Research on big data storage and compression method for dual core intelligent meter

Wei Ming^{1,a}, You Feng Xiang^{2,b}, Wang Qiang^{1,c}, Shen Xiao Feng^{1,d}

^{1}Department of electronic information, Applied Technology College of Soochow University, Suzhou, China
^{2}Institute of Technology, Applied Technology College of Soochow University, Suzhou, China
^{a}weim_anap@suda.edu.cn, ^{b}youfengxiang98@163.com, ^{c}wayne1024@163.com, ^{d}451340839@qq.com

Abstract—With the rapid integration and development of cloud computing, big data and smart grid, not only the type and accuracy requirements of the collected data are increasing, but also the storage depth is getting higher and higher. In hardware, the dual port RAM design is used to improve the real-time performance of reading and writing, digital filter is used for preprocessing in software, and "Equivalent Normalization (EN)" is used to improve storage depth. These schemes save hardware storage space and improve data transmission efficiency in practical application. So the advantages of dual-core intelligent power meter will be further exploited and brought into play.

1. INTRODUCTION

Smart grids (SGs) have emerged as a new architecture for the EPSs [1-2], which combines advanced sensing and measurement technology, information and communication technology, analysis and decision-making technology, automatic control technology and energy and power technology, and is highly integrated with power grid infrastructure. It can not only exchange data and information between power grid equipment, but also further complete the functions of real-time automatic control, intelligent regulation, online decision-making analysis and so on. As a digital and automated intelligent power meter has maintained power quality and reduce power quality problem in the network [3]. In smart grid, intelligent power meter is one of the main sources of the whole data, the real-time user-side metering data may be frequently collected for monitoring and controlling electricity consumption [4]. The large amounts of data collected by smart meters [5], such as temperature, date, voltage (V) and current (I), while derived parameters could be active power (P), reactive power (Q), power factor (PF), energy (E), total harmonic distortion (THD), and frequency (Hz) [6]. Using these data, power consumption behaviors of electricity consumers can be revealed [7], which facilitates applications such as load forecasting, demand response, anomaly detection, electricity price design, and energy efficiency promotion [8]. When a single parameter is stored per half-second in 16 bits with time axis, it needs at least 1.5 M bytes every day. The "big data analysis [9,10]" demand of smart grid requires not only large amount of data and many kinds of data, but also high sampling density. According to the requirements of IR46, there should be corresponding data backup. It can be said that the demand of storage space is geometrically doubled. Therefore, data compression storage and security measures are urgent matters to be solved. Based on the
dual core intelligent power meter of "metering processor" and "management processor", this paper makes a systematic and in-depth analysis from three aspects, and puts forward the corresponding improvement scheme and optimization measures, in three major areas: hardware storage circuit, data acquisition and preprocessing, data compression and storage, which not only saves hardware storage space, but also improves data transmission efficiency.

2. IMPROVEMENT MEASURES

2.1 Improvement of hardware memory circuit
At present, SPI, serial port, I2C and so on are mostly used in dual core meter system (metering core and management core). The real-time performance of this communication is not high. In the "control operation" interaction process, this kind of communication real-time performance is not high, suitable for small amount of data interaction, and its anti-interference performance is not obvious. However, in the "big data" smart grid, a large number of real-time reading will directly affect the data results. For example, the metering core of a single core smart meter must interrupt the current task, and the data reading and writing operation or other control operations of the management core should be handled, but the continuous collection of electric energy cannot be carried out. In this paper, a dual port RAM is proposed to make the dual core read and write memory data at the same time. Dual port RAM has two sets of completely independent data lines, address lines and read-write control lines, and allows two processors to access the memory randomly at the same time, i.e. shared dual port RAM. According to the agreement, the fixed address is allocated as the command interaction area and shared data area. In this way, the metering core can read the memory of the command interaction area regularly and operate according to the "instructions" of the area. This kind of active mode is helpful to improve the periodicity and stability of data acquisition, and when real-time control is needed, the hardware interrupt mode is used to coordinate. The management core can be more convenient to read the data stored in the metering core in real time without interfering with the work of the metering core. The hardware communication is only used for real-time emergency processing. The schematic diagram of dual core interaction is shown in Fig 1’.

![Fig 1 Schematic diagram of data interaction of dual core intelligent meter with dual port RAM](image)

Note: due to the limitation of memory space, the metering core should be able to cycle storage within a certain range of the storage area, and the management core should be able to read out in time. In order to prevent the later data coverage, the status word of data reading and writing can be opened up. When the data coverage occurs, on the one hand, it will send out error prompt; on the other hand, it will use hardware communication link to coordinate processing to improve the reliability of data. Through the improvement of the above circuit, it can better reflect the advantages of dual core, and is more conducive to the accuracy, real-time and stability of data acquisition.

2.2 Data security and backup
The data storage structure must be determined according to the design cost and security level requirements. For example, the data partition and backup are performed by single chip high capacity chip. The advantage of this scheme is that the hardware circuit is simple and the programming is convenient.
The disadvantage is that the memory chip is damaged, the price is expensive, and the backup data is damaged together. The advantage of using multiple independent chips for backup is that when one memory chip is damaged, other backup chips will not be affected, and the single chip cost is low, the disadvantage is that the circuit structure takes up multiple I/O port resources. Generally, in order to ensure the absolute security of data, which is usually backed up by three chips in the very important system.

2.3 Data acquisition and preprocessings

By using the dual port RAM as the "shared" unit, the measurement processor can focus on measuring data and have more time for periodic data acquisition and pre-processing. Due to the variety of user equipment, as well as the uncertainty and randomness of working period and state, the power grid has more noise, which is not conducive to data acquisition. The common interference factors are: ambient temperature (causing measurement error to drift to negative direction), grid harmonic voltage and inter harmonic voltage and frequency fluctuation, etc.; In order to ensure the stability, accuracy and compressibility (the higher the equivalent repetition rate, the more conducive to compression storage) of collected data, in addition to using hardware filter circuit, it is also necessary to select appropriate filtering algorithm in software. There are many data filtering methods applied to intelligent meter [11]. In order to reduce the amount of calculation, this paper mainly selects the limiting filtering method, arithmetic average filtering method and median value average filtering method. The characteristics and applicability of several filters are analyzed as follows.

- Characteristics of limiting filtering method: firstly, according to the collected data, determine the maximum allowable deviation value of two sampling (set as x), and judge when a new value is detected: if the difference between the current value and the last value is not greater than x, the current value is valid; otherwise, the current value is invalid, and the previous value is replaced by the last value. The advantage of this algorithm is that it can effectively overcome the impulse interference caused by accidental factors, such as the instantaneous impulse noise of household appliances.

- Features of arithmetic average filtering method: n sampling values are continuously taken for arithmetic average operation. When n value is large, signal smoothness is higher but sensitivity is lower; when n value is small, signal smoothness is low but sensitivity is high. The advantages are: it is suitable for filtering the signals with random interference; for example, the frequent switching of various electrical appliances in the daytime and the irregular change of the data collected by the watt hour meter are more suitable for selecting this algorithm.

- The characteristics of median filtering method: continuously sample n data, remove a maximum value and a minimum value, and then calculate the arithmetic mean value of n-2 data. The advantages of this method are: it can effectively overcome the fluctuation interference caused by accidental factors, and has a good filtering effect on the slowly changing measured parameters; it is conducive to the overall power measurement of household appliances when the meter is running at night.

Software filtering algorithm is far more than these, but no matter what method is selected, the collected data should have better smoothness, which is conducive to the compression and storage of later data. The software filter of the metering core can be selected by the management core intelligently. That is, after obtaining a certain amount of data, the management core can make use of the advantages of big data analysis. In addition to the purpose of data compression and storage, it should also summarize and analyze the data status in different periods, especially select a reasonable filter type for the metering core, and reflect the "self-optimization consciousness" of the dual core system, such as: a virtual sensor based self-adjusting control for hvac fast demand response in commercial buildings towards smart grid applications [12].

3. DATA COMPRESSION

With the rapid development of smart grid, more and more data types are collected by intelligent power meters, and the storage depth is higher and higher; the higher the sampling density, the more accurate the calculation value is, the more conducive to "big data analysis"; Therefore, to store data for one year (or
longer), the amount of data is large, which undoubtedly increases the hardware cost and the difficulty of circuit design, so the data must be compressed and then stored.

3.1 Data compression

The compression scheme is shown in Fig 2.

Analysis of power consumption of household appliances: under normal circumstances, in the process of residential electricity consumption, the current of refrigerators, washing machines, television sets, air conditioners, water heaters, lamps, etc. during the use of these devices, the current changes little in a short period of time. If \( \Delta t = 1 \) minute is taken as the data storage interval, and the integration takes the average value, the probability that the values of adjacent time points are equal is very high, which provides conditions for data compression. Especially in people's sleep time, in addition to the power changes caused by intelligent control of electrical appliances, there are few human interference factors, then the points of equivalent data are very denser. Therefore, the "Equivalent Normalization (EN)" method is used for compression storage, and only the equivalent array subscript is stored if the electric energy data is equal within the test accuracy range. From the Figure 2 scheme diagram, we can see that: for the first 16 bytes, the original data before compression is 16 × 2 byte = 32 bytes; Data volume after compression: 16 bytes + 4 × 2 byte = 24 byte, then 25% of space are saved; Especially for the night time, with the reduction of people's activities, most electrical equipment are in standby, shutdown, or stable operation state, the load current measured by the meter is basically unchanged, the measured data repetition rate is higher, based on the normal 7-hour sleep, in the worst case, a day of data can save 29.17% storage space. The flow chart of program algorithm is shown in Figure 3. In Figure 3, the array M[i] represents the time point to be stored, and its value matches the subscript of the array data [J], i.e., M [i] = j. In order to cope with more complex situations, with the development of intelligent power meter, especially for dual core meters, this data storage scheme has more obvious advantages. The data management core should reasonably and intelligently select the storage interval of 1, 2, 3 points according to the current change rate of users 15 points (or set by the power supplier).
3.2 Data storage
The data storage classification process algorithm is shown in Fig 4.

Due to the complexity of user equipment and randomness of power consumption habits, the power changes rapidly. In a certain period of time, there are a large number of unequal values, which is not conducive to the implementation of the "Equivalent Normalization (EN)" compression algorithm. Therefore, these data must be classified and stored in the private array. Generally, the address subscript is byte variable, the maximum compression interval is 256 data, which is a "Data Block". If the data per minute is different, the actual use of storage space settlement result is: 256 subscript variables + 256 double byte data, i.e. 768 bytes; Compared with the conventional storage scheme, it wastes 256 bytes of subscript space. In order to eliminate this situation, at the same time, it takes 4 hours and 240 minutes to round up. Convention: in a data block, when the amount of data is greater than 240, it will be treated as special data, and exactly 16 special marks will be used. Each flag corresponds to a data type, so the time point data can be identified with the array subscript. In this way, the data management MCU should select the data of this period and store it in the normal return scheme stored procedure. The storage scheme is shown in Fig 5.
4 hour data that can be compressed

The flag F1 area can be stored for 4 hours and 240 data

The flag F2 area can be stored for 4 hours and 240 data

The flag F0 area can be stored for 4 hours and 240 data

Fig 5 Partition storage and query graph of data

In the data block from F1 to F6 (F3 to FF are not fully drawn), each area is stored for 4 hours, and the 24-hour record can be completed by using 6 areas. Of course, the data difference between two adjacent time points of load current is very low during the whole day. The index of array M corresponds to the time point, which is used to record time. The value is a variable and does not occupy memory space. Data of array m and query algorithm of "F1 area - FF area": first judge whether the value of member variable of array M is in the range of 0XF1-0XFF, find the member M [241] whose first value is F1 in the array M, and then map it with the subscript of the first member address of the data in F1 area, and so on. Through the above-mentioned data compression, the actual need to upload less data, as long as the system carries out data matching and mapping according to the agreed protocol, the correct data can be obtained.

4. SUMMARY

The intelligent meter designed with dual port RAM can improve the accuracy, real-time performance and stability of data acquisition of metering core, and has certain reference value for other intelligent power electronic devices. In terms of data acquisition, the management core can dynamically control the software filtering algorithm of the metering core, improve the stability of the data and its own error correction rate, and better reflect the "intelligence" and advantages of the dual core. The "Equivalent Normalization" has a high compression rate, does not change the authenticity of the original data, and the data recovery is very easy, which not only promotes the development of dual core smart intelligent power meter under the IR46 standard, but also provides guarantee for cloud computing, real-time analysis and control of big data in smart grid.

REFERENCES

[1] Douglas L.S. Mendes, Ricardo A.L. Rabelo, Artur F.S. Veloso, Joel J.P.C. Rodrigues, Jose V. dos Reis Junior. An adaptive data compression mechanism for smart meters considering a demand side management scenario[J]. Journal of Cleaner Production, 2020, 255.

[2] Foad H. Gandoman, Abdollah Ahmadi, Adel M. Sharaf, Pierluigi Siano, Josep Pou, Branislav Hredzak, Vassilios G. Agelidis. Review of FACTS technologies and applications for power quality in smart grids under the IR46 standard, and also provides guarantee for cloud computing, real-time analysis and control of big data in smart grid.

[3] Amam Hossain Bagdadee, Md Zahirul Hoque, Li Zhang. IoT Based Wireless Sensor Network for Power Quality Control in Smart Grid[J]. Procedia Computer Science, 2020, 167.
[4] Ding Yong, Wang Bingyao, Wang Yujue. Secure Metering Data Aggregation With Batch Verification in Industrial Smart Grid. IEEE Transactions on Industrial Informatics 2020;10: 6607-16.

[5] Wilson L. Rodrigues Junior, Fabbio A.S. Borges, Artur F. da S. Veloso, Ricardo de A.L. Rabêlo, Joel J.P.C. Rodrigues. Low voltage smart meter for monitoring of power quality disturbances applied in smart grid[J]. Measurement, 2019, 147.

[6] Sawsan Sayed, Tasnim Hussain, Adel Gastli, Mohieddine Benammar. Design and realization of an open-source and modular smart meter[J]. Energy Science & Engineering, 2019, 7(4).

[7] Wang Y, Chen Q, Kang C, Xia Q, Luo M. Sparse and redundant representation-based smart meter data compression and pattern extraction. IEEE Trans Power Syst 2017;32:2142–51.

Wang Y, Chen Q, Hong T, Kang C. Review of smart meter data analytics: applications, methodologies, and challenges. IEEE Trans Smart Grid 2018..

[8] Shouxiang Wang, Haiwen Chen, Lei Wu, Jianfeng Wang. A novel smart meter data compression method via stacked convolutional sparse auto-encoder[J]. International Journal of Electrical Power and Energy Systems, 2020, 118.

[9] Sulaiman S. M., Aruna Jeyanthi P., Devaraj D.. Smart Meter Data Analysis Using Big Data Tools[J]. Journal of Computational and Theoretical Nanoscience, 2019, 11(2).

[10] Zafar A. Khan, Dilan Jayaweera, Manuel S. Alvarez-Alvarado. A novel approach for load profiling in smart power grids using smart meter data[J]. Electric Power Systems Research, 2018, 165.

[11] Lewis Mitchell, Georg A. Gottwald. Controlling model error of underdamped forecast models in sparse observational networks using a variance-limiting Kalman filter[J]. Quarterly Journal of the Royal Meteorological Society, 2013, 139(670).

[12] Energy; Findings on Energy Reported by Investigators at Sun Yat-sen University (A Virtual Sensor Based Self-adjusting Control for Hvac Fast Demand Response In Commercial Buildings Towards Smart Grid Applications)[J]. Energy & Ecology, 2020.