in Figure 2.

Table 2: Conditional Probability of Non-root Node G_{11} of Bayesian Network

Root node status			Conditional probability of node G_{11}	
X_{12}	X_{13}	X_{14}	$P(Y)$	$P(N)$
0	0	0	0	1
0	0	1	0	1
0	1	0	0	1
0	1	1	0	1
1	0	0	0	1
1	0	1	0	1
1	1	0	0	1
1	1	1	1	0

Table 3: Conditional Probability of Non-root Node G_6 of Bayesian Network

Root node status			Conditional probability of node G_6	
X_5	X_6	X_7	$P(Y)$	$P(N)$
0	0	0	0	1
0	0	1	1	0
0	1	0	1	0
0	1	1	1	0
1	0	0	1	0
1	0	1	1	0
1	1	0	1	0
1	1	1	1	0

Table 4: Conditional Probability of Non-root Node G_{10} of Bayesian Network

Parent node status			Conditional probability of node G_{10}	
X_9	X_{10}	X_{11}	$P(Y)$	$P(N)$
0	0	0	0.05	0.95
0	0	1	0.6	0.4
0	1	0	0.3	0.7
0	1	1	0.75	0.25
1	0	0	0.2	0.8
1	0	1	0.8	0.2
1	1	0	0.7	0.3
1	1	1	0.95	0.05

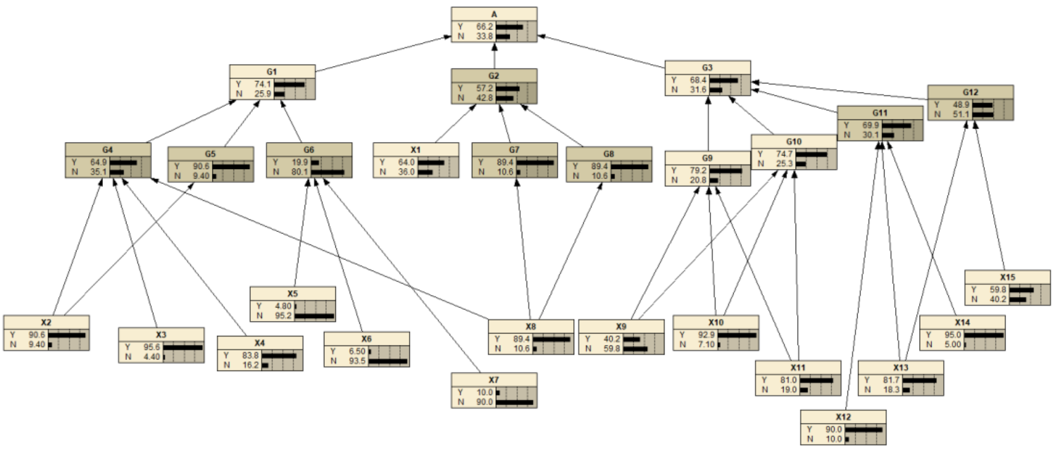
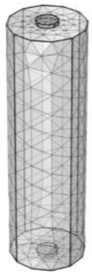


Figure 2: The Overall Success of the System Bayesian Network

V. APPLICATION OF DYNAMIC PROBABILITY NETWORK METHOD IN ENERGY STORAGE MANAGEMENT OF LITHIUM BATTERY PACK

This section first analyzes the heat distribution during battery discharge based on the battery pack model built by COMSOL. In order to establish an analytical model for batteries, the following simplified assumptions are made for batteries: Neglect convective heat transfer due to electrolyte flow. The effect of thermal radiation is not considered. The heat generation inside the battery is uniform and the current density is evenly distributed. The physical properties of the internal components of the battery are the same. By simplifying the modeling of the shape of the 2170 cylindrical lithium battery, a cylinder with two cylindrical plates can be obtained. Among them, the geometric shape of the battery casing can be equivalent to a cylinder with a height of 20mm and a diameter of 21mm, and the geometric shape of the positive and negative electrode ears of the battery can be equivalent to a small cylinder with a height of 1mm and a diameter of 6mm [15]. By using COMSOL simulation software to construct and mesh the 3D geometric model, a single 2170 cylindrical geometric model and mesh model can be obtained, as shown in Figure 3.

(a) Geometric Model



(b) Grid Model

Figure 3: 2170 Single Cylinder Battery Diagram

In geometric modeling, the battery pack is often a combination of a plastic film wrapped around the outer layer, an air medium filling the inner gaps, and the battery pack itself. In the simulation calculation, the battery pack model can be generally divided into three components: air domain, connector, and battery pack. And the material settings for these three parts are set as air medium, metallic aluminum, and active battery material.

In order to reduce the computational complexity of the model and improve the simulation speed, this section calculates the heat generation of a quarter of the individual cells of the lithium-ion battery pack, and obtains the final simulation results of the overall temperature field and heat distribution of the battery pack. In this model, it is believed that the heating situation of each individual battery is completely the same, and the heat generation rate is consistent, while achieving complete discharge. For the battery pack model, discharge it continuously from a state of charge (SOC) of 100% at a rate of 4 for 12 minutes, and discharge it to a state of 20% charge to obtain its temperature field thermal distribution, terminal voltage of each individual battery, and temperature. As can be seen in Figure 4, the analysis of COMSOL-based simulation results of the heat distribution of the battery pack shows that during the discharge process of the battery pack, the heat inside the battery is mainly concentrated in the center region of the battery, while the edge region is relatively low because of the lack of heat dissipation methods, which makes it difficult to emit heat. At the same time, under higher discharge current, the temperature of the battery pack rises more significantly. By simulating and analyzing the heat distribution, the heat distribution characteristics of the battery pack can be derived to provide a reference for the design and control of the battery pack. Meanwhile, optimizing the heat dissipation measures is an effective way to reduce the temperature of the battery pack and improve the service life and safety of the battery pack. The analysis of COMSOL-based battery pack heat distribution simulation results can provide an important theoretical basis for battery design, optimization and fault diagnosis.

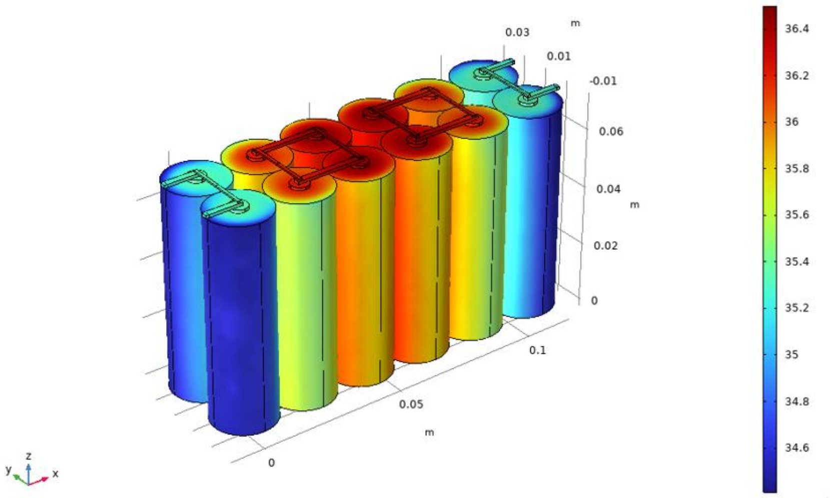


Figure 4: Simulation of the Thermal Distribution of the Battery Pack during 4C Rate Discharge (°C)

Model the model in Figure 2 of the Dynamic probabilistic network method in Simulink and control the energy storage heat of the lithium battery. The constructed model is shown in Figure 5.

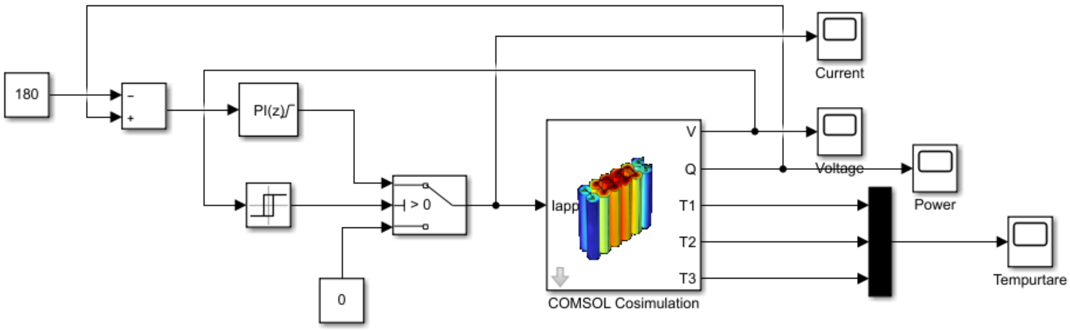


Figure 5: Simulink Simulation Circuit Control Diagram Based on Dynamic Probabilistic Network Method

During the discharge process of a lithium-ion battery pack, the voltage of each single cell will gradually decrease. Since the output current remains constant, the output power of the battery pack will gradually decrease. Since the heat transfer coefficient of the forced convection of the battery pack remains constant, the heat dissipation performance of the battery pack is relatively stable. As a result, the temperature of each single cell also keeps rising, and there is a large temperature difference between each single cell, and an obvious thermal consistency problem occurs.

Under constant power discharge conditions, the temperature of the battery pack shows a gradual increase with the extension of the discharge time, which is due to the fact that the heat is generated from the inside when the battery is discharged, and the rate of heat release is relatively fast, which leads to a rise in the temperature of the battery pack. Meanwhile, the thermal analysis results show that the heat distribution of the battery pack is not uniform during the battery discharge process, with the center temperature being relatively high and the edge temperature being relatively low.

The simulation results based on Dynamic probabilistic network method control management are shown in Fig. 6 and Figure 7. Figure 6 shows the average power values of all batteries, which are controlled within a reasonable range. Figure 7 shows the temperature curves of three different batteries, with convergence reaching a higher standard. Therefore, the accuracy of the Dynamic probabilistic network method in control management was verified.

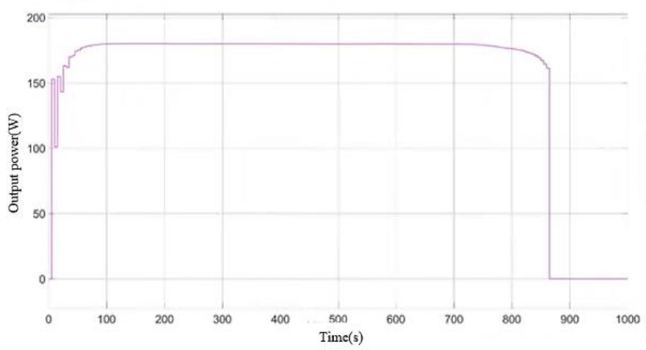


Figure 6: Average Power Values of All Batteries

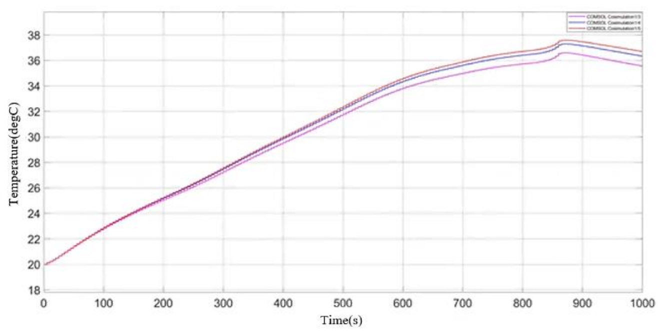


Figure 7: Temperature Curves of Three Different Batteries

VI. CONCLUSIONS

Through the integration of big data technology and the State Grid, this article aims to promote the integration of renewable energy and improve the efficiency and reliability of the power system. In this paper, a system integration principle based on the Bayesian probabilistic approach is proposed to realize the engineering application of the theory of "Comprehensive Integration". Under the Bayesian network framework, the network model for management system integration of State Grid Corporation of China is constructed around mutual information entropy by considering different objectives at different stages of the system integration principle. The network model takes the maximization of the probability of achieving the holistic goal as the basic goal, expresses the tension between macroscopic holism and microscopic complexity at each stage in terms of mutual information, constructs different computational models of the goal probability according to the tasks at different stages of the system integration, and proposes the corresponding algorithms. The algorithm proposed in this paper is implemented in Netica, a commercial statistical analysis software, and applied to the management of State Grid Corporation of China for intelligent decision-making on small sample data, thus proposing a feasible framework for the future application of the principle of system integration, and providing an important theoretical basis for the engineering practice of the theory of "Comprehensive Integration". Finally, with the help of COMSOL commercial simulation software, the accuracy of the Dynamic probabilistic network method was verified in a lithium battery pack energy storage management example.

REFERENCES

[1] M. S. Mahmud, J. Z. Huang, R. Ruby, A. Ngueilbaye and K. Wu, "Approximate Clustering Ensemble Method for Big Data," IEEE Transactions on Big Data, vol. 9, no. 4, pp. 1142-1155, 2023.

[2] H. Ahmed and M. A. Ismail, "A Structured Approach Towards Big Data Identification," IEEE Transactions on Big Data, vol. 9, no. 1, pp. 147-159, 2023.

[3] S. Zhiyong, W. Caixia, C. Ning, Y. Xiaoning, and W. Si, "Policy Requirements and Economic Affordability of Energy Storage for New Energy," in 2022 6th International Conference on Power and Energy Engineering (ICPEE), Shanghai, China: IEEE, Nov. 2022, pp. 330–333.

[4] Ma Yingjie, Luo Yiqing, Zhang Shuo, Yuan Xigang, "Simultaneous optimization of complex distillation systems and heat integration using pseudo-transient continuation models," Computers & chemical engineering, vol.108, pp.337-348, 2018.

[5] K. Gharehbaghi and M. Myers, "Intelligent System Intricacies: Safety, Security and Risk Management Apprehensions of ITS," International Conference on Industrial Technology and Management, pp. 37–40, 2019.

[6] Lakemond, Nicolette Holmberg, Gunnar, Pettersson, Anders, "Digital Transformation in Complex Systems," IEEE transactions on engineering management, pp. 1-13, 2021.

[7] S. Nuccio and C. Spataro, "Uncertainty Management in the measurements performed by means of virtual instruments," IEEE Iinternational Workshop on Advanced Methods for Uncertainty Estimation in Measurement, pp. 40–45, 2008.

[8] Y. Hata, N. Hayashi, Y. Makino, A. Takada, and K. Yamagoe, "Alarm Correlation Method Using Bayesian Network in Telecommunications Networks," in 2022 23rd Asia-Pacific Network Operations and Management Symposium (APNOMS), Takamatsu, Japan: IEEE, Sep. 2022, pp. 1–4.

[9] H. Yuan, J. Yang, J. Yang, R. Zhu, and W. Xue, "Study on the Influencing Factors of Energy Consumption of Nearly Zero Energy Public Buildings in Cold and Arid Regions of Northwest China," in 2022 3rd International Conference on Advanced Electrical and Energy Systems (AEES), Lanzhou, China: IEEE, Sep. 2022, pp. 487–491.

[10] H. Khoshkbari, V. Pourahmadi, and H. Sheikhzadeh, "Power Allocation in Cellular Network Without Global CSI: Bayesian Reinforcement Learning Approach," in 2020 28th Iranian Conference on Electrical Engineering (ICEE), Tabriz, Iran: IEEE, Aug. 2020, pp. 1–6.

[11] V. Senthilkumar and V. Jayalakshmi, "Radial Basis Function Networks for Image Restoration with Stochastic Normalizations as Bayesian Learning in Deep Conventional Neural Network," in 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India: IEEE, Mar. 2022, pp. 1298–1302.

[12] J. Zheng and Y. Liu, "Research on Online Learning Achievement Prediction Based on Dynamic Bayesian Network," in 2023 IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqing, China: IEEE, Feb. 2023, pp. 862–866.

[13] W. R. Scott, R. E. Levitt, and R. J. Orr, Eds., "Global projects: institutional and political challenges." Cambridge University Press, 2011.

[14] R. Pakdel, M. Yavarinasab, M. R. Zibad, and M. R. Almohaddesn, "Design and Implementation of Lithium Battery Management System for Electric Vehicles," in 2022 9th Iranian Conference on Renewable Energy & Distributed Generation (ICREDG), Mashhad, Iran, Islamic Republic of: IEEE, Feb. 2022, pp. 1–6.

[15] A. S and L. N, "Effect of Input Data and Parameters of Genetic Algorithm on Lithium-Ion Battery Modeling," in 2023 IEEE World Conference on Applied Intelligence and Computing (AIC), Sonbhadra, India: IEEE, Jul. 2023, pp. 524–528.