Integrated energy system planning research based on big data load prediction method

Yongli Wang¹, Hekun Shen*¹,a, Jialin Yang², Nan Wang², Yuze Ma¹, Pengxiang Zhao², Zhen Li², Xichao Zhou², Suhang Yao¹

¹ North China Electric Power University, Beijing, 102206, China
² State Grid Integrated Energy Service Group Co., Ltd., Beijing, 100032, China

Abstract. The planning of integrated energy system is a very complex multi-target, multi-constraint, nonlinear, random uncertainty mixed integrated combination optimization problem, its planning and design process should not only consider the interdependence between the system capacity, energy conversion, energy storage, energy use and other links, but also consider the interaction and integration of cold, hot, electricity and other multi-energy flows, which is essentially a non-deterministic polynomial difficult problem. China's energy continues to develop rapidly, all kinds of sensors and intelligent equipment data is increasing, the data obtained in the equipment and all kinds of sensors collected energy load prediction related factors such as temperature, weather, wind speed and other data volume increased dramatically, the data dimension is also increasing, the scale of data has also increased from GB to TB or even higher, based on the traditional prediction methods and intelligent prediction methods, has been far below the load forecast desired to achieve accuracy and speed requirements, Therefore, the use of big data technology to predict energy demand is an important future direction.

1 Introduction

With the development of big data technology, the high-speed processing of load curve data has been successfully realized, and energy consumption can be predicted in short-term time. Through the improvement of traditional energy big data technology in the processing of large amounts of data structure, the complex and parallel processing power of energy big data technology within the specified delay is successfully realized. At the same time, some applications even have the ability to reflect real-time processing in real time by using different scales [1]. The current energy big data technology has been studied in depth in the aspects of power grid frame optimization, power scheduling and load prediction. Using energy big data technology to dig deep into the large amount of load data collected, so as to make accurate load prediction, can provide a strong technical support for the planning and operation of integrated energy system.

A large number of researchers in the framework of big data technology to carry out load prediction, taking into account the integrated energy system in the big data environment of power load factors of multi-source, the temperature and other factors as factors, the use of two-layer multi-core learning algorithm, the establishment of support vector load prediction algorithm, is a comparative breakthrough [2]. A large number of researchers carry out load prediction under the framework of big data technology. Considering the multi-source of power load factors in the integrated energy system, taking temperature and other factors as factors, using two-layer multi-core learning algorithm to establish support vector load prediction algorithm is a breakthrough. Using graph clustering algorithm to segment different characteristics of each industry users, on this basis, the big data analysis and processing technology combined with support vector machine algorithm applied to load prediction, designed a set of load prediction architecture [3], and made the implementation and comparative study of the algorithm, showing that the prediction results and the actual situation of high consistency, and the operating speed advantage is very obvious, has a strong practicality [4].

There are many commonly used load prediction methods, such as regression analysis, exponential smoothing, time series, Kalman filtering and so on [5]. However, due to the large change of load curve and many influencing factors, the traditional method has obvious limitations in predicting speed and accuracy, and the actual application effect is relatively poor [6]. With the development of artificial intelligence in recent years, it has been applied in various fields [7]. Among them, expert system, artificial neural network, fuzzy prediction method, wavelet analysis method and other intelligent methods are widely used, in load prediction, the use of intelligent algorithm has a good advantage [8]. The accuracy of prediction is improved by applying fuzzy theory to neural networks, and the complexity of the original prediction model is reduced by combining empirical modal decomposition with deep learning theory, and the speed of algorithm load prediction is accelerated [9]. Applying the mixing wheel theory to load prediction, using the traversal
and uncertainty of chaos, constructing the prediction model, which better reflects the load change of the user, and the literature takes into account the factors affecting the load change, which is more targeted [10].

2 Short-term power load prediction model based on attention mechanism

The accuracy of the forecast results of the integrated energy system load is directly related to whether the planning and operation are reasonable and effective, so it is an important prerequisite for the comprehensive energy system planning. The traditional energy system load prediction method is limited to the independent prediction of various forms of energy, although it has formed a mature theoretical method and has been proved by practice, but because of neglecting the coupling and mutual influence between plpoints, it can not be applied to the load prediction of the future integrated energy system. The load forecasting difficulties of integrated energy systems are mainly reflected in the following two aspects:(1) There are many influencing factors. Since integrated energy systems cover all forms of energy such as electricity, gas, heat and cold, they are affected by a number of factors, such as climate, socio-economics, regional layout structures, architectural design characteristics, etc., resulting in difficulties in modeling (2) difficulty in solving predictive models. Multi-dimensional and multi-time scales make it more difficult to solve the load prediction model of integrated energy system. In recent years, scholars at home and abroad have made extensive and in-depth exploration of the load prediction method of integrated energy system, and achieved many effective results.

The model is mainly divided into data pre-processing, feature extraction, prediction output 3 parts. The feature extraction part is mainly composed of CNNs, LSTM and Attention module 3, CNNs part through a series of convolution and pooling operations, the characteristic extraction of data, and then through the full connection layer to convert the feature data into one-dimensional data, input into LSTM; The LSTM and Attention modules combine to model input features in the time dimension. The Predictive Output section is responsible for decoding the features of the feature extraction module output and predicting the output for the next moment. Data preprocessing: As the first part of the entire algorithm, power high-dimensional data needs to be entered into the model in chronological order. For example, a piece of high-dimensional power data with a length of T is processed and converted to $X = [x', x'\cdots, x'^t, x']$. Feature Extraction Module: The CNNs section of the feature extraction layer requires feature extraction of the high-dimensional time series entered. The entire CNNs section consists of two one-dimensional convolution layers, two pooled layers, and one fully connected layer. In order to inhibit the occurrence of "neuronal degeneration" in the course of ReLU training, Leaky-ReLU is used as an activation function to replace ReLU. The role of the pooling layer is to reduce peacekeeping and increase the feeling of wild, in order to prevent information confusion, so the model chooses the maximum pooling as the method of downsampling. Dropout technology is also used to improve the generalization of the model. Where the output of CNNs can be used, $H_c = [h_1, h_2, h_3, \cdots, h_i, h_j]$ where i represents the CNNs output dimension.

LSTM can model and analyze the time dimension for features extracted by CNNs, and LSTM's output at the t-moment is shown in formula (1).

$$h_t = LSTM(H_{c_{t-1}}, H_{c_t}), t \in [1,1]$$ (1)

The entry of the Attention module is the result of the LSTM structure, and the Attention module automates the learning and judging the importance of different characteristics and assigns weights based on the importance relationship. The Attention module can be embedded directly into the deep learning model for end-to-end training. The weight coefficient in Attention is shown in formula (2).

$$a_i = \frac{\exp(e_i)}{\sum_{j} \exp(e_j)}$$ (4)

In, $e_i$ represents that t-moments have characteristics output by the LSTM, $\mu$ and $w$ are constant terms.

Predictive output: This section requires decoding the features output by the Attention module. This is shown in the public style (5). The $y$ represents the model's prediction of t-moment results, and $w$ and $b$ represent weights and biases, respectively.

$$y_t = Sigmoid(w's + b)$$ (5)

3 Integrated energy system planning

3.1 RIES can flow modeling

The initial energy P (electricity, gas, heat, light, wind, etc.) supplied to the external region by the total energy flow of the region is converted, stored, transmitted, and managed on the demand side, and finally meets the end demand L within the region. The procedure can be expressed as:

$$L = f(P, C, N, D)$$ (6)

C, S, N, and D represent the change process of energy conversion, storage, transmission, and demand side management. Where D is related to the specific control strategy of demand-side management, it can be regarded as a modification of the end requirements, expressed as:

$$L = D(L)$$ (7)
N contains electrical, gas, heat (cold) and other systems for the transmission of the network transmission process, its characteristics are different, discussed below.

1) Power system: Power system current calculation takes node voltage amplitude and phase as the state variables, and establishes the active and reactive balance equation:

\[
P_{m} - P_{n} = U_{1} \sum_{i} U_{i} \left( G_{i} \cos \theta_{i} + B_{i} \sin \theta_{i} \right)
\]

\[
Q_{m} - Q_{n} = U_{1} \sum_{i} U_{i} \left( G_{i} \sin \theta_{i} + B_{i} \cos \theta_{i} \right)
\]

(8)

Formula (8) equations have the following characteristics:(1) smooth continuous;(2) the desired solution is near; \(U_{1} = 1, \theta_{i} = 0\) (3) the structure is sparse.

2) Natural gas system: natural gas system can flow to node air pressure as the state variable, establish the node gas balance equation:

\[
Q = \sum f_{i} + \sum f_{m} - \sum F_{m} + F_{e} = 0
\]

(9)

In the formula: \(Q_{m}\) is the net injection of natural gas at node i, \(f_{m}\) and \(f_{m}\) are the injection of flow from upstream node n and the output flow to downstream node m, respectively, \(F_{m}\) is the gas consumption of the compressor, and the correlation coefficient \(F_{e}\) of one when the compressor extracts gas from node i, otherwise zero. Branch flow is calculated as follows

\[
f_{m} = \begin{cases} 
W_{m} \sqrt{p_{i}^2 - p_{m}^2} & p_{i} \geq p_{m} \\
-W_{m} \sqrt{p_{m}^2 - p_{i}^2} & p_{i} < p_{m}
\end{cases}
\]

(10)

In the formula: \(f_{m}\) is the flow of the branch \(im\), \(W_{m}\) is the damping coefficient of the branch \(im\), \(p_{i}\) and \(p_{m}\) is the pressure of the node.

The compressor gas consumption is:

\[
F_{e} = \begin{cases} 
\frac{\alpha_{f} m_{i}}{n_{c}} \left[ k_{1} \left( \frac{p_{m}}{p_{i}} \right)^{\alpha} - k_{2} \right] & p_{m} \geq p_{i} \\
\frac{\alpha_{f} m_{i}}{n_{c}} \left[ k_{1} \left( \frac{p_{i}}{p_{m}} \right)^{\beta} - k_{2} \right] & p_{m} < p_{i}
\end{cases}
\]

(11)

For compressor-free gas distribution networks, formula (9) can also be simplified to:

\[
p_{i} - p_{m} = s_{w} m_{i}^{\beta-1} f_{m}
\]

(12)

In the formula: \(n_{c}\) is compressor efficiency; \(\alpha, \beta, k_{1}, k_{2}\) is relevant constant coefficient, \(s_{w}\) is pipe resistance coefficient, and \(\beta\) is normal coefficient related to airflow and pipe characteristics.

3) Heating (cold) system: Heating (cold) system analysis includes hydraulic analysis and thermal analysis of two parts. Its hydraulic analysis selects branch flow M as the state variable, and establishes a flow balance equation and a loop pressure drop equation for a water supply network with N nodes and b branches: \(N - 1<br-N+1

\[
A_{m} M = Q
\]

\[
B_{e} \left( S^{M} M + Z - H \right) = 0
\]

(13)

In the formula: \(A_{m}\) is the node-branch correlation matrix, \(Q\) is the net injection flow at each node, \(B_{e}\) is the basic loop branch association matrix, \(S^{M}\) for each pipeline resistance coefficient diagonal array, \(Z\) for the pipe elevation difference vector, \(H\) for the branch pump head vector.

Thermal analysis establishes a temperature balance equation for water supply and return networks:

\[
A_{T} T_{s} = 0
\]

\[
A_{T} T_{h} + B_{T} T_{r} = 0
\]

(14)

In the system: \(T_{s}\), \(T_{h}\) and \(T_{r}\) are the water supply network, the return water network node and the user return water temperature vector, \(A_{T}\) and \(B_{T}\) are the matrices related to flow, temperature and loss respectively.

4) Solution method: Although the above energy systems have different characteristics, they all meet Kirchhoff’s law and energy conservation law, and can be solved by The Newton-Raphson method or other iterative algorithms. The solution of hybrid trend can be solved by constructing a comprehensive equation of hybrid network, using the unified iteration of the increasingly wide and comparable matrix, or by distributing the mixed trend information to different energy sources separately by the energy equivalent of the energy coupling part.

3.2 Integrated energy system planning objectives

Economy

\[
F_{i} = \min C_{e} = C_{w} + C_{w-c} + C_{w-t}
\]

(15)

in the form, \(C_{e}\) is the operating cost of the system, yuan, \(C_{w}\) is the annualized investment cost, the yuan, \(C_{w-c}\) is the purchase cost, the yuan, and \(C_{w-t}\) is the equipment operation and maintenance cost, the yuan.

1) Power grid purchase costs

\[
C_{w} = \sum_{i} (\alpha P_{t}) T
\]

(16)

In the formula, \(\alpha_{t}\) is the purchase price, yuan, \(P_{t}\) is the time period t, kW, T is the scheduling cycle.

2) Equipment operating maintenance costs

\[
C_{w-t} = \sum_{i} C_{e} P_{t} T
\]

(17)

In the formula, \(C_{e}\) is the distributed power generation operating maintenance costs, yuan/kW, \(P_{t}\) is the distributed power generation output power, kW.

4 Case validation

This article uses load data from July 1, 2019 to August 20, 2019 (30 days per month) in a region of Sichuan, China,
as training data, sampled once a day for a total of 50 load data points, and 40 points for test data from August 21, 2019 to September 30, 2019. The traditional planning scheme and the short-term power load prediction model based on attention mechanism are trained and forecasted. In order to verify the planning scheme effect of the short-term power load prediction model based on attention mechanism, the two are compared, and the simulation effect diagram is shown in Figure 1. As can be seen from Figure 1, the planning scheme of the short-term power load prediction model based on attention mechanism has significantly improved compared with the traditional planning scheme, and the prediction curve fits better, indicating that the optimization effect is significant and the prediction accuracy is higher.

According to the load forecast data for the park planning, select a typical day for equipment power analysis, typical under-day equipment power as shown in the figure.

At the same time, comparing the impact of different load prediction methods on the planning results, Scenario 1 and Scenario 2, although the device elements are the same, but because of the size of the load affects the operating costs of the equipment, thus affecting the size of carbon emissions. Scenario 2 is a short-term power load prediction based on the attention mechanism of this article, and Scenario 1 is a traditional power load prediction. From Table 1 above, you can get the value of the operating cost and carbon emissions of Scenario 1 and Scenario 2, which is 8.14% lower than Scenario 1, and Scenario 2, which is 4.81% lower than Scenario 1.

5 Conclusions

How to make accurate prediction of energy load by big data technology is a forward-looking subject, and the accurate prediction results are of great significance to the formulation of reasonable planning and scheduling plan, improving equipment utilization rate and ensuring the safe and stable operation of the system. RIES covers a wide range of energy forms, such as electricity, gas, heat and cold, and its uncertainty complexity far exceeds that of any single system, and may exist in all aspects of energy production, transmission, conversion, consumption, etc. Therefore, in the planning research of integrated energy system, in addition to new energy contribution, load, equipment failure and other factors, but also need to consider such as meteorological conditions on the heat and cold load, external temperature on the heat pipe transmission efficiency, building layout on the cooling heating demand and so on. This paper describes the classical methods and intelligent methods, and points out the advantages and disadvantages of different prediction methods, introduces big data technology to complete load prediction, can effectively increase the accuracy and speed of prediction, is conducive to the realization of good integrated energy system planning, and contributes to the improvement of stability and economy.

Acknowledgment

The authors would like to thank the support of the project “the State Grid Corporation Science and Technology Project (SGSDDK00PDJS2000383)”.

References

1. Barukčić I., Wen Zhou, Meng Liang, Yang Zhengfu, Liu Zhibin, Liu Yajing. Research on complex integrated energy planning and operation optimization[J]. MATEC Web of Conferences, 2021, 336.
2. Zhang Xiaohui, Liu Xiaoyan, Zhong Jiaqing, Gao Wenbo. Electricity–gas-integrated energy planning based on reward and penalty ladder-type carbon trading cost[J]. IET Generation, Transmission & Distribution, 2019, 13(23).
3. YU Sheng, ZOU Hongbo, YU Fan, et al. Application of fuzzy neural network in power system short-term
4. YUE Youjun, LIU Yinghan, ZHAO Hui, et al. Short term electric load forecasting based on CEEMDAN-SE and DBN [J/OL]. Electrical measurement and instrumentation. [2019-11-24].

5. DONG Zihan. Regional power network load forecasting based on chaotic time series [J]. Power System and Clean Energy, 2019, 35(5): 38-41.

6. ZHOU Xinghua, GENG Juncheng, DU Songhuai, et al. A summer short-term load forecasting method for temperature sensitive users considering demand response [J]. Power System and Clean Energy, 2019, 35(4): 16-22.

7. Tirupati Uttamrao Solanke, Vigna K. Ramachandaramurthy, Jia Ying Yong, Jagadeesh Pasupuleti, Padmanathan Kasinathan, Arul Rajagopalan. A review of strategic charging-discharging control of grid-connected electric vehicles [J]. Journal of Energy Storage, 2020, 28.

8. CHEN Linghui. Research on load balancing and decentralized dispatching method of renewable energy power system [J]. Power System and Clean Energy, 2019, 35(12): 60-66.

9. HU Wei, HU Yawei, YANG Feng. A wind power prediction neural network model based on improved flower pollination algorithm [J]. Journal of Systems Management, 2019, 28(5): 934-940.

10. TAO Zhiyong, CUI Xinxin. A hybrid improved flower pollination algorithm [J]. Transducer and Microsystems, 2019, 38(10): 139-142, 145.