1. Introduction

Chinese power network structure and operation/control characteristics have witnessed profound change, due to commission and operation of Ultra High Voltage (UHV) AC/DC power transmission, grid-connected massive new energy, and application of new types of devices of Flexible AC Transmission Systems (FACTS). As a result, power system stability problems under faults are more complicated and real-time dispatching and operation face great challenges. Among them, Low-Frequency Oscillation (LFO) is one of the key factors that threaten interconnected power grids’ security and stability and restrict transfer capability [1–3]. While power system incurs disturbance, LFO from 0.1 Hz to 2.5 Hz may occur in relative rotor angles and tie line power between different generator groups, which may grow sustainably and finally cause collapse of power system if damping is insufficient [4]. According to its mechanism, LFO is divided into three types: weak damping, forced, and nonlinear, among which weak damping and forced oscillation accidents have once happened in Henan, Yunnan, western Inner Mongolia, Xinjiang, and some other places in China [5–7]. Therefore, it is necessary to utilize multiple information sources and advanced algorithms to analyze and control LFO, to deal with damage to power system that LFO brings.

Wide area measurement system (WAMS) uses Phase Measurement Unit (PMU) to synchronously record dynamic operation trajectory. Many studies have been carried out on its application in state estimation, LFO, transient stability judgment, voltage stability estimation, and so forth [8–15]. However, as WAMS is applied in real power systems, many advanced functions are hard to be practical due to the problems of devices, communication, master station maintaining and management, and so forth [16]. The main reason is that WAMS data accuracy, continuity, and stability are impaired by multiple influence factors, which cause the methods depending heavily on data accuracy and continuity...
to lose practical value. In other aspects of information processing, such as radar signal processing, [17] applies a reduced-dimension multiple signal classification (MUSIC) algorithm to improve search performance, [18] generates large-scale antenna coarray by expanded coprime array structure, and [19] proposes that array data is expressed as triple linear model, and radar angle is estimated by 2-dimensional local search. In addition, an improved radar parameter estimation algorithm is proposed in [20]. These methods cannot be used for power Low-Frequency Oscillation due to tolerance of data estimation online calculation.

In this paper, the role that wide area measurements data plays in PFO analysis key technologies is analyzed. A new model of double-set elastic store is proposed to solve the problems in continuous and reliable access of the wide area measurements data which is dense and of large scale and has quick update rate. Comparison of the new model and conventional linear data store model and its practical application validate its accuracy.

2. Power System LFO Analysis Principle

The phenomenon of LFO and its influence factors in large-scale power system are complicated and diversified. Weak damping and forced oscillation mechanisms are among the common causes of LFO in power systems.

2.1. Online LFO Analysis Frame. To realize analysis and monitoring of LFO in power systems, on the one hand, forced oscillation source searching and positioning as well as modal analysis and model calculation are carried out, utilizing multiple information sources of WAMS data, Energy Management System (EMS) data, power system dynamic parameters, and so forth; on the other hand, frequency domain simulation analysis is carried out through integration of multiple information sources to search for the possible weak damping model and its control strategy. Through matching of the calculated modal and model with the frequency simulation output, LFO control strategies are obtained. The online analysis and control frame are presented in Figure 1.

2.2. Key Roles WAMS Data Playing in LFO Analysis. As data source of online analysis, WAMS data plays a key role in the accuracy of LFO analysis. The key roles are presented as follows:

(1) To Increase Accuracy of Online Data Integration

If upload rates of EMS data and WAMS data are considered as 1 s and 10 ms, respectively, the WAMS data is then 100 times as large as the EMS data. WAMS data acquisition has higher AD precision and hence is able to be used in calibration of EMS state estimation data. Take a node voltage as example and let EMS data at a time instant be \( v_1^{(E)} \), while WAMS has measured values of \( v_1, v_2, \ldots, v_n \); we obtain

\[
V_{\text{Max}} = \text{MAX}(v_1, v_2, \ldots, v_{n-1}).
\]

\[
V_{\text{Min}} = \text{MIN}(v_1, v_2, \ldots, v_{n-1}).
\]

\[
V_{\text{AVG}} = \frac{1}{n} \sum_{k=1}^{n} v_k.
\]

During the time window, transmission faults are excluded, and the relationship between the two types of data is

\[
V_{\text{Min}} < v_1^{(E)} < V_{\text{Max}}
\]

\[
V_{\text{Min}} < V_{\text{avg}} < V_{\text{Max}}
\]

\[
V_{\text{Min}} < V_N < V_{\text{Max}}
\]

where \( k \) is the maximum value, \( v_k \) is the minimum value, \( T \) is the average value, and \( \delta \) is the rated value. When EMS state data is calibrated, equations (1)–(3) are first calculated; if (4)–(6) are tenable, then data \( v_1^{(E)} \) is trusted and more strict calibration can be proceeded, and if there are generators on the node of power system topology, there should exist \( v_1^{(E)} \geq V_N \); if there is no generator but there is load, there should exist \( v_1^{(E)} \leq V_N \). If there exists \( v_1^{(E)} \leq V_{\text{Min}} \) or \( v_1^{(E)} > V_{\text{Max}} \), then the data is abnormal, and \( v_1^{(E)} \) should be calibrated by \( t \).

(2) To Calculate the LFO Model and Modal

Tie line power trajectory is extracted form WAMS data, (7) is calculated through Prony algorithm, magnitude, initial phase angle, and attenuation factor, and frequency can be obtained from the complex exponential linear model:

\[
P_n = \sum_{i=1}^{k} b_i z_i^T = \sum_{i=1}^{k} A_i e^{j \phi_i} e^{j (\alpha_i + j 2 \pi f_i) n T},
\]

where \( p_n \) is the power trajectory sampling value within computing time window, \( N \) is the number of samplings, \( k \) is the number of fitting functions of complex exponential, and \( A_i, \phi_i, \alpha_i, \) and \( f_i \) are magnitude, initial phase angle, attenuation factor, and frequency, respectively.

The \( k \) model magnitudes \( \{ A_1, A_2, \ldots, A_k \} \) are sorted, and the frequency \( f_i \) according to the maximum magnitudes \( A_{\text{max}} \) is considered as the oscillation dominant model.

Likewise, the generator rotor angle or voltage phase angle trajectories within the same computing time window can be extracted and analyzed by Prony algorithm. Together with power judgment, initial phase angles \( \theta_i \) are sorted, and the relative angle differences of the nodes with the maximum initial phase angle as reference are then calculated. The
power system modal distribution can be obtained through observation of the distance between the relative angle differences of the nodes and 0, π axis.

(3) To Calculate Energy Function for Forced Oscillation Source Searching

If power systems incur oscillation with good damping, it may be forced oscillation. Energy function can be used to search for the forced oscillation source, and it is expressed as

\[ W_{ij} = \int_0^t \Delta P_{ij} d\Delta \theta_i + \Delta Q_{ij} d\ln U_i, \]  

(8)

where the variable physical meaning can be seen in [11, 21, 22]. Through PMU information, the transmission line active power \( \Delta P_{ij} \), the transmission Line reactive power \( \Delta Q_{ij} \), the node voltage \( \Delta \ln U_i \), and the node phase angle \( \Delta \theta_i \) can be calculated. \( dW_{ij}/dt \) within continuous computing time windows is then obtained. If \( dW_{ij}/dt > 0 \), the forced oscillation source is considered as in this generator or node.

The generator potential energy or branch potential energy is considered as an index of the influence that generator has on oscillation. If the energy is positive or negative that flows from the generator to network, it can reflect whether the generator is exciting or suppressing the oscillation, then, through observation of the generators’ energy variation, the generator that the oscillation source is located at and the path that the energy flows through can be identified.

All of the above three key technologies need to use WAMS information as data source, which speaks volumes for WAMS big data playing a key role in LFO online analysis.

3. WAMS Big Data

3.1. WAMS Structure. WMAS is composed of PMU, communication system, and wide area measurement master station [23], as is shown in Figure 2. PMU device acquires information of voltage, current, rotor speed, and switch signal, receives order from WAMS master station to send equal interval dynamic data, and sends transient data if called by master station.

3.2. WAMS Data Characteristics. Wide area measurement data is from PMU devices, including transient data, dynamic data, and events marking data, as is shown in Figure 2. The transient data is a file saved by PMU of the instantaneous value records within a setting duration as power system is under contingency. As national standard GB/T26865.2-2011 [24] requires, the minimum sampling rate is 4 800 S/s. The data records should also meet the requirements of international standard ANSI/IEEE C37.111-1999 (COMTRADE). The dynamic data is a series of power system operation behavior values, which PMU obtains through one or multiple calculations based on the instantaneous value of direct acquisition. The events marking data is mainly applied in identification of power system disturbance form. Besides, relays and security automatic equipment actuation, manual operation, or synchronous clock abnormal may also generate events marking (Figure 3).

3.2.1. Wide Area Measurement Data Scale. Transient data is usually stored in PMU as binary data and is called by WAMS master station only as needed. Dynamic data such as three-phase voltage, current phase, excitation voltage and current, rotor speed, and frequency are centralized and stored at master station, also usually as binary data. The transient data and dynamic data form the subject of wide area measurement data. Events marking data’s storage format is versatile but usually is ASCII. It occupies a relatively small part of the quantity of wide area measurement data.

The calculation formula of the dynamic data scale stored in WAMS master station is as follows:
The left side of the equation represents storage space scale (unit: bytes). \( i \) represents a single PMU device, and WAMS receives \( N \) PMU device data. \( p \) is the number of recorded phases. \( a \) is the number of recorded analogs. \( s \) is the number of recorded switch signal groups and a group has 32 switch signal states. \( v_k \) is dynamic data transmission rate, and \( v_k \) is usually 25, 50, and 100 in China. \( T \) is the designated storage time length. \( \delta \) is the additional space as dynamic data is stored as file and may be taken as 0. \( C \) is the number of files needed for splitting dynamic data. All the variables are integer.

The scale of the transient data stored in PMU depends on transient wave recording frequency, the type and number of wave recording channels, wave recording length, and the number of start-ups. Its calculation formula is as follows:

\[
M_T = \left( \sum_{i=1}^{N} \left( (8p_i + 4a_i + 4s_i) + 20 \right) \right) v_k T + \delta * C, \quad (9)
\]

\((i = 1, \ldots, N, k = 1, 2, 3)\).

Let the additional space in (9) be 0; we obtain that the data quantity of a single data influx is \( \sum_{i=1}^{N} ((8p_i + 4a_i + 4s_i) + 20) \), as in (9). A single data influx costs at least 10 ms, but no more than 40 ms, as required.

### 3.3. Difficulty Analysis of WAMS Data Processing

WAMS data has rapid update rate and diversified data storage structures, as Section 3.2 expresses. How large its data scale is can be shown by the WMAS data of a province of China.

The province has 87 power plant stations, with a total of 155 installed PMU devices. WAMS master station is located in the province dispatching control center. According to the WMAS dynamic data configuration table, the power system has a total of 5 393 phases, 3 426 analogs, and 74 double-byte switch signals.

Let the additional space in (9) be 0; we obtain that the dynamic data quantity once uploaded is 60 256 bytes. If PMU sends data to master station at the rate of 50 times per second, WAMS master station receives 0.24 TB (1 TB = 1 099 511 627 765 bytes) data per day.

As for the transient data, if we insert 155 PMU measurement channels to (10) and let the additional space be 0, one sampling point of all the PMU devices costs 35 572 bytes of storage space. If the sampling rate is 4 800 Hz, if the recording time duration before and after event starts is 5 s and 15 s, respectively, and if all the wave recordings in power system have the same set, a wave recording of transient data costs \( 3.106 \times 10^{-3} \) TB of storage space. If transient data wave recording is 20 times a day, the WAMS data scale table of transient and dynamic data is presented in Table 1.

As Table 1 presents, the WAMS transient and dynamic data of a province in one month is closed to 9 TB in quantity. Without periodical data cleaning, the data scale of a year will be more than 100 TB. The data scale is so large that communication is difficult to guarantee a steady stream without broken link or reconnection, even if high speed communication network is used. Together with WAMS data’s characteristics of rapid update rate, diversified data formats, multiple formats conversion, and so forth, we can conclude that although WAMS data have high precision and good synchronization performance, its characteristics, high density, high processing rate, and diversified structure, make the processing of wide-area measurement data become the key and difficult point in application of LFO analysis. Conventional linear data store model is difficult to meet the
analysis requirements of continuous and reliable data access. Therefore, a new storage model needs to be constructed to process high density, multiple concurrency, and real-time response WAMS big data.

4. Double-Set Elastic Store Model

Online LFO analysis system, as Figure 1 shows, is established in which WAMS provides wide area measurement data source, having a data processing scheme of Single-Input Multi-Output (SIMO), as Figure 4 presents. Data analysis system and WAMS data concentrator establish connection interface channel as is required by the standard [24]. The power system dynamic data set uploaded to master station is forwarded to the data analysis system in 10 ms at highest rate. There are two difficulties in processing the data. First, data loss and deadlock in data processing must be avoided. If deadlock occurs in data processing, data loss will cause interruption of the analysis procedure. Second, no matter during input or output, simultaneous reading and writing of the same address of storage space will result in storage access conflict. Once the access conflict cannot be effectively avoided, it will cause the first fault and lead to malfunction of the system.

4.1. Structure of the Double-Set Elastic Store Model. The linear storage space of the WAMS data received by the analysis system is transformed to plane storage structure, as Figure 5 shows. The vertical space is segmented into multiple bulk storage areas under dynamic management, labeled as $y_i$, which points to a power system substation group set. One or multiple $y_i$ are corresponding to one WAMS data time instant. Their total space should be able to store at least the quantity as much as received whole power system dynamic data at one time instant. In horizontal direction, dynamical segmentation of the storage space into some spaces labeled as $x_i$ is carried out again, which is corresponding to the called dynamic data set of the PMU devices in one plant station. The value of $x_i$ needs to accommodate all the variables of the PMU devices of one substation. The storage space that $y_i$ and $x_i$ point to is virtual until PMU data really arrives; that is, storage space is opened up dynamically only on the arrival of data.

$x_i$ has two limits: allowable value $x_{\text{allowable}}$ and maximum value $x_{\text{max}}$. The allowable value $x_{\text{allowable}}$ needs to comply with the requirement of data synchronization at PMU measuring points. Within allowable time delay, communication delay, and conversion as well as link layer and network layer retransmission considered, PMU data is not able to be well synchronous and ordered. Therefore, there must be data buffer at measuring points, caching data and then synchronizing and ordering data in buffer zone before data loading. If $x_i$ reaches $x_{\text{allowable}}$, $y_i$ in vertical direction is ready to be changed. But whether or not $y_i$ needs to be changed depends also on the state, which is reading or writing. If $y_i$ is ready to be changed is in the state of reading, PMU data receiving and saving need to be continued in horizontal direction.

The maximum value $x_{\text{max}}$ is related to the maximum allowable time that waiting for $x_{\text{max}}$ to change consumes. There is the maximum time of waiting for $y_i$ to change between $x_{\text{allowable}}$ and $x_{\text{max}}$. If the received plant station number already reaches $x_{\text{max}}$ whereas $y_i$ is still not changed, the bulk storage area in horizontal direction is deemed to be too small and the maximum value of $y_i$ needs adjustment. Nevertheless, if $x_{\text{allowable}}$ and $x_{\text{max}}$ are made equal, the flexible mechanism of waiting for $y_i$ to change is lost, which leads to the unfavorable outcome of increasing the probability of concurrent access conflict.

The data capacity managed by one $x_i$ depends on the number of PMU devices, the number of PMU access channels, and the quantity of data sent by PMU to master station. Different $x_i$ may correspond to different space capacity, and hence the storage space occupied by one $y_i$ may be different. This is similar to spring, capable of expansion and contraction.

The abovementioned storage model can realize the management of the substation group collection in the vertical direction and can realize the management of the PMU sending data collection with the substation as the unit in the horizontal direction. Furthermore, the storage spaces in vertical and horizontal directions are capable of elastic extension and dynamic contraction. Therefore, the model is named as the double-set elastic store model.

4.2. Cyclic and Saltatory Working Principle. When $x_{\text{max}} \rightarrow \infty$ or $y_i \rightarrow \infty$, the data saved by double-set elastic store model will require a lot of memory over a long period of time, leading to extensive hardware consumption and finally resource exhaustion. To avoid this risk, the model adopts the cyclic and saltatory working mechanism.

If there is enough storage space to store multiple time instants data of power system substation group set, cyclic storage is carried out as Figure 6 presents. In Figure 6, each spot represents a $y_i$ and a cube corresponding to $y_i$ represents a $x_i$. The spot is a structure to manage the queue of $x_i$, including the queue state, $x_{\text{allowable}}$, $x_{\text{max}}$, $x_1$ final value, $x_1$ packet establishing time, and mutex semaphores. The spots of $y_i$ are distributed on an elliptical orbit, driven in cyclical motion by the PMU data blocks continuously received through channel. Each $y_i$ has 4 states, that is, idle, writing, postwrite, and reading, with only one state at a time instant. The access to space of data reading and writing is decoupled in plane storage structure. The accessing operation depends on the state and mutex semaphores. If $y_{\text{max}}$ is reached through traversing, the model will return to $y_1$ and search for available storage position.

There is concurrency in data reading and writing. Therefore, to avoid deadlock in storage procedure due to multiple concurrency access to the same address, saltatory
addressing model is adopted as Figure 7 presents. On the arrival of new PMU dynamic data, its storage space position searching is not always in the order of \( \{y_1, y_2, \ldots, y_{\text{max}}\} \). The space position to accommodate new data needs to switch from \( y_j \) to \( y_k \), as data receiving has made \( x_{\text{max}} \) reached in the space of \( y_j \); whereas at the same time the space \( y_k \) is being read by an application module. In this situation, if there is no new space position, the valuable data will be dropped to keep the system operating. The method of resolution is saltatory addressing storage. \( y_{k+1} \) can be first tried in the situation; the new data can be stored in \( y_{k+1} \). If \( y_{k+1} \) happens to be in idle state; otherwise, \( y_{k+2} \) is tried in the same way. The problem saltatory addressing brings is mismatch between the storage position order and
To avoid the data accessing disorder caused by the mismatch in structure to manage the queue of \( x_i \), there are variables of the \( x_i \) packet establishing time, by which the data writing order of \( y_k \) and \( y_{k+1} \) can be distinguished accurately.

4.3. Comparison of Performance of Models. To test performance of the double-set elastic store model, it is compared with the conventional linear data store model.

Taking a province WAMS data as example, data frame is 62 592 bytes, and configuration frame is 191 722 bytes. The number of PMUs is 154, the number of phases is 5 445, the number of analogs is 3 452, and the number of switch signals is 533.

The two-store model receives 5000, 10000, and 100000 pieces of data at the frequencies of 10 ms, 20 ms, and 40 ms, respectively. The maximum tolerated transmission times are 1 000 s, 2 000, and 4 000 s of transmitting 10 000 pieces of data at the frequencies of 10 ms, 20 ms, and 40 ms, respectively. The comparisons of the tested performance of the two models as 100 000 pieces of data are received are shown in Tables 2 and 3.

As can be seen in Tables 2 and 3, the double-set elastic store model is able to finish data transmission without frame dropping within limited time as uploading frequency increases. Meanwhile, although the conventional linear data store model has no frame dropping when data interval is long, it may drop data to keep continuous operation as uploading frequency increases due to structure flaw, which is not tolerated in online analysis. The double-set elastic store model requires a little more memory than the conventional model, which is, however, negligible considering rich available hardware resource.

The new double-set elastic store model of WAMS data aims at dealing with high rate and high-density data and has advantages of high reliability, no multiple concurrency access conflict, and highly elastic interface to applications. It has already been applied in multiple LFO analysis systems of provincial power grids to guarantee data transmission reliability of WAMS big data under 10 ms data interval and to provide technical support of LFO analysis accuracy.

5. Conclusions

Due to exploration and grid connection of massive new energy, commission and operation of DC power transmissions of different voltage levels, and wide application of power electronic devices, power system operation characteristics become more complicated and more difficult to control. LFO problem is still among the difficulties that power system operation and control have to face. The principles of two types of LFO online calculation and analysis are studied in this paper, which are forced and weak damping oscillation. The important roles that wide area measurements data plays in PFO analysis key technologies are analyzed in depth. First, wide area measurements data is able to increase accuracy of online data integration and hence increase accuracy of control strategy calculation of weak damping enhancement. Second, it is able to extract...
power response trajectories in time sequence and rotor angle curves, which are used in oscillation model and modal identification to provide criterion of oscillation control strategies. Moreover, it can be used to calculate energy function for forced oscillation source and propagation path searching. To ensure accuracy and continuity of LFO online calculation and to meet the processing requirements of high rate, multiple concurrency, and high reliability, a new model of double-set elastic store is proposed, transforming storage space linear model to plane model, adopting cyclic and salutatory storage mechanism, and hence successfully solving the problems in continuous and reliable access of the wide area measurements data which is dense and of large scale and has quick update rate. The new model provides technical support of accuracy and robustness of LFO online analysis and is totally applicable in other wide area measurements data applications such as power system rapid operation monitoring and security early warning and control.

**Data Availability**

The processed data required to reproduce the findings in this paper cannot be shared at this time as the data also form part of an ongoing study.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**References**


