Data Collection Method for Energy Storage Device of Distributed Integrated Energy Station based on Double Decision Tree

Hao Chen*, Guilian Wu, Linyao Zhang, Jieyun Zheng
Economic and Technology Institute, State Grid Fujian Electric Power Co.,Ltd., Fuzhou Fujian 350013, China

***Abstract*—The distributed integrated energy station includes an electric energy storage device, heat storage device, cold storage device and other devices. Aiming at the problem of low data acquisition accuracy of energy storage device caused by using a single sensor or acquisition scheme in the existing methods, a new data acquisition method of energy storage device of distributed integrated energy station is designed based on double decision tree algorithm. The data acquisition process of double decision tree algorithm is constructed. On the basis of the process, the mathematical models of electric energy storage device, heat storage device, cold storage device and hybrid energy storage device are established. Then the double decision tree algorithm is used to solve the constructed model, and the acquisition pseudo code is given. So far, the data acquisition of distributed integrated energy station energy storage device based on double decision tree has been completed. The results of case analysis show that the accuracy of this method is higher than 98%, and the collection time is less than 30 ms.**

***Keywords*—Double decision tree; distributed; integrated; energy station; energy storage device; data collection**

I. Introduction

The energy storage devices of integrated energy stations such as photovoltaic, electric energy storage, cold and heat energy storage, and natural gas combined cooling, heating and power supply can improve the reliability and energy utilization efficiency of the power supply system [1-2], and reduce system line losses. However, each energy system usually operates independently, resulting in the inability of many energy systems to coordinate effectively, which is prone to safety hazards and low energy utilization problems [3-4]. Therefore, the efficient collection of energy storage device data in distributed integrated energy stations is of great significance to its rapid development [5].

A decision tree algorithm is a method of constructing a decision tree for preliminary screening data based on data attribute information. The algorithm has higher learning performance, higher computing precision and higher computing efficiency. This algorithm has been widely used in many fields [6]. However, the decision tree operation is more complex and the parallel computing capability is low. Based on this, the dual decision tree algorithm came into being. It has the advantage of enhancing the parallel computing capability of the decision tree, and can be applied to collect the data from the energy storage device of the distributed energy station. Among them, the C4.5 algorithm is an efficient algorithm in the decision tree algorithm. Using the C4.5 algorithm to generate a decision tree helps to accurately select data collection points from the data source and facilitate subsequent data processing [7]. The dual decision tree algorithm can use decision trees to filter the data required by the user from the data set. After it establishes the data set, it is processed again through another decision tree to achieve high-quality and high-efficiency data collection.

Li Junnan et al. studied the collection and application of power energy big data based on the big data cloud platform [8]. This method utilizes the big data cloud platform to realize the collection of power and energy big data. Du Peng et al. studied a wide-area data acquisition scheme based on the power dispatching data network [9]. Certain results have been obtained using the wide area method of data collection in the power dispatching data network. However, the above two methods are applied to the data collection of energy storage devices in distributed integrated energy stations. Because the device needs to collect data on four aspects of electricity storage, cold storage, heat storage and hybrid energy storage during the data collection process. As a result, the acquisition accuracy is low and the time is long.

In order to solve the above problems, this paper studies the data acquisition method of the energy storage device of the distributed integrated energy station based on the double decision tree. It establishes the mathematical model of each energy storage device in the distributed integrated energy station. The dual decision tree algorithm is applied to the data collection of the energy storage device of the distributed comprehensive energy station, which improves the data collection effect of the energy storage device of the distributed comprehensive energy station.

II. Data Collection of Energy Storage Devices in Distributed Integrated Energy Stations

In this paper, in the process of researching the data acquisition method of the energy storage device of the distributed integrated energy station, the data sampling idea based on the double decision tree is first designed. Then, on this basis, the mathematical model of each energy storage device in the distributed integrated energy station is established. Finally, the dual decision tree algorithm is applied to the data collection of the energy storage device of the distributed integrated energy station, and the pseudo code is generated.

*Corresponding Author.

A. Data Collection Idea of Double Decision Tree Algorithm

The process of the data collection method of the energy storage device of the distributed integrated energy station based on the double decision tree is as follows:

In the sample, a decision tree for initially collecting the original data of the energy storage device is established through the C4.5 algorithm. It uses the established decision tree to filter and sort out the original data, realize the preliminary screening of the data, and make the re-collected data samples show a balanced distribution state [10]. The established decision tree can be reused, reducing data training and collection time. A second tree is established based on the data contained in the first established tree, and the collected data is further analyzed. The specific process is shown in Fig. 1.

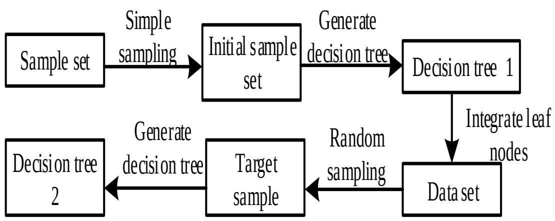

Fig. 1. Double Decision Tree Collection Process.

In Fig. 1, the sample set is the energy storage device data of the target distributed integrated energy station. It includes historical operation data of the energy storage device and corresponding file information.

On the basis of the acquisition process of the double decision tree in Fig. 1, a mathematical model of the energy storage device of the distributed integrated energy station is established.

B. Constructing Mathematical Model of Energy Storage Device in Distributed Integrated Energy Station

Distributed integrated energy stations usually include three loads of heat, cooling and electricity. Therefore, before collecting the data of the energy storage device of the distributed integrated energy station, the mathematical model of the energy storage device of the distributed integrated energy station is established from the aspects of electricity storage, heat storage, cold storage and hybrid energy storage.

1) Electric energy storage device model: The establishment of the model of the electric energy storage device in the distributed integrated energy station should focus on considering the electric energy capacity constraints, the maximum power, the complementary constraints and the charging and discharging efficiency [11]. In the process of model establishment, the energy consumption of the electric energy storage device caused by the change in time during the operation of the electric energy storage device is not considered.

When the time is set as t, the charging power and discharging power of the electric energy storage device are represented by $P_{char}(t)$ and $P_{dis}(t)$ respectively, and $W_E(t)$ is the remaining power of the energy storage device. In this paper, the mathematical model of the electric energy storage device of the distributed integrated energy station is established as follows:

$$\begin{cases} 0 \le P_{dis}(t) \le P_{\max}, 0 \le P_{char}(t) \le P_{\max} \\ W_E(t) = W_E(0) + \int_0^t \left[\eta_c P_{char}(t) - P_{dis}(t) / \eta_d \right] dt \\ W_{E_{\min}} \le W_E(t) \le W_{E_{\max}} \\ P_{dis}(t) P_{char}(t) = 0 \end{cases} \tag{1}$$

In formula (1), η_c represents the charging efficiency of the electric energy storage device. η_d represents the discharge efficiency of the electric energy storage device. $W_E(0)$ and $P_{\max}$ represent the initial remaining power and the maximum charging and discharging power of the electric energy storage device, respectively. $W_{E_{\min}}$, $W_{E_{\max}}$ and $P_{dis}(t)P_{char}(t)=0$ respectively represent the remaining power operation area of the electric energy storage device and the complementary constraint, and the complementary constraint is used to limit the state unity of the electric energy storage device.

2) Case analysis Model of heat and cold storage device: The distributed integrated energy station adopts cold storage and thermal storage to realize sensible heat energy storage and phase change energy storage. Sensible heat energy storage is a low-cost energy storage method. Phase change energy storage realizes energy storage by storing and releasing heat in the process of changing the physical state of materials. It has a higher cost [12], but the temperature fluctuation is small. The phase change energy storage method is selected as the energy storage method of the heat storage and cold storage devices of the distributed integrated energy station. This method has a high energy storage density, and when the heat storage and temperature difference are the same, it can ensure that the heat storage of the phase change material is only one-quarter to one-fifth of the sensible heat material. The temperature during energy storage and release can be controlled and kept constant.

Both cold storage and heat storage devices have energy consumption characteristics when storing and releasing energy. And the energy has a certain dissipation with time.

The model formula of the heat storage (cold storage) device of the distributed integrated energy station is as follows:

$$\begin{cases} 0 \le P_{TI}(t) \le P_{\max}, 0 \le P_{TO}(t) \le P_{\max} \\ W_P(t) = \eta_T W_P(t-1) + \eta_{TI} P_{TI}(t) - P_{TO}(t) / \eta_{TO} \\ 0 \le W_P(t) \le R_{HS} \\ P_{TI}(t) P_{TO}(t) = 0 \end{cases} \tag{2}$$

In formula (2), $W_P(t)$ and $1-\eta_T$ respectively represent the residual heat of the heat storage (cold storage) device at time t and the loss rate of the heat storage (cold storage) device per unit time. η_{TI} represents the heat storage (cold storage) efficiency of the heat storage (cold storage) device. η_{TO} represents the heat (cold) efficiency of the heat storage (cold storage) device.

R_{HS} and $P_{\max}$ represent the maximum capacity and maximum heat release (cold) power of the heat storage (cold storage) device, respectively, $P_{TI}(t)$ represents the heat storage (cold storage) power of the heat storage (cold storage) device when the time t is time; $P_{TO}(t)$ represents the heat storage (cold storage) power when the time t is Cold storage) device to release heat (cold) power.

The calculation formulas of the two heat storage and cooling efficiencies are as follows:

$$\eta_I = \frac{1.16G(\mathrm{T}2-\mathrm{T}1)\times 100\%}{Pt_1/60} \tag{3}$$

$$\eta_O = \frac{1.16G(\mathrm{T}1-\mathrm{T}2)\times 100\%}{Pt_2/60} \tag{4}$$

In the formula, the total amount of hot water collected after 90s of water supply to the end of heat storage or cold storage is described by G . The power is denoted by P , and the collected hot water and cold water time are denoted by t_1 and t_2 , respectively. The mean value of outlet water temperature and the value of cold water temperature are described by $\mathrm{T}1$ and $\mathrm{T}2$, respectively.

1) Hybrid energy storage system model: The internal energy of the energy storage system of the distributed integrated energy station is in a complementary state [13]. Compressed air energy storage and molten salt heat storage are used in molten salt heat storage non-supplementary combustion compressed air energy storage systems. Combining multiple energy stores can improve the economics of energy storage systems.

The turbine inlet is heated by the heat of the molten salt heat storage system of the hybrid energy storage device. Convert thermal energy into electrical energy, and set an electrothermal device to convert electrical energy into thermal energy [14]. It realizes the four-quadrant operation of the hybrid energy storage device of the distributed integrated energy station for two types of energy flows, electric energy and thermal energy.

According to the four-quadrant operation principle of the hybrid energy storage device, P_{in} is represented by P_o and the electric energy input power and output power of the hybrid energy storage device, respectively. H_{in} and H_o represent the electric energy input power and output power of the hybrid energy storage device, respectively. τ_1 represents the energy input time interval. τ_2 represents the energy output time interval, and the model of the hybrid energy storage device can be obtained as follows:

$$\begin{bmatrix} P_o(t) \\ H_o(t) \end{bmatrix} = \begin{bmatrix} \eta_{Com}\eta_{Tur} & \eta_{Tur}\eta_{Hs}\eta_{H-E} \\ \eta_{Heat}(1-\eta_T\tau_1) & \eta_{Hs}(1-\eta_T\tau_2) \end{bmatrix} \begin{bmatrix} P_{in}(t-\tau_1) \\ H_{in}(t-\tau_2) \end{bmatrix} \tag{5}$$

In the formula, η_{Com} represents the compressor efficiency. η_{Tur} stands for turbine efficiency. η_{Hs} and η_{H-E} represent heat storage efficiency and turbine inlet air heating efficiency, respectively. η_{Heat} represents the heater efficiency.

C. Data Acquisition for Energy Storage Equipment of Distributed Comprehensive Energy Station

The dual decision tree algorithm is applied to the established model of each energy storage device of the distributed integrated energy station. The data collection of the energy storage device is realized by using the double decision tree algorithm.

The decision tree algorithm is a typical inductive reasoning algorithm. The decision tree algorithm needs to clarify the attributes of each node of the tree and obtain the required attribute data. In order to improve the operation performance of the decision tree algorithm, the concept of information gain is introduced into the decision tree algorithm. We use the amount of information gained in the operation process to clarify the test attributes required for each node of the decision tree [15], and the decision tree algorithm is usually used to select attributes with a larger number in the operation process.

The C4.5 algorithm is a decision tree method that replaces the information gain of the attribute classification level evaluation index with the information gain rate. This method can effectively improve the defect that the traditional decision tree algorithm is limited to local optimization. The C4.5 algorithm uses the automatic discretization method to process the attributes with continuous values, and avoids the over-learning of the decision tree by pruning the decision tree. Algorithm C4.5 builds a decision tree using information that is relevant to the collection and classification. Let L represent the case set, C_i represent the case sample class label, and $i=1,2,\cdots,n$, and the information entropy formula for data collection can be obtained as follows:

$$I(L) = -\sum_{i=1}^{n} \frac{F(C_i,L)}{|L|}\log_2\frac{F(C_i,L)}{|L|} \tag{6}$$

In the formula, $|L|$ and $F(C_i,L)$ represent the number of samples in the case set L and the number of cases belonging to the C category in the case set A, respectively.

In the case where the number of values k exists in the selected attribute X . According to the probability of each information obtained from the training set, the formula for the conditional entropy of the decision tree is formed as follows:

$$E_x = \sum_{i=1}^{k} \frac{|L_i|}{|L|} I(L_i) \tag{7}$$

In the formula, $|L_i|$ is the number of cases of various subtrees in the attribute X , and the formula for obtaining mutual information is as follows:

$$G(X) = I(L) - E_x \tag{8}$$

Algorithm C4.5 selects the heuristic search extended attribute score, and the extended attribute selects the attribute with the largest information gain. The heuristic method can be effectively applied to the normalization process and in the presence of different attribute values, the attributes that reflect high-quality information gain can be selected. It obtains the attribute information gain rate formula as follows:

$$g_\tau(X) = \frac{G(X)}{Z(X)} \tag{9}$$

Branch the energy storage device data by the value of attribute X to obtain the $Z(X)$ value of the dataset, $Z(X) = -\sum_{i=1}^{k} \frac{|L_i|}{|L|} \log_2\left(\frac{|L_i|}{|L|}\right)$. After completing the above calculations, based on the flow of the dual decision tree data collection method in Fig. 1, simple random sampling is performed on the original data set to obtain an initialization sample B . It applies the C4.5 algorithm to build the first decision tree within the initialization sample. After traversing all the leaf nodes of the first decision tree, the corresponding samples are stored in the sample data set, and the data is sent to the next data set $B_i\{i=1,2,L\ ,m\}$ until the set requirements are met. Sampling stops after all data traversal is completed. Randomly draw samples from the data set obtained by the first decision tree, and use $Z * \frac{|B_i|}{\sum_i^m B_i}$ to represent the number of samples from the data set i , where Z and $|B_i|$ represent the number of target samples and the number of samples in the dataset i , respectively; The target sample Z_t is all the samples extracted, and the second decision tree is generated by using the target sample Z_t that has been extracted. Through the second decision tree, the final collection of the data of the energy storage device of the distributed integrated energy station is realized.

The pseudo code for generating the data collection method of the energy storage device of the distributed integrated energy station with the dual decision tree is shown in Fig. 2.

```
Input: dataset Z
Output: selected data
Initialize Tree←NULL;
If Z is empty or meet other end conditions, then terminate;
End if
For all attributes in the dataset a do
Calculate the information gain rate;
End For
a_best← the attribute with the largest information gain rate;
Tree←a_best is the root decision node;
Zv←Sub-dataset based on a_best division;
For all Zv do
Treev←C4.5(Zv)
Add Treev to the corresponding branch of Tree;
End for
Return Tree
Find Last Child Node(Tree Node node , ArrayList Zi(i=1,2,…,m));
If no child node then
do Zi.Add(node);
If the number of Zi samples < initial limit value then
do i ++;
End if
Else
For Tree Node n in node. Child Nodes do
this Find Last Child Node (n,Zi)
End for
For all Zi do
Random sampling, the number of samples is Z * |B_i| / ∑_i^m B_i ;
Integrate into the target sample set Zt
End for
Input dataset Zt;
Initialize Tree←NULL
Terminate if Zt is empty or encounter other end conditions
End if
For all attributes in the dataset a do
Calculate the information gain rate;
End for
a_best← The attribute with the largest information gain rate;
Tree←a_best is the root decision node;
Zv←Sub-dataset based on a_best division;
For all Zv do
Add the Tree to the corresponding branch of the Tree
Treev ← C4.5 (Zv)
End for
Return Tree
```

Fig. 2. Data Acquisition Pseudocode.

So far, the design of the data acquisition method for the energy storage device of the distributed integrated energy station based on the double decision tree is completed. The next step will verify the effectiveness of the proposed method through case analysis.

III. Case Analysis

In order to validate the data acquisition method of distributed comprehensive energy storage device based on double decision tree is effective. A distributed integrated energy station in a certain industrial park is selected as the experimental object. The distributed comprehensive energy station includes three electric energy storage devices, two heat storage devices, two cold storage devices, and one compressed air hybrid energy storage device. The method of this paper is used to collect the data of each energy storage device of the distributed integrated energy station, and the validity of the data collection method of this paper is verified.

A. Energy Storage Device Data Acquisition Test

We set the sampling time to 20s, and the actual sampling interval is 2s. The LMG671 conventional broadband power detection instrument and the designed data acquisition method of the energy storage device of the distributed integrated energy station based on the double decision tree are used to collect the power of each energy storage device. The results within the statistical sampling time are shown in Tables I to III.

TABLE I. ACTUAL POWER（KW）

Sampling time/s	Electric energy storage device 1	Electric energy storage device 2	Electric energy storage device 3	Heat storage device 1	Heat storage device 2	Cold storage device1	Cold storage device2	Hybrid energy storage device
2	1626	2217	2366	3127	2856	2535	2685	2353
4	1653	2234	2385	3153	2846	2517	2676	2342
6	1624	2285	2350	3148	2862	2538	2648	2364
8	1626	2265	2358	3122	2816	2583	2649	2313
10	1653	2235	2363	3117	2836	2568	2636	2379
12	1635	2220	2349	3155	2866	2531	2657	2365
14	1688	2296	2361	3194	2839	2519	2687	2393
16	1626	2248	2348	3159	2863	2564	2675	2354
18	1687	2236	2317	3176	2818	2576	2632	2359
20	1647	2265	2376	3150	2837	2599	2686	2341

TABLE II. LMG671 CONVENTIONAL POWER ACQUISITION RESULTS（KW）

Sampling time/s	Electric energy storage device 1	Electric energy storage device 2	Electric energy storage device 3	Heat storage device 1	Heat storage device 2	Cold storage device1	Cold storage device2	Hybrid energy storage device
2	1628	2220	2369	3131	2859	2539	2682	2356
4	1650	2237	2388	3150	2842	2520	2672	2346
6	1627	2281	2349	3144	2866	2541	2651	2367
8	1623	2269	2354	3126	2813	2586	2653	2317
10	1657	2231	2367	3113	2840	2564	2632	2382
12	1637	2223	2346	3159	2862	2534	2653	2362
14	1682	2293	2365	3191	2836	2519	2682	2397
16	1623	2243	2345	3163	2860	2561	2672	2350
18	1681	2239	2321	3172	2822	2571	2628	2363
20	1642	2269	2372	3154	2831	2602	2682	2338

TABLE III. THE POWER ACQUISITION RESULTS OF THE METHOD IN THIS PAPER（KW）

Sampling time/s	Electric energy storage device 1	Electric energy storage device 2	Electric energy storage device 3	Heat storage device 1	Heat storage device 2	Cold storage device 1	Cold storage device 2	Hybrid energy storage device
2	1625	2215	2364	3125	2854	2534	2685	2352
4	1652	2234	2385	3152	2843	2516	2675	2341
6	1623	2285	2349	3147	2861	2537	2647	2364
8	1624	2264	2357	3124	2814	2584	2651	2315
10	1652	2234	2361	3117	2834	2567	2635	2378
12	1634	2218	2347	3152	2864	2534	2654	2364
14	1685	2296	2359	3194	2837	2517	2685	2391
16	1625	2247	2347	3158	2861	2564	2675	2354
18	1685	2234	2315	3175	2816	2574	2634	2361
20	1647	2264	2374	3149	2837	2598	2684	2341

Comparing and analyzing the three tables above, it can be seen that the power value of the energy storage device collected by the method in this paper is basically the same as the actual value. However, the difference between the power value detected by the conventional detection device and the actual value is higher than the difference between the method in this paper and the actual value. It shows that compared with the conventional detection device, the power value of the energy storage device collected by the method in this paper is more accurate.

On the basis of the above collection results, in order to further verify the data collection performance of the method in this paper, accuracy is selected as the evaluation index to evaluate the data collection performance. The accuracy of the power data collected by the method of this paper, the big data cloud platform method [8] and the wide-area method [9] is collected for the 8 energy storage devices of the distributed integrated energy station. In order to visually demonstrate the data acquisition performance of the method in this paper, the MATLAB tool is used to generate the comparison results of the power acquisition accuracy of the energy storage device shown in Fig. 3.

From the experimental results in Fig. 3, it can be seen that the power collection accuracy of the distributed integrated energy station energy storage device power using the method in this paper is higher than that of the big data cloud platform method and the wide area method. The accuracy of the method in this paper to collect the data of the energy storage device of the distributed integrated energy station is higher than 98%. The experimental results effectively verify that the method in this paper has high data acquisition performance. The main reason is that the method in this paper uses the double decision tree algorithm to effectively improve the data collection performance through two decision tree processing. It has high data acquisition accuracy and high practicability of data acquisition of energy storage equipment in distributed integrated energy stations.

B. Real-time Test of Data Acquisition of Energy Storage Device

The real-time data collection of each energy storage device in a distributed integrated energy station affects the normal operation of the energy storage device. Set the attribute data to be collected to 2500. The method of this paper is used to collect the average collection time of each sample of the power, DC current and DC voltage of the energy storage equipment of the distributed integrated energy station. The statistical results are shown in Tables IV to VI.

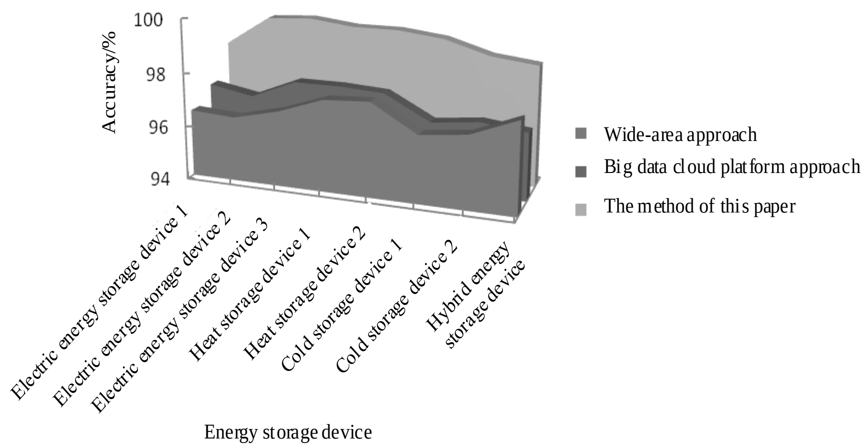

Fig. 3. Accuracy of Power Collection of Energy Storage Devices.

TABLE IV. ENERGY STORAGE DEVICE POWER COLLECTION RESULTS

Attribute data/pcs	Electric energy storage device 1/kW	Electric energy storage device 2/kW	Electric energy storage device 3/kW	Heat storage device1/kW	Heat storage device2/kW	Cold storage device1/kW	Cold storage device2/kW	Hybrid energy storage device/kW
500	1625	2215	2364	3125	2854	2534	2685	2352
1000	1652	2234	2385	3152	2843	2516	2675	2341
1500	1623	2285	2349	3147	2861	2537	2647	2364
2000	1624	2264	2357	3124	2814	2584	2651	2315
2500	1652	2234	2361	3117	2834	2567	2635	2378

TABLE V. DC VOLTAGE ACQUISITION RESULTS OF ENERGY STORAGE DEVICE

Sampling time/s	Electric energy storage device 1/V	Electric energy storage device 2/V	Electric energy storage device 3/V	Heat storage device1/V	Heat storage device2/V	Cold storage device1/V	Cold storage device2/V	Hybrid energy storage device/V
500	125.4	135.6	151.2	141.5	131.2	128.5	171.5	161.9
1000	124.2	134.5	150.4	141.6	131.5	128.6	170.8	160.24
1500	125.2	135.2	151.2	142.2	131.6	128.9	171.2	161.3
2000	125.4	134.4	151.3	141.6	131.8	128.7	171.4	161.8
2500	125.5	135.5	151.4	142.5	131.7	128.3	171.2	161.5

TABLE VI. DC CURRENT ACQUISITION RESULTS OF ENERGY STORAGE DEVICE

Sampling time/s	Electric energy storage device 1/A	Electric energy storage device 2/A	Electric energy storage device 3/A	Heat storage device1/A	Heat storage device2/A	Cold storage device1/A	Cold storage device2/A	Hybrid energy storage device/A
500	5.75	6.35	6.25	5.64	6.35	9.52	7.52	8.25
1000	5.35	6.52	6.34	5.28	6.85	9.45	7.46	8.64
1500	5.46	6.45	6.15	5.46	6.28	9.25	7.18	8.15
2000	5.27	6.85	6.28	5.31	6.45	9.34	7.64	8.34
2500	5.81	6.34	6.54	5.28	6.75	9.52	7.58	8.62

From the analysis of Tables IV to VI, it can be seen that the method in this paper can complete the target collection attribute data volume. And it counts the time required to collect different data volumes during the collection process, and draws its statistical results as Fig. 4.

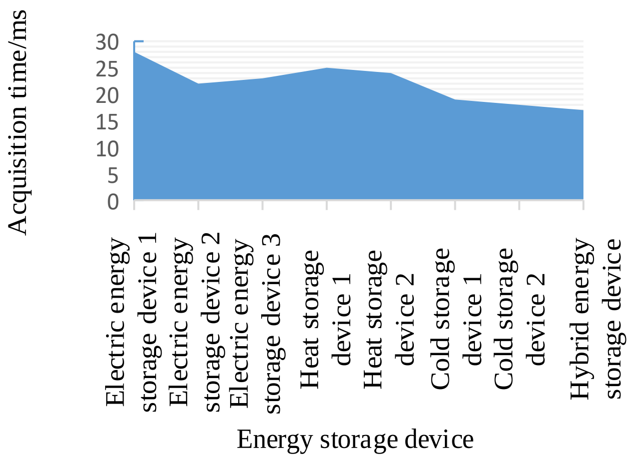

Fig. 4. Real-time Data Collection

The experimental results in Fig. 4 show that the power, DC current and DC voltage data collected by the method in this paper are all less than 30 ms. The reason for the short acquisition time of the method in this paper is that the decision tree for preliminary acquisition of the original data of the energy storage device is established by the C4.5 algorithm. It uses the established decision tree to filter and sort out the raw data. It realizes the preliminary screening of data and improves the collection efficiency. Using the method in this paper to collect various data from energy storage equipment in distributed integrated energy stations not only has high acquisition accuracy. This method has a high real-time acquisition, which again verifies the high data acquisition performance of the method in this paper. It can be applied to the practical application of data acquisition of energy storage devices in distributed integrated energy stations.

IV. CONCLUSION

Distributed comprehensive energy stations are an important direction for integrating and optimizing the energy Internet. In this paper, a mathematical model of each energy storage device in a distributed integrated energy station is established. It collects energy storage device data through a dual decision tree method and draws the following conclusions:

1) The accuracy of the proposed method is higher than 98%, and the collection time is less than 30ms, which can improve the data collection efficiency and collection accuracy. The method in this paper can obtain ideal data collection results in a short time, and the tree-like decision tree structure can intuitively reflect the status of the data to be collected and improve the accuracy of data collection. The method is applied to the data collection of energy storage devices in distributed integrated energy stations, and has high engineering practice value.

2) The proposed method improves the defect of inaccurate collection caused by too much useless information in massive data. It avoids unbalanced sampling space caused by differences in data distribution between different samples, and provides greater convenience for subsequent data analysis.

3) In this study, it is not considered that the electric energy storage device can provide heat and cooling data through the heat pump in the process of outputting electricity. The accuracy of power collection of energy storage devices needs to be further improved and studied.

REFERENCES

[1] Y. H. Jia and F. Zhang, "A bi-level optimal configuration of multiple storage in reginal integrated energy system with distribution wind power inclusion," Renewable Energy Resources, 2019, 37(10), pp. 1524-1532.

[2] C. L. Wang, H. Liu and J. F. Gong, "Joint Scheduling of Different Energy Storage for Improving Wind Power Accommodation Ability in Integrated Community Energy System," Electric Power Construction, 2018, 039(4), pp. 35-44.

[3] Z. H. Jiang, Y. Q. He, L. L. Cao, "Reconfiguration of distribution network with distributed generations and energy storing devices based on improved genetic algorithm," Power System Protection and Control, 2018, 046(5), pp. 68-72.

[4] W. Xiong, Y. Q. Liu, W. H. Su, "Optimal configuration of multi-energy storage in regional integrated energy system considering multi-energy complementation," Electric Power Automation Equipment, 2019, 39(1), pp. 124-132.

[5] Y. Lu, Y. Dai, W. Z. Ma, "Decentralized Dynamic Optimal Power Flow in Distribution Networks With Distributed Generation and Energy Storage Devices," Power System Technology, 2019, 43(2), pp. 434-442.

[6] Y. W. Liu, Y. Hu, N. L. Tai, "Rule extraction method of operation and maintenance expert system for an intelligent substation based on the decision tree," Journal of Electric Power Science and Technology, 2019, 34(1), pp. 125-130.

[7] W. Q. Sun, Z. Li, Y. M. Tan, "Method of Power System Energy Storage Configuration Based on Flexibility Promotion," Journal of System Simulation, 2018, 30(1), pp. 235-241.

[8] J. N. Li, W. Li, H. J. Li, "Research on big data acquisition and application of power energy based on big data cloud platform," Electrical Measurement & Instrumentation, 2019, 56(12), pp. 104-109.

[9] P. Du, L. Yan, B. C. Gao, "Wide-area Data Acquisition Scheme Based on Power Dispatching Data Network," Automation of Electric Power Systems, 2019, 43(13), pp. 156-161.

[10] L. Chen, H. X. Fei, H. L. Ding, "A data sampling method based on double decision tree," Computer Engineering and Science, 2019, 41(01), pp. 134-139.

[11] L. P. Zhang, X. Ye, J. Wang, "Research on method for data collection based on distribute power generation," Renewable Energy Resources, 2017, 35(8), pp. 1203-1207.

[12] T. F. Ma, J. Y. Wu, L. L.Hao, "Energy Flow Modeling and Optimal Operation Analysis of Micro Energy Grid Based on Energy Hub," Power System Technology, 2018, 42(1), pp. 179-186.

[13] K. K. Gu, W. S. Hu, K. Zhao, "Design and implementation of a hand-held device for power data acquisition and analysis based on mobile network," Power System Protection and Control, 2018, 506(8), pp. 115-121.

[14] H. B. Kang, Y. N. Qu, L. Zhao, "Research on microgrid control technology with distributed power supply and energy storage device," Chinese Journal of Power Sources, 2017, 41(4), pp. 627-629.

[15] L. L. Chen, L. H. Mu, X. F. Xu, "Influences of energy storage operational strategy and characteristic on microgrid reliability," Electric Power Automation Equipment, 2017, 37(7), pp. 70-76.