Chapter 7
Research and Implementation of Building a Digital Twin Model for Electric Grid Based on Deep Learning

Junfeng Qiao, Zhimin He, Hai Yu, LianTeng Shen, and Xiaodong Du

Abstract In the production of the power industry, digital twin technology is mainly used for the dispatch and distribution of electricity. Through real-time monitoring and analysis of key indicators in the dispatch center of the power company, the operation level and power supply service quality of the power company are improved. At present, the digital twin applications in the power industry are mainly used for indicator monitoring and process analysis, with the main goal of monitoring being existing collected data, lacking real-time analysis of higher-order data. With the continuous expansion of electricity consumption scale, the electricity demand of electricity customers is also becoming increasingly rich, which correspondingly drives the data generated in various business processes of power enterprises to develop towards a trend of high-dimensional and high-order. Traditional data processing and mining technologies are no longer able to quickly characterize and describe such high-order and high-dimensional data. Only by utilizing more efficient data analysis methods can we meet the current demand for power digital twin services. This article proposes a method for constructing a power grid digital twin model based on deep learning methods, which conducts deep learning on high-dimensional and high-order data in the power grid, identifies its features, and constructs a data category classifier to promote the construction of a power grid digital twin model that is more needed for load power business analysis, and improve the usability and applicability of the power grid digital twin model.

J. Qiao (✉) · Z. He · H. Yu
State Grid Laboratory of Power Cyber-Security Protection & Monitoring Technology, State Grid Smart Grid Research Institute Co., LTD, Nanjing, Jiangsu 210003, China
e-mail: 1318558905@qq.com

L. Shen
China Electric Power Research Institute Co.Ltd, Beijing 100192, China

X. Du
State Grid Hebei Electric Power Co., LTD, Shijiazhuang, Hebei 050021, China
e-mail: 806293765@qq.com

R. Kountchev et al. (eds.), *Proceedings of International Conference on Artificial Intelligence and Communication Technologies (ICAICT 2023)*, Smart Innovation, Systems and Technologies 369, https://doi.org/10.1007/978-981-99-6956-2_7

7.1 Introduction

The variety and quantity of equipment in power production are abundant, and using digital twin technology to model and simulate the business process of power will undoubtedly face huge performance bottlenecks [1]. In order to overcome this problem, it is necessary to conduct in-depth research and analysis on existing digital twin models and related technologies, and identify the key factors that prevent digital twin models from quickly and accurately describing power business scenarios [2]. Intelligent sensors are the grippers of data collection, and digital twin models accurately and real-time describe the state of the power grid by obtaining the most real-time data from intelligent sensors. Only when the real-time data in the digital twin model can be consistent with the current situation of power operation in reality, can the power business analysis based on the digital twin model have practical significance. Otherwise, the lagged data description model will lose most of its analytical significance [3]. Taking the distribution network planning business as an example, the distribution network is directly associated with users to allocate and supply electricity, which is the most direct window for interaction with users. Electricity customers use the distribution network to perceive the power supply services of the grid [4].

Building a scientific and effective digital twin model for distribution networks and providing guidance for distribution network planning is the most urgent need for distribution network planning work. The expansion and design of future distribution network development must rely on real-time distribution network related data. However, with the large-scale integration of new energy, the state of the distribution network is rapidly changing, with frequencies reaching the minute level. The traditional evaluation method uses the entire network data of time section as the basis for planning the distribution network, and in the current development situation of the distribution network, it is no longer possible to implement truly effective evaluation [5]. Therefore, more efficient data collection strategies and methods are needed, first ensuring the accuracy and real-time nature of the basic data. Based on this, a planning analysis model is constructed to achieve the planning of the distribution network [6]. Therefore, the key to the success of the construction of the digital twin model proposed in this article lies in the rapid acquisition of transient, steady-state data, and other related real-time state variables of the power grid.

7.2 Related Work

The key to digital twinning of power equipment currently lies in preparing for real-time and complete data collection, abstracting power business from scenarios into digital description models in virtual space. The most critical entity in the power industry is power equipment, which is the entire embodiment of the power grid entity. Whether it is in the power generation or supply process, the carrier of all business is power equipment. Therefore, in order to abstract and simulate the process of electricity, it is first necessary to develop a scientific and efficient data collection plan for device description [7]. At present, the main collection devices for power equipment are wired collectors, including energy meters and Android based vertical and mobile handheld terminals [8]. The sensors of these collection devices are very sensitive, and their collection frequency and speed are first-class, reaching the second level, which is sufficient for business analysis based on digital twin models. However, there are still some data, such as device temperature, device images, videos, and other descriptive data that cannot achieve a minute level collection frequency [9]. And the speed of transmitting data back to data centers such as data centers after data collection is also constrained by bandwidth transmission bottlenecks, so corresponding new methods must be developed to capture these types of data. Since it is not possible to directly and quickly obtain the required relevant data, other processing methods can only be used to process or simulate the missing data, or even create fake data to replace the required real data [10]. False data does not mean that the data cannot be used. If simulated false data can reflect the real power grid status or parameters, then these data can also meet the needs of business analysis. The data simulation method proposed in this article requires the use of the currently hottest and most popular artificial intelligence method, namely the deep learning algorithm. Through deep learning algorithm, the features of real data that cannot be obtained in reality are solved and simulated, and then the necessary basic data is generated based on these features.

Deep learning algorithms are currently widely used in the field of image processing, such as the widely used face recognition technology, which widely uses deep learning algorithms. Human beings are almost unable to understand and imagine data that exceeds three dimensions, and their analysis of data that exceeds three dimensions is also powerless. Deep learning algorithms can replace humans in abstracting data from multiple dimensions, completing analysis of high-dimensional data, and outputting results according to preset analysis logic [11]. This article applies deep learning algorithms to learn historical data and obtain the desired data features of the future state. Then, based on the sample information in the historical data, the two are combined to obtain a complete target data-set of the future state. The United States initially applied deep learning algorithms to the military and aerospace fields, by collecting a large amount of surrounding environmental data of spacecraft and combining it with internal state data to predict the future attitude and state of spacecraft, providing decision support for ground command systems.

7.3 Research on Power Grid Digital Twin Model Based on Deep Learning

This article uses multiple types of databases to store the operating status and equipment parameters of the power grid, as well as other related archival data. This is because the data generation carriers and description objects in various fields of power operation are different. Therefore, the data types and structures are rich, with over 4 types of data structures. Therefore, the data platform of this article also supports the access of various types of power data.

7.3.1 Data Description Model for Power Equipment Collection

The data of power equipment includes static data and floating state data. Static data refers to fixed and unchanging data such as equipment type, equipment manufacturer, equipment address, equipment capacity, and voltage level. Floating dynamic data includes data such as equipment operating voltage, passing current, equipment resistance value, and equipment opening and closing state. These data change with changes in power operation status, and the speed of data change is even lower than 1 s.

There are also some power data that can be collected at low frequencies or even cannot be collected, such as real-time images and video data of devices. In power voltage transformation and transformer stations, there are video monitoring devices that monitor the core important equipment in the station. For some devices that are not of high voltage level or not important to your mother outdoors, there is no corresponding video and image acquisition equipment for data collection, which is due to economic and other reasons. However, with the rapid development of the power grid, these data have become increasingly necessary. Only with real-time image and video data can the equipment in the power grid be fully and accurately described, and a scientific and effective digital twin model can be constructed to achieve objective and effective evaluation and analysis of the power grid.

As is shown in Fig. 7.1, the data collection in the power system is mainly divided into static collection devices and dynamic collection devices, and meteorological, temperature, and video data must be dynamically collected. Even if the corresponding image and video data can be collected, the transmission and storage of these data still have significant problems under the current power grid state. This is because image and video data usually occupy a large storage space and require the purchase of expensive and large amounts of storage devices, which will undoubtedly hinder the cost-effectiveness of the analysis results.

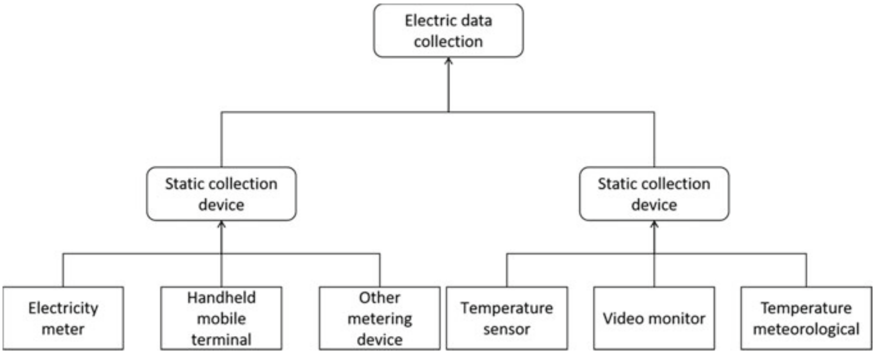

Fig. 7.1 Schematic diagram of electric power data collection

7.3.2 Research on the Description Method of Power Equipment Simulation Data

As mentioned in the previous section, for device images, videos, and other state parameters that cannot be obtained but must be obtained, they must be simulated to obtain the required data. The application of data simulation methods must conform to the true and objective image of the described object in order to obtain effective data results. Otherwise, data simulation will fail, and the results of data simulation cannot be truly applied, so these tasks will become useless.

As is shown in Fig. 7.2, this article utilizes deep learning as an intelligent algorithm for feature extraction and sample training of target data requiring solutions. Building a data feature library is a difficult task, and the method adopted in this article is machine learning, which requires analysis in two different situations. If the collection frequency is relatively low and you want to obtain higher frequency data, you can directly use the data simulation algorithm to extract the data from the two collection sections, and smooth it based on historical data. You can take the average value or use similarity algorithms to determine the missing high-frequency data based on the trend of data changes. This method treats the data obtained from the required solution as missing data, and uses rich data completion algorithms to complete the missing data, resulting in the uncollected data in the high-frequency data. For data that is difficult to collect, it is necessary to perform directional extraction of its features. Firstly, based on other available data that can be collected, manually annotate and fill in the data that cannot be collected. Then, using artificial intelligence machine learning algorithms, a connection feature library is constructed between the collectable data and the annotated result data, and continuous learning and training are conducted to ultimately obtain a data classifier. Continuously iterate new sample data and improve the classifier until its classification results are relatively stable. Finally, by inputting other data that cannot be collected at the same time as the required solution, the solver outputs the target data to obtain the data object to be simulated.

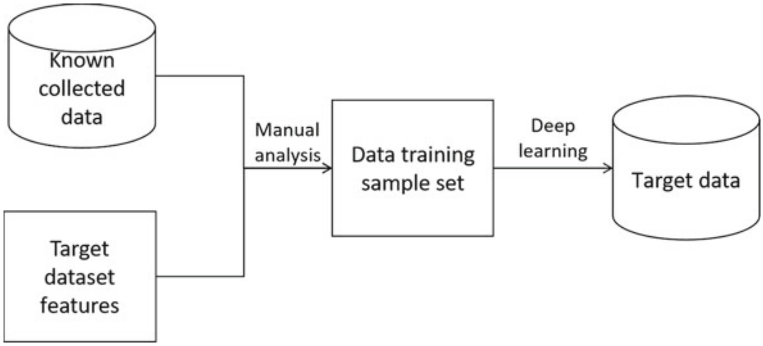

Fig. 7.2 Schematic diagram of electric power data collection

7.4 Implementation and Application of Power Grid Digital Twin Model Based on Deep Learning

7.4.1 Implementation of Power Grid Digital Twin Model Based on Deep Learning

As is shown in Fig. 7.3, after obtaining the full amount of power entity description data, it is simulated and solved to obtain a standardized power description model, which is then transformed into a power digital twin model. The first step is to establish a standardized description information model for power equipment and power operation data, and use the graph description model to construct a unified description of power equipment. For a data description model that was originally relational, convert it into a corresponding graph description model. If the relationship description model cannot be one-to-one mapped to the topology structure, manual intervention is needed to transform its relationship structure and find its corresponding and mapping method. Then generate a unified graph structure conversion method, forming an automated tool that can achieve automated conversion in future data conversion. At the same time, it is also necessary to consider the corresponding methods of different conversion methods. For different conversion methods, corresponding identification should be provided to identify which data the method can apply to. When encountering data with load conversion conditions, the conversion should be carried out directly. When encountering data that does not meet the conditions, it should be pushed to the conversion method with load conversion conditions.

The second step is to construct a digital twin model of the power grid, and through deep learning, integrate the obtained analog data into the entire digital twin model. At the same time, it is necessary to sort out the correlation between simulated data and objective data to obtain its data association model. For a well constructed digital twin model, it is not static and needs to be continuously adjusted according to business needs. For different business data call requests, it is necessary to parse the business

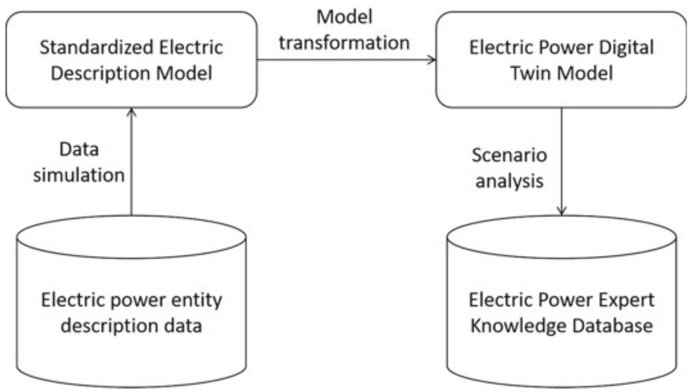

Fig. 7.3 Building digital twin model for electric power scenarios

request analysis work order to determine the required data type, data structure, and data volume. Then update the existing digital twin model, identify the missing parts in the existing digital twin model, and extract them from the data source. Meanwhile, if excess data is not used for a long time or its usage frequency is lower than the set threshold, it will be removed from the digital twin model.

Finally, in the completed digital twin model, new business call requests are continuously iterated, and the existing digital twin model is constantly updated. For newly added new businesses, the corresponding digital twin model needs to be rebuilt. Repeat the above process to improve the usability and scientific of the digital twin model.

7.4.2 Application of Power Grid Digital Twin Model Based on Deep Learning

There are many real-time analysis requirements in the power business, leaving a lot of room for digital twin models. Firstly, in the field of marketing, power customer service requires a large amount of real-time data to grasp and analyze customers' electricity needs and complaint lists. Therefore, power customer service requires real-time digital twin models to provide efficient and fast analysis services.

In the maintenance business of power equipment, business analysis services provided by digital twin models are also needed. Power equipment maintenance requires real-time mastery and analysis of the operating status and environmental parameters of all equipment in the power grid. Through indicator comparison and mining analysis, hidden dangers of the equipment in operation are identified, and early warning and judgment of the equipment's faults are carried out based on certain business guidance opinions.

As is shown in Fig. 7.4, the digital twin model is applied in power customer service, power equipment maintenance, and power grid operation control. The digital twin model based on deep learning studied in this article has been applied in power marketing and power equipment operation and maintenance business, and has been connected to the historical data of over 50,000 low-voltage equipment in the distribution network over the past year. It provides basic model support for equipment maintenance, fault warning, and improvement of power customer service efficiency, and provides decision support opinions. Through demonstration applications, the digital twin model based on deep learning proposed in this article can improve the efficiency of fault warning by over 2%, achieve a fault prediction accuracy of 99% in urban distribution networks, and over 85% in rural power grids. At the customer service center, the digital twin model studied in this article can provide real-time work order data for customer service personnel, provide second level data extraction services for customer service personnel to grasp the electricity needs of electricity customers, improve customer service efficiency, and improve customer service satisfaction.

As shown in Fig. 7.5, in electric power equipment fault prediction, this model obtains data corresponding to all factors affecting power equipment fault, including transformer oil and gas data, voltage, current and power data carried by the equipment, and electrical attribute data of the equipment itself. These data are input into the digital twin model as the dependent data set of power equipment fault prediction. Then, based on the long-term and short-term memory neural network prediction algorithm, the correlation characteristics between the data corresponding to these influencing factors are calculated, and then mapped to equipment faults in the time series dimension. Construct a refined fault diagnosis method for power equipment using a hybrid driven approach of data mechanism models and data models. The application of this model provides an algorithmic basis for power equipment fault prediction,

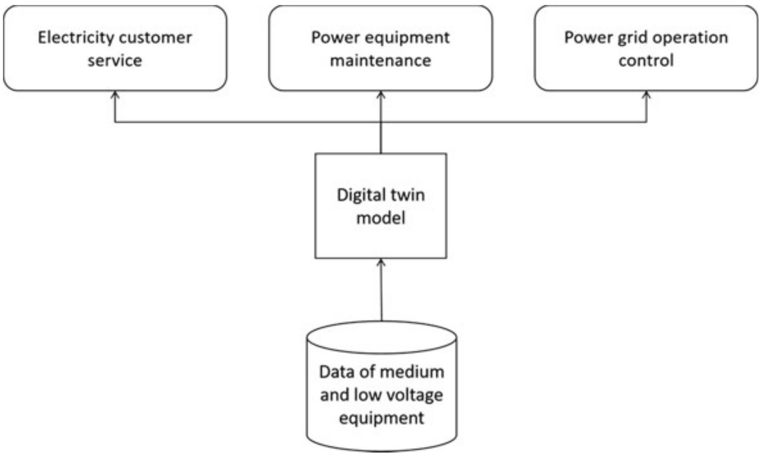

Fig. 7.4 Application of digital twin model in electric power scenarios

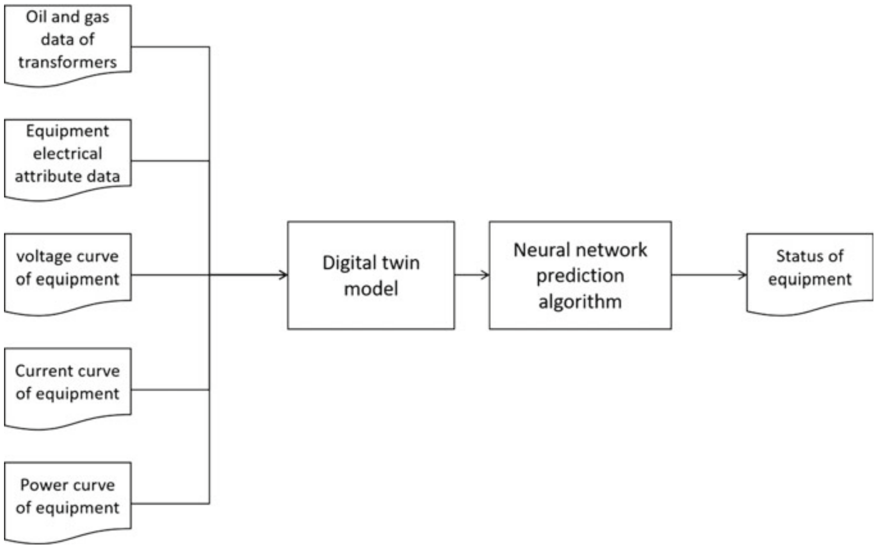

Fig. 7.5 Application of digital twin model in prediction of electric equipment

and has higher prediction accuracy compared to traditional single indicator prediction methods such as time series.

7.5 Conclusion

In summary, the digital twin model construction method proposed in this article can accelerate the performance of data analysis, simplify the process of data analysis structurally, and meet the needs of digital twin models for power equipment maintenance and power customer service businesses functionally. The key to the success of this method lies in the application of deep learning models, which incorporate more power data, fill in some low frequency data, improve data frequency, and provide more accurate data for data analysis. For some data that cannot be obtained, artificial intelligence deep learning algorithms are used to simulate and learn features based on data features, and combined with manual annotation results, the data feature classifier is constructed. By inputting the relevant data of the new data object to be simulated, the data set of the mock object is obtained, which fills the data gap. Whether in data completion or data simulation, this model can meet the actual needs of power analysis.

Acknowledgements This work was supported by State Grid Corporation of China's Science and Technology Project (5108-202218280A-2-401-XG) which is 'Research and Application of Basic Common Service Technology for Digital Twins in Power Grid'.

References

1. Yang, L., Yang, Y., Xu, W., et al.: Research on the technical system and application practice of electric power artificial intelligence. Sci. Inf. Technol. (2020)
2. Yang, T., Zhao, L., Wang, C.: Overview of the application of artificial intelligence in power systems and integrated energy systems. Power Syst. Autom. **43**(1), 13 (2019)
3. Zhang, B., Liu, H., Yu, Z., et al.: Overview of research on network attack detection in power systems based on artificial intelligence. High Voltage Technol. **48**(11), 14 (2022)
4. Liu, D., Li, B., Shao, G., et al.: A framework for intelligent control system of large power grid based on big data and artificial intelligence. Electr. Power Informatiz. **017**(003), 14–21 (2019)
5. Huang, A.: Research and application of key technologies of electric power artificial intelligence open platform. Autom. Instrum. (2020)
6. Hole Assembly Ring Research on Power Artificial Intelligence Image Recognition Technology and Its Application in Overhead Transmission Line Inspection Business. Anhui University (2019)
7. Qi, B., Zhang, P., Zhang, S., et al.: The application status and development prospects of digital twin technology in state assessment of transmission and transformation equipment. High Voltage Technol. (2022)
8. Wang, J., Xiang, Y., He, Z.: Spacecraft system engineering model and implementation based on digital twins. Comput. Integr. Manuf. Syst. (2019)
9. Li, L., Mao, C., Lei, B.: Multi station linkage optimization of IoT hybrid single packaging line driven by digital twins. Packag. Eng. **42**(1), 163–170 (2021)
10. Wang, Y., Liu, B., Xu, L., et al.: Reflection on Digital Twin Technology and the Construction of Smart Water Conservancy Hub. Jiangsu Water Conservancy (2022)
11. Jiang, L., Xie, S., Tian, H.: Blockchain sharding and resource adaptive optimization mechanism for digital twin edge networks. J. Commun. **44**(3), 12–23 (2022)