IEEE Access
SPECIAL SECTION ON INTELLIGENT DATA SENSING, COLLECTION AND DISSEMINATION IN MOBILE COMPUTING

Received November 13, 2019, accepted November 20, 2019, date of publication December 2, 2019, date of current version December 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2957181

Review of Power Spatio-Temporal Big Data Technologies for Mobile Computing in Smart Grid

YING MA[1,4], (Member, IEEE), CHAO HUANG[2], YU SUN[3], (Member, IEEE), GUANG ZHAO[2], AND YUNJIE LEI[1]

[1]School of Computer and Information Engineering, Xiamen University of Technology, Xiamen 361024, China
[2]Xiamen Great Power Geo Information Technology Company Ltd., Xiamen 361008, China
[3]Department of Education and Learning Technology, National Tsing Hua University, Hsinchu 30013, Taiwan
[4]Key Laboratory of Data Mining and Intelligent Recommendation, Fujian Province University, Xiamen 361024, China

Corresponding author: Chao Huang (10386057@qq.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61502404, in part by the Natural Science Foundation of Fujian Province of China under Grant 2019J01851, in part by the Distinguished Young Scholars Foundation of Fujian Educational Committee under Grant DYS201707, in part by the Xiamen Science and Technology Program under Grant 3502Z20183059, in part by the Open Fund of Engineering Research Center for Software Testing and Evaluation of Fujian Province, and in part by the Open Fund of Key Laboratory of Data mining and Intelligent Recommendation, Fujian Province University.

ABSTRACT Mobile computing adds computing, storage, processing and other functions in wireless network ends to provide customized and differentiated services, so that it can be widely applied in the fields of internet of things, video, medical treatment, retail and so on. Recently, power spatio-temporal big data (PSTBD) technology of smart grid based on mobile computing has experienced explosive growth. This paper emphasizes the specific requirements, technologies, applications, and challenges of the current PSTBD for mobile computing in smart grid. Based on current development status of PSTBD companies in representative countries in the world, we introduce PSTBD technology based on the characteristics of mobile computing based smart grid, and conduct a comprehensive investigation and analysis of relevant articles in this field. After comparing the differences between the traditional and the PSTBD based platform in the aspects of important features, platform goal, and platform architecture, we describe the key technologies and algorithms of the current PSTBD in detail. Then, based on the requirements of each link and field of the power grid, the typical application of PSTBD technology in various aspects of smart grid application based on mobile computing is discussed. Finally, the development direction and challenges of PSTBD are given. Through data analysis and technical discussion, it hopes to provide technical support and decision support for relevant practitioners in the PSTBD field.

INDEX TERMS Mobile computing, data processing, smart grids, spatio-temporal big data.

I. INTRODUCTION

Power spatio-temporal big data based on mobile computing refers to "power+mobile equipment+big data", which collects and processes multi-source, heterogeneous, multi-dimensional and multi-form spatio-temporal big data in various links from generation, transmission, transformation, distribution, power consumption to dispatching power production and power service. The characteristics of Power Spatio-Temporal Big Data (PSTBD) meet the "5V3E" characteristics [4]–[6], as shown in Fig.1. In addition to "3E" which is energy, exchange, and empathy, the "5V" is as follows.

Volume: Conventional power dispatching system includes hundreds of thousands of data collection points; the number of power distribution data centers often reaches tens of millions; data volume is often above TB and PB.

Velocity: Decision support requires analysis of large amounts of data in a fraction of a second; real-time processing requires continuous real-time data generation.

Variety: Data types are structured, semi-structured, and unstructured data, including real-time data, historical data, text data, multimedia data, time-series data and so on.

The associate editor coordinating the review of this manuscript and approving it for publication was Xuxun Liu.

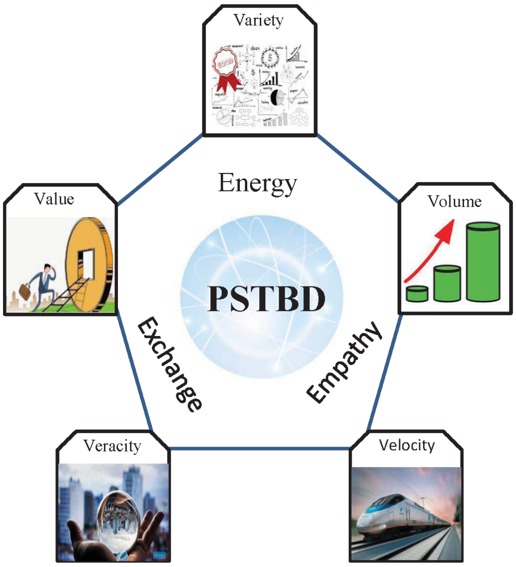


FIGURE 1. **Power patio-temporal big data features.**

Value: Electric power enterprises realize business trend analysis, prediction and decision support, through a series of means such as data mining.

Veracity: Due to the generation technologies, acquisition means and various forms, there are a large number of inferior data in power data.

Recently, PSTBD technologies for mobile computing are used in many engineering applications [7]–[9], involving user behavior analysis, demand response analysis, equipment risk analysis, system risk assessment, energy efficiency analysis, decision support, and other fields.

For example, C3 energy power grid is a large data analysis of time and space. It mainly serves the smart grids owners, operators and users on the supply side, such as utility companies, dispatching agencies, transmission and distribution companies. It can be used to reduce costs, predict and respond to system failures, and grasp the energy consumption of users in power grid operation. BG&E uses two application modules of the C3 energy analysis engine platform to integrate its internal 12 data source systems and data from 2 million smart meters in its service area. A total of 10 TB cloud image data, integrated analysis of 35 million pieces of data, about 8 GB/220 million pieces of data per day. Con Edison's distribution network fault risk assessment system was piloted at the New York Power Supply Company to provide fault risk assessments for feeders and equipment (cables, distribution transformers, etc.) to guide outages, improve facility maintenance efficiency, and improve distribution network reliability. PSE&G also uses advanced analysis of real-time sensors to track various operational metrics. The application of the analysis helps the company identify and remediate problems before they occur, saving millions of dollars in avoiding equipment failures.

Tokyo Electric Power and General Electric Company of the United States recently announced that to improve the power generation efficiency of thermal power plants, a system that uses the ''Internet of Things (IOT)'' will be jointly developed. The Fujitsu Thermal Power Plant of Tokyo Electric Power will introduce the industrial Internet of Things technology of General Electric Company of the United States. The two organizations will collect and analyze big data such as the operation of power generation equipment, and finally develop the corresponding system. NEC, Mitsubishi Electric IFS cloud computing service, which is a cloud computing service that manages and analyzes the data such as the operating status and energy consumption of equipment collected by factory automation control equipment in the manufacturing plant. Omron's sensing technology is combined with Oracle's information visualization software in Japan to provide services from the NSSOL cloud computing environment ''Absonne''. The power consumption is measured by various sensors located in factories, warehouses, office buildings, etc. Then, the visualization of the relationship between equipment operation status and power consumption is realized by combining the production scale and power consumption data in the production management system.

The French energy company owns Europe's largest electricity production system, mainly for transmission, distribution, gas supply and engineering consultancy, with stable markets in Italy and UK. At the same time, France Power actively expands its international business and participates in power projects in Asia, Africa, and America through wholly-owned or joint ventures. As a leading power infrastructure service provider in the world, the French Electric Power Company attaches great importance to the role of big data in enterprise operation analysis and management. By establishing professional institutions, improving data base and enhancing analysis ability, it constantly excavates the value of data assets to provide effective decision support for enterprise strategic transformation and service upgrading.

China's National Development and Reform Commission and the National Energy Administration issued the Guiding Opinions of the State Council on Actively Promoting the ''Internet+Action'', one of which is to promote the deep integration of energy and information and communication infrastructure. In 2017, the Energy Bureau issued the ''13th Five-Year Plan for Energy Development'', emphasizing more emphasis on system optimization and actively building a large grid intelligent monitoring system. During the ''Thirteenth Five-Year Plan'' period, it should actively promote the deep integration of new technologies in the fields of energy, information and big data, promote the efficient integration and intelligent regulation of grid information physics systems, and promote the construction of intelligent monitoring systems for large power grids. In 2018, it has formed a comprehensive and controllable big data and large grid intelligent monitoring system comprehensive solution, technical system and (international) standard specifications, environmental awareness, special communication, intelligent cloud, intelligent analysis, intelligent service, etc. in the intelligent monitoring of large power grids. It has

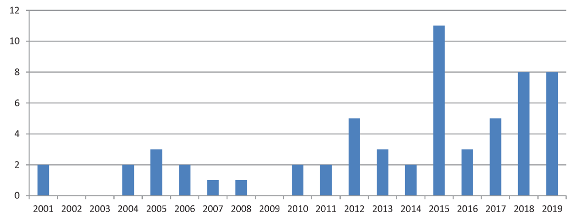


FIGURE 2. The number of articles published during 2001 and 2019.

broken through a number of key technologies and laid the foundation for building a giant energy intelligent management system that is information-driven, active defense, and precise control. Promote the construction of major engineering demonstration projects for large-scale intelligent monitoring robots based on big data and artificial intelligence technologies.

This paper also attempts to make a preliminary analysis and summary of the development opportunities and technical challenges of PSTBD Technologie from different perspectives. The rest of this paper is organized as follows. Section 2 investigates relevant articles and analysis the state of the power spatio-temporal big data technology research. After describes the PSTBD platform architecture associated with mobile edge computing in Section 3, algorithms for PSTBD have been reviewed in Section 4. The state of applications PSTBD technology are provided in Section 5. Section 6 finalizes the paper with challenges and the research directions of PSTBD technology.

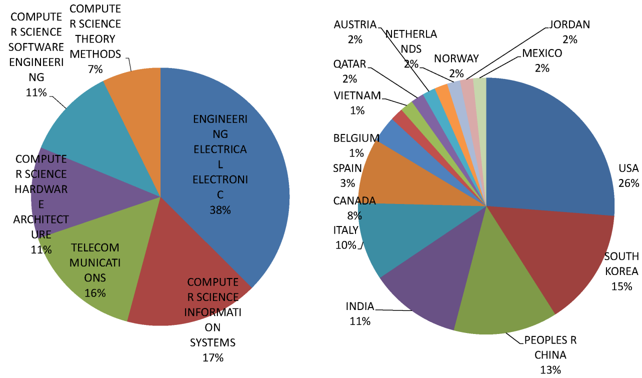
II. RELEVANT ARTICLES ANALYSIS

In order to investigate the existing research in mobile computing based power spatio-temporal big data, an article analysis was conducted on 1 September 2019 using the well-established and acknowledged databases, Web of Science (WoS). The query for WoS is as follows:

R1: TS=(power spatio temporal data and mobile computing); R2: TS=(power space time data and mobile computing); R3: TS=(space time data and mobile computing and power network); R4: TS=(space time data and mobile computing and power system); R5: TS=(space time data and mobile computing and power grid); R6: TS=(space time data and mobile computing and smart grid).

The final query is R= (R6 OR R5 OR R4 OR R3 OR R2 OR R1). The number of articles reached 58 from 2001 to 2019 in the query results. Fig.2 depicts the number of articles published about mobile computing based power spatio-temporal big data technology in these twenty years. We can see the number of publications has grown rapidly in the last five years.

We can see that the engineering electrical electronic is the largest group among all research directions, from the Fig.3(a). It can be seen from Fig.3(b) that the number of articles in USA, SOUTH KOREA, and PEOPLES R CHINA is 16, 9, and 8 respectively. It means the number of paper proposed by these three countries exceeds the sum of other countries.



(a) Direction of the articles (b) Articles from different countries

FIGURE 3. The analysis of the articles in WoS.

TABLE 1. The number of articles from different research organizations.

Research organization name	Number
PCSHE	3
UNIVERSITY OF WATERLOO	3
YONSEI UNIVERSITY	3
ANDHRA UNIVERSITY	2
CALIFORNIA STATE UNIVERSITY SYSTEM	2
ACADEMMIA NACIONAL DE CIENEIAS	2
COLLEGE OF ENGINEERING PUNE	2
ETRI	2
PENN STATE UNIVERSITY	2
SEOUL NATIONAL UNIVERSITY SNU	2
SUNGKYUNKWAN UNIVERSITY SKKU	2
UNIVERSITY OF CALIFORNIA SYSTEM	2

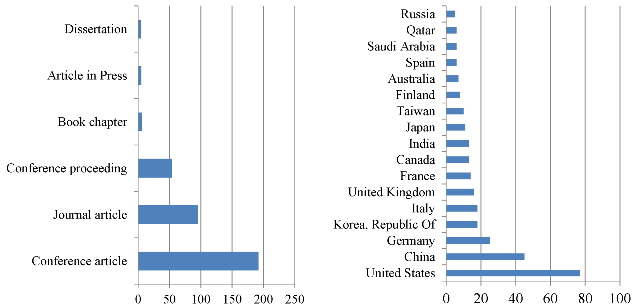
TABLE 2. Top 5 journals in the number of publication.

Source	Count
SENSORS	3
COMPUTERS ELECTRICAL ENGINEERING	2
FRONTIERS OF INFORMATION TECHNOLOGY ELECTRONIC ENGINEERING	2
IEEE TRANSACTIONS ON MOBILE COMPUTING	2
IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS	2

Research institutions in this field have published the most papers for PCSHE, UNIVERSITY OF WATERLOO, and YONSEI UNIVERSITY. All of these organizations published three articles, as shown in Table 1.

From Table 2, among the top 5 most popular journals in this field, SENSORS has published three papers. Although not a traditional journal, it is an interdisciplinary field, which makes this journal a natural outlet for mobile computing based PSTBD technologies.

We also extend analysis on the retrieval of Engineering Village library, and the final query is as follows: (((((mobile computing) WN KY) AND ((power system or smart grid or power grid or power network) WN KY)) AND ((space or spatio) WN KY)) AND ((time or temporal) WN KY)). 375 related papers were retrieved, and their types were shown in Fig.4(a). Unlike papers retrieved in WoS, most of the papers retrieved by IEEE are conference papers, 192 conference papers and 54 proceeding papers. According to the search results, the United States is still the country that publishes the most papers in this field, as shown in Fig.4(b).



(a) Types of these articles (b) Articles from different countries

FIGURE 4. The analysis of the articles in query results.

TABLE 3. Top 20 directions in the number of publication.

Controlled vocabulary	Count
Mobile Telecommunication Systems	67
Wireless Telecommunication Systems	51
Mobile Computing	41
Energy Efficiency	35
Digital Storage	32
Energy Utilization	27
Computer Simulation	27
Distributed Computer Systems	23
Communication Channels (Information Theory)	21
Real Time Systems	21
Bit Error Rate	21
Orthogonal Frequency Division Multiplexing	20
Ubiquitous Computing	20
Embedded Systems	19
Green Computing	18
Hardware	18
Data Acquisition	17
Data Communication Systems	17
Mobile Devices	16
Code Division Multiple Access	16

From the perspective of controlled vocabulary, these papers mainly come from the fields of mobile computing, communication, and data processing, as shown in Table 3.

III. POWER SPATIO-TEMPORAL BIG DATA PLATFORM

In view of the wide-area space-time measurement information and on-line security control of large power grid, the key is to quickly identify the dominant link and its state variables from the large spatial-temporal data according to the specific disturbance scenarios, and then to effectively assess the stability situation and calculate the prevention and control strategy for specific problems in real time. So it is necessary to study the PSTBD Technologies for mobile computing in smart grids, such as analysis, extraction, cleaning and so on.

A. BIG DATA ANALYSIS BASED ON MOBILE COMPUTING

The most important part of big data analysis is data mining algorithms, which can display the characteristics of data more scientifically and intuitively based on different data types and formats. For online security prevention and control of large power grids, data mining and machine learning methods are used to study the spatiotemporal correlation characteristics of large-scale measurement information of large power grids and efficient data mining algorithms.

Predictive analysis is the most common application of big data analysis. At present, the prediction problems in the power grid are often developed for a certain local electric quantity trajectory information. The analysis and utilization of the spatio-temporal correlation characteristics of the wide-area measurement trajectory information of the power grid is seriously insufficient, and it is difficult to grasp the overall movement behavior and change trend of the power grid. The research focuses on the prediction method of wide-area spatio-temporal big data. By establishing a scientific model, the spatio-temporal relationship and change law of the power grid are mined from the big data, so as to predict the future running trajectory and weak links of the power grid through the model, which is effective for the future stable situation of the power grid. Predict and adaptive prevention gains valuable time. It also helps to identify the main factors and prevention and control points (such as cutting machine or load shedding) in grid operation, and provides targets for reflection virtual modeling and adaptive defense control based on wide-area spatio-temporal information.

In addition, in order to improve the timeliness of online preprocessing, data mining and predictive analysis of big data running on the grid, it is necessary to study real-time big data analysis and management techniques, specifically related to mobile computing. And key technologies such as graph calculation and memory calculation are also crucially important for mobile computing based smart grids.

B. PLATFORM GOAL

Under the traditional security control mode of power systems, relying on the ‘‘modeling + simulation’’ model, the level of intelligence is not high, and it is restricted by parameters and models. The depth of mining wide-area measurement information is insufficient, and timeliness is difficult to guarantee. The above four aspects of research will provide basic theoretical support for the online intelligent security defense of a large grid based on a smart grid dispatching technical support system.

The main goal of establishing an information-driven model based on big data technology is to meet the efficient analysis and processing of large-scale spatio-temporal sequence data of large power grids, and to dynamically track the evolution of space-time sequences of power grids. At the same time, it combines the theory of machine learning and complex network to quantitatively evaluate the multi-dimensional spatiotemporal dynamic behavior of large power grids and adaptive wide-area coordinated control.

As shown in Fig.5, the two are combined to complement each other, which can further deepen the relationship between spatio-temporal sequence information and mobile computing characteristics of the smart grid.

C. PLATFORM ARCHITECTURE

Based on the wide-area space-time sequence data of the power grid, a data analysis platform centered on Spark is

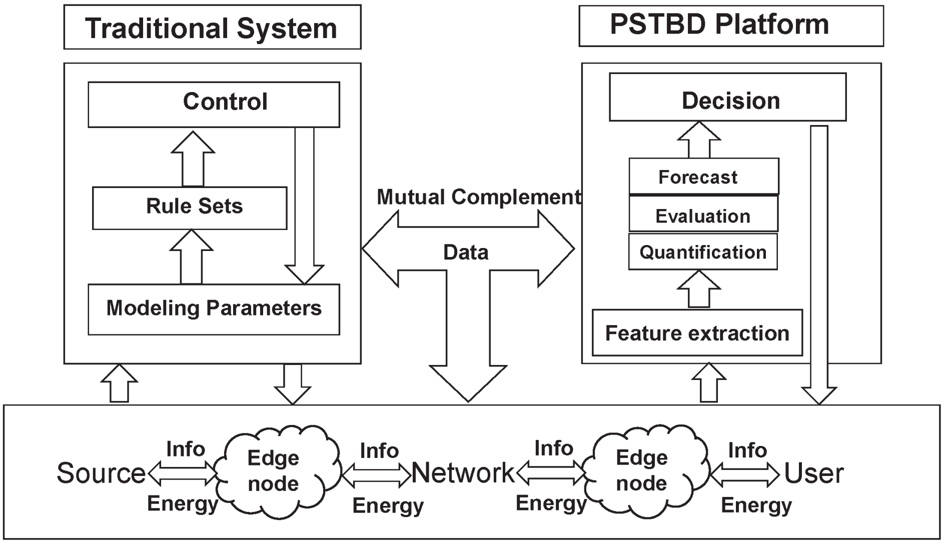


FIGURE 5. Platform construction purpose.

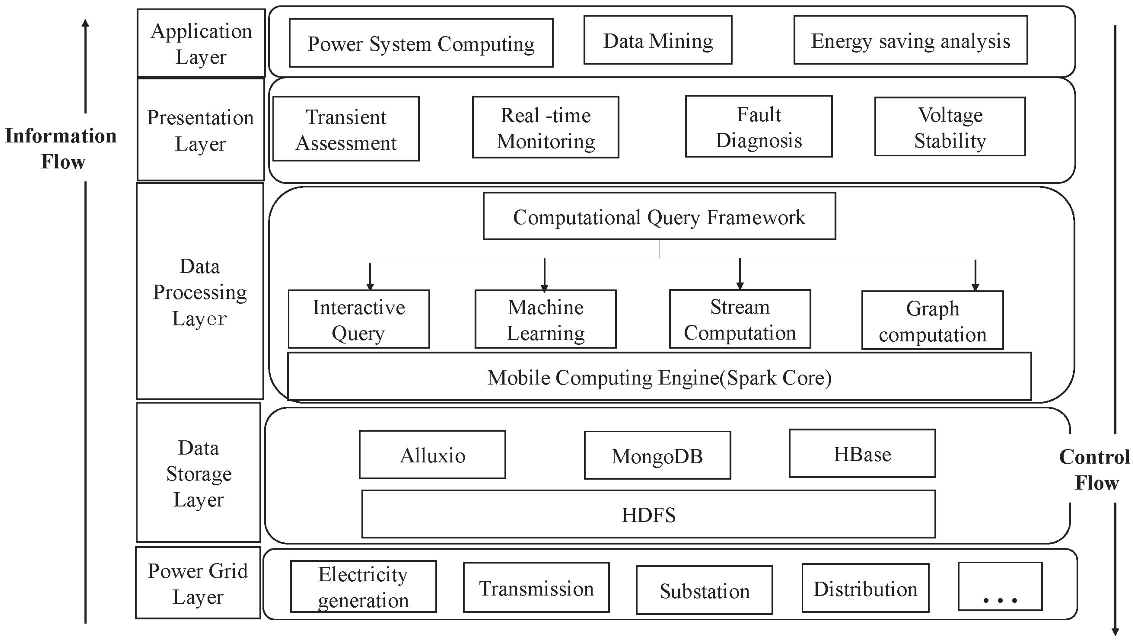


FIGURE 6. Mobile computing based big data platform.

constructed, as shown in Fig.6, which mainly includes the following levels.

The platform aims to build an information-driven grid information-physical coupling system, extract key features of the dominant grid operating state from complex information networks, and utilize information-physical interactions through computation, communication, and control technologies. And feedback, improves the intelligent real-time perception of the power grid and wide-area coordinated control capabilities to ensure the safe and stable operation of the power grid [10].

Application Layer: The application layer is an abstraction layer of communication methods designed for communication between processes and processes of the entire Internet computer network. Data application layer addition or improvement techniques includes data mining, etc., through online analysis, intelligent processing, and machine learning to achieve application goals.

Presentation Layer: Graphical display of calculation results, allowing dispatchers to visually identify the real-time operating status of the grid. In order to facilitate the front-end display, the analysis results of historical and real-time data are respectively written into different databases.

Data processing layer: The data processing layer is the core part of the platform. The computing and query framework mainly uses Spark Core and four components running on it:

MLlib, Spark SQL, Spark Streaming, GraphX. The computing framework adopts a unified programming model, which fully exploits the advantages of Spark. The Redis in-memory database acts as a data buffer pool to ease database load. The algorithm library is used to store commonly used high performance parallel algorithms, and the knowledge inventory is stored in the cognitive model obtained by machine learning training.

Data storage layer: In order to meet the functions of high-quality acquisition and integration of grid time-space sequence information, high-speed indexing, and storage of streaming data, automatic error detection, etc., the platform uses Hadoop distributed file system (HDFS) as the underlying distributed storage. The system, combined with a variety of NoSQL databases, provides powerful underlying support for large-scale mass data storage. As a distributed memory file system, Alluxio can store multiple user data in shared memory, avoiding a large number of disk I/O operations and improving data processing efficiency.

Power Grid layer: The grid can use the business insights obtained from the data processing layer for grid abnormal event monitoring and real-time decision making to control the physical layer of the grid in real time.

IV. ALGORITHMS FOR PSTBD

PSTBD based on mobile computing has its own characteristics and needs corresponding algorithms to solve the problems in the application. These algorithms mainly include: real-time control technology, temporal-GIS technology, data presentation technology, transmission and storage technology, parallel analysis technology, and other mobile computing related technologies.

A. REAL-TIME CONTROL TECHNOLOGY

With the deepening of the interconnection degree of the power grid and the increasingly complicated operation mode at present, the real-time analysis and control of the grid state become more and more critical [11]. Wide area measurement system (WAMS) uses synchronous phase angle measurement technology to realize real-time and high-speed acquisition of synchronous phase angle and main data of power network by gradually laying out synchronous phase angle measurement unit (PMU) of key measuring points in the whole network. PMU can ensure the synchronization of the whole network data through the global positioning system (GPS) [12]. Timing information and data are stored and sent to the main station at the same time. Therefore, the scheduler in WAMS can monitor the dynamic process of the power grid in real-time. Wide area measurement system coverage has also theoretically realized considerable real-time monitoring of the whole network, making real-time stability analysis and control of the power grid possible [13]. In practical engineering applications, due to uncertainties, stray components may appear in the WAMS system, which affects the accuracy of identification results of low-frequency oscillation. In order to improve the accuracy of the results, it is generally necessary to pre-process the read data. Then, the order estimation of the identification model, Prony analysis and comprehensive analysis are carried out. The amplitude, frequency and damping ratio of power oscillation are obtained, and then the small disturbance stability of power system based on WAMS online data is obtained [14].

Recently, most of the latest systems in this aspect adopt (GPS) synchronous clock technology to conduct stable analysis and control of PMU phase difference data according to hyperplane clustering method based on big data analysis. [15].

B. TEMPORAL-GIS TECHNOLOGY

The development of the power system makes the power grid increasingly dense, the power equipment increasingly complex, the load and building information constantly changing, the requirement for high electricity quality increasing, and other problems are becoming more prominent. Therefore, power system planning, operation, and business departments must collect, store, analyze, and quickly process their large and complex information including graphics, maps, and data. In order to realize the rational planning of transmission network construction and distribution network transformation, it is necessary to improve the quality of electric energy and power supply reliability [16].

The GIS (geographic information system) is an integrated information management system for the electric power production process [17]. It connects power equipment, substations, transmission and distribution networks, power units and power companies. The unified system provides information on natural equipment such as power equipment, grid operation, power technology, production management, electricity market, mountain rivers, topographic town roads, and meteorological, hydrological, geological, and resource sources. Through GIS, you can query related data, pictures, charts, maps, management information, etc.

The temporal geographic information system quickly reflects the historical changes of the power grid topology through time playback, and compares and analyzes the differences between the two historical nodes in different periods [18]. The emergence of temporal geographic information systems will play an important role in grid operation management and planning prediction, and will enhance the functional positioning of grid equipment and topological network data as ‘‘digital’’ in grid information systems.

Spatial objects have three basic properties: spatial properties, topic properties, and time properties. Traditional geographic information systems only focus on spatial and topic attributes when dealing with spatial objects. Therefore, the traditional geographic information system can only process the current state of the data or a snapshot of a certain historical state, can not record the dynamic changes of the data with time, and can not perform time series analysis and time processing [19]. Changes in geographic objects are closely related to time. In order to track and record these changes, it is necessary to introduce time information into the GIS, which is the main problem to be solved in the temporal GIS.

The tense grid data in the GIS uses the temporal and multi-level version of the ground state correction model to support the full lifecycle management of the grid model, meeting the historical evolution of the recorded grid model and the planning and design application requirements [20]. The performance requirements for fast access to historical data are in the form of monthly baseline and engineering delta. That is, on the 1st of each month, the current state of the grid resources is stored as a full snapshot, as the baseline for the month; then the variable data generated by each project is incrementally stored, so that the horizontal query of the section and the vertical query of the device history can be realized.

C. DATA PRESENTATION TECHNOLOGY

Data presentation technology refers to a technology that uses graphical means to effectively and clearly express and exchange information. Since the massive data sets formed by various systems in the smart grids contain various multivariate data, time-varying data, and high-precision high-resolution data, even a common data set can reach the order of TB. Therefore, facing massive smart grids big data, how to quickly and effectively extract useful information from users, how to present them to users through a limited screen in an intuitive and easy to understand way is a key technical difficulty in smart grid big data applications. Smart grid big data presentation technology can significantly improve the visibility and intuitiveness of power data, which helps managers to understand the operating status of the power system more intuitively, clearly and accurately.

Visual representation of spatiotemporal data can be divided into static visualization and dynamic visualization. The static visualization of spatio-temporal data is generally based on superimposing elements that can describe time changes on a two-dimensional map, and then describes the temporal and spatial attribute data and the variation characteristics in the spatial range [21]. Dynamic visualization can display spatiotemporal data by various means such as dynamic map and 3D GIS [22]. The spatio-temporal data is presented in a dynamically changing map or three-dimensional scene, which can visually and vividly represent the changing process of various spatial information. The development of 3D GIS enables spatio-temporal data to be presented in 3D geospatial. The spatial object based on spatiotemporal data intuitively expresses the motion process of spatial scale in the three-dimensional world, and solves the key problem that the planar map is difficult to express the motion trajectory of objects with different spatial elevations. A dynamic map is an electronic map that can reflect the changing process of spatio-temporal information centrally and intuitively. Its production and development is an important basis for visualizing time and space data [23]. The development of data visualization technology makes dynamic data rendering more and more powerful. EChart and D3 visual frameworks have also been used for reference in the field of spatio-temporal data representation [24].

Visualization technology is widely used in smart grids to monitor and control the operation of the grid in real time, which can effectively improve the automation level of the power system [25]. Spatial information flow display technology is usually embodied in the fusion of grid parameters and existing GIS, such as three-dimensional display technology and virtual reality technology. Historical flow display technology is often applied to grid historical data management and display. It can realize real-time monitoring data of power production site or forecasting function of data trends such as power grid planning and load forecasting data. It can be seen that this technology has great application value.

D. TRANSMISSION AND STORAGE TECHNOLOGY

With the development of computer technology, the storage and management mode of spatio-temporal data is constantly changing and updating. At each stage, it is affected by the current status of computer hardware and software, the characteristics of data scale and the actual application requirements [26]. The spatio-temporal data storage and management model has evolved from a traditional centralized file storage/space database to a distributed file system management represented by Hadoop HDF, and then to a distributed NoSQL non-relational database.

The amount of data generated by the power grid system during normal operation is extremely large. In order to facilitate the monitoring and management of the staff, it is necessary to store the data, which utilizes big data transmission and storage technology. The trend of contemporary big data is the storage of massive data and the existence of data for more and more things. By combing the key technologies of big data processing, the types of existing storage technologies and databases in the context of big data are summarized. In addition, in the context of the big data era, the research of grid big data and big data related technologies provide effective support for the modernization and information of grid work.

In terms of data storage, smart grids generally adopt distributed file systems to access power big data, such as Hadoop's HDFS storage system. However, these systems can only be used to store data but cannot meet the real-time requirements of the field. In order to meet real-time requirements, big data management and processing technologies need to focus on complex structure, semi-structured and unstructured data. Sadiq et al. mainly solved several key problems of big data storage, representation, processing, reliability and effective transmission [27]. Cai et al. Developed a reliable distributed file system (DFS), energy-optimized storage, compute integration into storage, big data duplication and efficient low-cost big data storage technology [28]. And it can break through distributed non-relational big data management and processing technology, data fusion technology [29], data organization technology [30], [31], research big data modeling technology [32]; breakthrough big data index, backup and other technologies [33].

E. PARALLEL ANALYSIS TECHNOLOGY

Compared with traditional algorithms, data-driven big data analysis enables us to analyze the interrelationships and interactions of all elements in the system, which is regarded as the correlation of high-dimensional parameter representation [34]–[36].

Smart grid operation has independent resources and decentralized control functions. Aggregating distributed data sources into centralized mining sites is a systematic prohibition due to potential transmission costs and privacy issues [37]. Although model-based mining can be performed at each distributed site, the decoupling process of the connection source is closely related to the simplification and assumptions. As a result, the results are often almost unsatisfactory and lead to biased views and decisions.

Load forecasting issues involve intelligent ML solutions related to the generation of large numbers of data sources, making them ideal for MapReduce deployments. The proposed framework can minimize data movement, thereby reducing the occurrence of network contention, which will ultimately improve the efficiency of the cloud system. The two main mechanisms that must be explicitly considered are:

- Data location-aware scheduling algorithm;
- Application-specific resource allocation mechanism.

In particular, tasks that require a common data set are dispatched to computers (computing nodes) that are very close to these data sets. For most data processing applications, storage capacity (such as disk and storage) is more important than computing power (ie, CPU).

In order to make the MapReduce parallel programming model easier to use, there are a variety of big data processing advanced query languages, such as Facebook's Hive [38], Yahoo's Pig [39], song Sawzall [40] and so on. These high-level query languages parse the query into a series of MapReduce jobs through a parser and execute them in parallel on a distributed file system. Compared with the basic MapReduce system, the high-level query language is more suitable for users to perform parallel processing of large-scale data [41].

F. EFFICIENT CALCULATION TECHNOLOGY

The key issues that need to be solved in efficient time-space big data calculation include how to use efficient data in the database; how to use the heterogeneous computing resources in the system to coordinate computing; how to realize the graphical customization of tools. Temporal efficient data for the IO, the data-model, combined with node traffic interface developed a data communication module, the details of the mask data distributed storage, the new class package, provides operator interface similar data goal adaptive platform with local data to hide data details. Through data multi-copy parallel read and write, calculation, transmission and data read and write overlap, data read-ahead and other technologies to achieve rapid read and write response data in a distributed environment [42]–[44]. Hidden internal hardware, data and parallel implementation details, according to parallel, master-slave parallel, pipeline parallel, work pool parallel and divide and conquer parallelism and other parallel programming modes provide five different programming templates, developers On the basis of the inherent multi-level concurrency of the mining algorithm, the appropriate programming template is selected to realize the fast parallelization of the processing algorithm. According to a certain functional granularity, the software is divided into a series of geographical space-time big data tools. The tools are loosely coupled, can independently complete each operation, and can also complete a complex remote sensing processing and calculation through loose integration. The tool developer implements the interface predetermined by the tool, and stores the metadata of the tool into the database of the integrated platform. The integration platform automatically generates the interface and the control of the tool by obtaining the meta information of the tool. Strictly define the input, parameters, and output format of the tool so that the tools can be correlated with each other through the output and input, and the integration of the data stream constitutes a processing pipeline. In the specific implementation of the tool customization, the user constructs the tool flow graphically, and automatically generates the tool flow script in the background. The integrated platform automatically invokes the corresponding tool according to the script and the data information provided by the database.

G. DATA FUSION TECHNOLOGY

In the future, smart grids require multiple links such as power generation, transmission, substation, power distribution, power consumption, and dispatching to achieve comprehensive information collection, smooth transmission, and efficient processing, and support the high integration of power flow, information flow, and business flow. For the efficient management of high-frequency, real-time, multi-source and heterogeneous allocation of large-scale data, there is currently no accurate definition of the integrated information model, metadata specification and unified conversion format for heterogeneous data [45]. It is impossible to allocate electricity big data. For efficient integration and integration, In order to provide smart grid intensive resource allocation, some researchers can expand multi-source heterogeneous data fusion technology [46]–[48].

Power big data is characterized by massive multi-source heterogeneity and industry-specific data access, so it is necessary to study scalable, efficient and reliable power big data management technology. The technical mapping of REST architecture-style Web Service as a CIS interface will greatly improve performance and efficiency, opening up a new dimension for power system information integration. path. Faced with the reality of increased grid infrastructure and local distribution, how to efficiently manage these infrastructures and process heterogeneous data and reduce grid costs will be a huge challenge.

H. MOBILE COMPUTING TECHNOLOGY

Real-time computing: the extensive application of mobile computing in the power Internet of things makes the real-time computing of large-scale power sensors and mobile device data become the key of the system [38]. At present, the real-time processing technology of spatio-temporal big data is mainly applied in the fields of smart city, smart transportation and smart, and medical treatment. In the process of power system information processing, there is a strong response decision and a short time delay, which can meet the sensor terminal real-time input. On the other hand, the data and instructions collected by the power sensor terminal must reach the mobile processing terminal in real-time and make corresponding decisions. The real-time information loss caused by decision information is also immeasurable, so the real-time calculation of spatio-temporal big data of power is particularly important.

Batch computing: batch computing in the power Internet of things refers to batch processing of static massive spatio-temporal data of power. In other words, data should be put in place before the calculation, and data should be mined and the business model should be verified. Apache Hadoop was originally used as a platform for analyzing spatio-temporal data for smart grids. In this model, large power spatio-temporal data sets can be divided into many small data sets for processing on the corresponding microdata units. A simple example is the analysis of the minimum reading of static data from all smart meters obtained using a map-reduce job in a map task. However, since map-reduce is designed for conventional batch processing, it is not suitable for frequent reprocessing of large dynamic data sets and cannot be used for real-time sensor data and stream data processing. Apache Spark is another generic clustered computing platform that provides flexibility, scalability, and speed to meet the challenges of large data in a smart grid. In addition to the usual batch processing, Apache spark is also capable of performing iteration and flow processing. Olston *et al.* [39] implemented the operation of processing Amazon data set using batch processing technology.

Flow computing: the value of power data decreases over time, so it is required to process new data or events as soon as they occur. Power network big data stream computing technology can mask the underlying details of data stream processing, with high-performance computing capability, time-efficient data stream online analysis capability, correlation analysis capability integrating multiple data sources, and resource management and deployment capability supporting flow processing. EALSOA(Extensible Advanced Large Scale Online Analytics) is an open source platform for mining big data streams. The platform can accomplish various tasks of data mining based on the distributed time-flow algorithm and provide developers with the secondary development of the original algorithm. At present, the amount of data obtained by the power grid increases exponentially, which makes the interactive processing information of the data server of power distribution scheduling management more and more. A large amount of real-time monitoring information is stored centrally before processing, which is easy to cause large processing delay or even message accumulation, which affects the scheduling processing efficiency. Research on relevant technologies of power big data distribution computing, memory computing, and flow computing is of great significance to break through the bottleneck of big data storage and computing in distribution network monitoring. Researchers have conceptualized and abstracted this setup in the flow model [51]. In this model, data runs at high speed, one instance at a time, and the algorithm must process it under very strict spatial and temporal constraints. Streaming algorithms can use probabilistic data structures to provide quick approximation answers.

V. STATE OF THE APPLICATIONS

The applications of power big data cover all aspects of power industry such as transmission, transmission, transformation, distribution, use, and regulation, as shown in Fig.7 in real-time energy management, power generation control center, transmission operation, state estimation, network physical system, network security defense, strategies, end-user real-time power demand, real-time price forecasting and other areas have very strong achievability. With the further advancement of smart grid construction, big data technology will play an increasingly important role in the smart grid. The following is a brief description of the application prospects of big data in smart grids through several typical application cases.

A. INTELLIGENT POWER DATA ANALYSIS

The bidirectional flow of power information in smart grid provides the possibility for power production and power operation to actively participate in the management and service of each link of power system. Diamantoulakis *et al.* [52] proposed to obtain information from spatio-temporal data of power, so as to facilitate management to make transaction decisions on smart grid. This method can improve the design of predictive ML algorithm by extracting load-carrying mode algorithm from large-scale data sets (k-means, ANN, etc.), compress data into low memory requirements, design scalable real-time performance and develop distributed computing architecture. Cao *et al.* [53] proposes a household power forecasting method based on deep load, which provides a good reference for household energy management system. In order to support real-time energy management, the authors in [54] intelligently process large amounts of power data in an efficient way. In the aspects of energy resources, load forecasting, disaster recovery, efficient data management, and analysis frameworks are critical to grid-optimized operations [55]–[57]. Public utilities use a variety of disaster recovery technologies to enhance customer engagement in power grid management [58], [59]. Through the big data obtained from smart meters and household equipment, public utilities can easily obtain real-time consumption information. It can

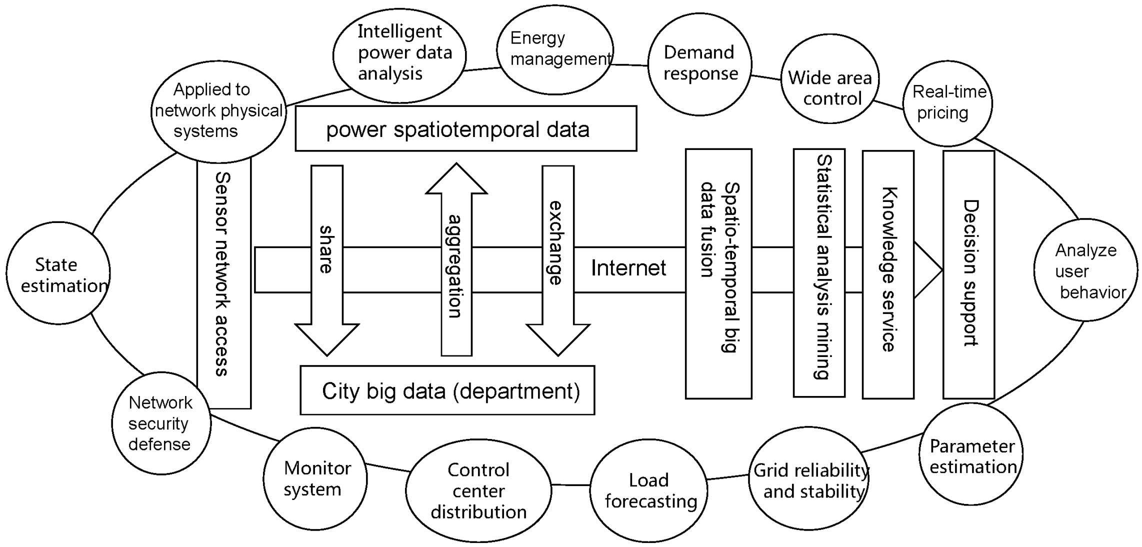


FIGURE 7. **The applications of power spatio-temporal big data based on mobile computing.**

also develop appropriate Incentive measures and operational strategies to better utilize power resources [60].

Spatio-temporal big data analysis can dynamically classify consumer behavior and power grid characteristics to help utility companies make better operational decisions [61], [62]. Chelmis *et al.* [63] developed a method of clustering customers using time series data to identify suitable customers. Yu *et al.* [64] using the big data for distribution power grid operation and planning, the importance of high-granularity load forecasting and customer consumer behavior modeling were determined. Disaster recovery in smart cities using big data has evolved in [65].

B. GRID RELIABILITY AND STABILITY

Data collected from social media are of great help for reliability prediction of power systems. For example, twitter data can be used to identify and locate specific blackout cases and regions [66]. Chen *et al.* [67] listed the importance of geographic information systems, global positioning system and meteorological data in power failure processing. The paper [68] USES more reliable voltage stability detection data than previous methods, which can be well used for power network monitoring. The data collected by the terminal device can also be used for stability prediction and real-time detection [69]. The big data method proposed in [70], [71] is helpful to improve the reliability and stability of power grid. In the article [72], an event detection application is developed by using the collected big data. Power spatio-temporal data mining can manage power transmission and monitor the motor. For example, [73] proposes a way to identify anomalies in power networks, which effectively utilizes data collected by intelligent devices.

C. WIDE AREA CONTROL

Based on the "stratification, zoning, local control" system scheme, a multi-level wide-area intelligent protection and control system is constructed. A time synchronization method based on Coordinated Universal Time and synchronous phasor measurement technology is proposed to achieve accurate time synchronization and real-time interaction between stations in the region. The proposed protection method combining wide-area current differential protection and comprehensive direction protection, wide-area adaptive standby power automatic switching technology and intelligent collective station protection, realizes fault intelligent judgment, system decision and self-healing control of multi-terminal and multi-component of regional power grid. The successful application of the two regional power grids in Duyun and Liupanshui in Guizhou Province proves the practicability and reliability of the developed system [74].

The power system stabilizer can not only solve the system oscillation problem, but also collect data to establish the linear model [75] and other optimal models [76]. In order to solve the problem of local optimization of a single model and improve the stability effect, the data acquisition and controller [77], [78] with higher accuracy can be designed to form a reliable closed loop of the whole power system.

D. STATE ESTIMATION

Any decision in the power system control, depends on the actual state of the system in real-time. The state of the power system is difficult to obtain, but the state variable can be predicted by using the method based on data estimation, and these predicted values can be obtained periodically [79], [80]. The accuracy and robustness of execution in the power real-time control process are very important, and some progress has been made in recent years. Although it is possible to develop reliable software for large power systems, the state estimation of power systems is still a problem of great concern to researchers [81]–[83].

The concept of modern power systems is subject to deregulated power market conditions and the impact of the opening of non-electric power companies. As a result, some fundamental business changes have been achieved through the new regulatory framework. These new developments have particularly affected the distribution network in which distributed power plants are beginning to frantically install, as well as ongoing demands for service quality and operational efficiency.

E. MONITOR SYSTEM

Spatio-temporal data visualization can be applied in many places, one of the important directions is advanced visualization. After the visualization of the smart grid, the overall evaluation level can be improved. If we want very accurate grid connection information and power system operation status to be obtained in real time, one of the methods is to apply visualization technology and then analyze. In general, there are multiple techniques for grid visualization, such as three-dimensional, two-dimensional, and single-line diagrams. However, there are also problems, because the number of variables is increasing and there are dependencies between variables, so advanced visualization technology must be used to realize the visualization of big data of power grid. By combining real-time monitoring, parallel coordinates and scatter plots, and the Andrew curve, the problem of increasing the number of variables and their dependencies will be solved. High-dimensional data visualization usually commercial tools such as real-time dynamic monitoring system [84]. Real-time dynamic detection systems have many visual options to choose from, including voltage amplitude diagrams, voltage Angle contour maps, dashboards for situational awareness, oscillation pattern diagrams, and frequency diagrams [85].

F. PARAMETER ESTIMATION

The key of power system control, operation and planning is power system state estimation and power system parameters. Evaluation is widely used in applications, such as flexible control design for network or physical attacks and trial system monitoring and operational resource planning [86], [87]. State estimation requires a large amount of data, and the smart grid framework is just right to provide it. Because large amounts of data generated by various smart devices and sensors can be captured using grids, the system will become transparent, enabling more accurate and perfect state estimation. However, due to the introduction of a large number of active nodes, the system will be affected by the power system optimization problem into nonlinear, mixed integer and non-convex [88]. At the same time, the power quality and reliability of power system can be improved by using prediction tools and real-time monitoring to optimize the use of resources [89]. Spatio-temporal big data can optimize the parameter estimation, take corrective measures for emergencies and analyze late emergencies. For example, the application of many mixed voltage regulation resources on feeders (such as solid-state transformers, voltage regulators, static synchronous compensators, on-load tap switches, intelligent inverters) is a trend of voltage/reactive regulation [90]–[93].

G. APPLIED TO NETWORK PHYSICAL SYSTEMS

Any physical and network vulnerability will have a significant impact on the smart grid, an important infrastructure. In the event of a natural disaster or system failure, traditional power personnel can perform emergency analysis of the grid to provide flexibility [94]. However, most of the existing power systems are not memory networks and are intentionally targeted at a large number of important components of the energy system. Cascading failures of the system can result from this structural attack. Therefore, it is necessary to achieve tight network-physical coupling, which can make the power network resistant to such physical and network attacks [95], [96]. Once big data is timely analyzed and identified such malicious attacks, the possibility of large losses can be greatly reduced [97].

H. NETWORK SECURITY DEFENSE

In the aspect of security, technology has been proposed for exploring network security defense strategies based on power system emergency sequencing systems [98]. Although there are cybersecurity defense standards, these methods are based primarily on expert opinions rather than systematic risk studies. ARCADES proposes a method based on graph-based network physical security model to identify improved network security defense strategy. The model is evaluated by the resistance distance metric and then prioritized according to the system's emergency analysis. This paper also proposes a technology to identify the most critical network security mechanisms to protect the grid. Planning and Operations: This paper focuses on two applications of grid security. For planning, a systematic approach was developed to verify the effectiveness of security policies and operations, prioritizing security mechanisms for auditing and monitoring.

I. LOAD FORECASTING

Load forecasting (LF) plays an important role in the planning and operation of power systems [99]. Future smart grids will leverage LF and dynamic pricing-based technologies for effective demand-side management (DSM). In this paper, a comprehensive comparative study of LF and dynamic pricing schemes in smart grid environments is conducted. Real-time pricing (RTP), time of use (ToU), and critical peak pricing (CPP) are discussed in detail. The two main categories of LF: computational models based on mathematics and artificial intelligence are described in subcategories. The mathematical models for effective DSM are discussed, including automatic recursion, moving average, automatic recursive moving average, automatic recursive integral moving average, exponential smoothing, iterative reweighting mean square, multiple regression, etc.

Electricity demand forecasting is extremely important for energy suppliers and other participants in power generation, transmission, distribution, and marketing. Accurate power load forecasting models are critical to the operation and planning of utility companies [100]. Load forecasting is extremely important for energy suppliers and other participants in power generation, transmission, distribution and marketing. This paper describes the power demand forecasting technology. Various types of methods and models are included in the literature. Load forecasting can be roughly divided into three categories: short-term forecasting, which is usually one hour to one week, and media forecasts are usually from one week to one year, and long-term predictions are more than one year. Based on the various types of research presented in these papers, load forecasting techniques can be divided into three broad categories: traditional forecasting techniques, improved traditional techniques, and soft computing techniques.

J. DEMAND RESPONSE

Considering the two difficulties in prediction, article [101] proposed a data-driven load reduction prediction method. The first problem is that the data is very sparse. Each customer receives only a few load reduction requests per year. Therefore, the k-nearest neighbor method that requires relatively small amounts of data is mainly used in our proposed method. The second difficulty is that each customer's characteristics are so different that a single forecasting method cannot cover all customers. Forecasting methods that provide significant predictive performance to a customer may provide poor performance to other customers. Therefore, the proposed prediction method uses a weighted integration model to apply different models to different customers. The forecast is based entirely on the history of power consumption data and demand response events, without any additional internal information from each customer. The actual data obtained from the demand response service provider verified that the proposed framework is suitable for predicting the load reduction for each customer. The amount of load reduction in the demand response is neither predetermined nor reliable, as evidenced by the text-especially in the liberalized energy market. In this case, a virtual power generation company consisting of many demand-responsive customers cannot provide a stable energy reduction. Therefore, the prediction of demand response should be carefully considered, because the amount of load reduction for each demand response customer is predicted based on past data.

The optimization of demand response will be based on the prediction of each customer's response. If the DRSP can know the future load reduction for each customer, the Demand Response Service Provider (DRSP) can select participating customers in the demand response event to avoid wasting resources and avoiding penalties. In addition, the prediction method can be applied to other energy applications, such as distributed energy sources, as they are likely to also feature small amounts of data and different characteristics of each resource. The proposed method can be used as one of several prediction blocks in an energy optimization system [102].

K. ENERGY MANAGEMENT

Energy system is what we have learned that one of the technology innovation and its important and complicated enough, the optimization of the improvement is the energy management, with the development of the industry, this way to gather more valued gaze, despite the current optimization on energy production and distribution with advanced technology and years of experience, but still can't meet the needs of today.

Hourly electricity price is usually used to adjust the hourly load level of a given user to improve the utility of the user. In practice, the utility of the user is not only determined by the hourly electricity price, but also affected by other factors, such as the lowest energy consumption level, the highest or lowest load level within the hour, etc. In [104], the model is proposed to maximize the utility of users, and a simple linear programming method is realized through a two-way communication device between power supply and users.

Due to the need to process huge data, the development of smart grid requires a series of big data analysis, management and monitoring technologies. The article [105] investigated HPC processing and focused on improving cost-effectiveness and security in the context of spatio-temporal data. They not only give a detailed explanation of the commonly used technical methods, but also focus on the prediction of electricity consumption, that is, the data of power supply use by smart electricity meter.

Researchers in the energy field has produced amazing results and accumulated a large amount of data. Zhang *et al.* [106] made an in-depth study of intelligent energy management, and their work is as follows: firstly, the source and characteristics of energy big data were thoroughly discussed, and the process model of managing intelligent energy driven by big data was proposed. The second step is to evaluate the big data analysis of intelligent energy management. The third step is to analyze the industry development of big data-driven intelligent energy management. Step four considers the current smart energy challenges, such as collection, governance, data integration, security, and privacy.

L. ANALYZING USER BEHAVIOR

In the power market, smart electricity meters, mobile devices, and sensors can acquire a large amount of data, and can analyze electricity consumption and other behaviors through load data mining. This paper [107] proposes a dynamic clustering method to analyze the consumption behavior of electricity consumption based on the consumption behavior and level of electricity customers in adjacent time periods. The specific operation of this method is to firstly use the clustering method to reduce the data, and then use Markov model to simulate the consumption dynamics. The key steps include obtaining the state transition probability matrix based on the power load data fitting curve. Finally, based on the behavior dynamics of typical users, real-time search algorithm and density estimation clustering technology are used to divide and conquer large-scale spatio-temporal data. Numerical experiments show that the model is efficient.

Based on the massive data generated by the intelligent community, researchers [108] studied the residential electricity behavior analysis model. Combining mobile computing and K-means algorithm, this model extracts multiple power features and calculates the feature weights by entropy weight method to complete the user classification tasks. Different from the previous methods, this algorithm is the first application in the field of residential users in the power industry. In the next step, the results of the user classification model will be studied to study the demand side response for different user groups.

M. CONTROL CENTER DISTRIBUTION

In recent years, the penetration of advanced sensor systems (such as voltage sensors, etc.) in power distribution systems has increased significantly [109]. The amount of data obtained by these sensors is also explosive, and new power big data technology is needed to explore the value of these complex data sets.

The future Intelligent Power System Control Center [110] presents the vision for next generation monitoring, analysis, and control functions. The article begins with a review of current control center technologies and then introduces the next-generation visual functions. The vision of the intelligent control center is expected to become an important part of the future intelligent transmission network. Future work will focus on expanding the smart grid monitoring modeling framework, [111] considering monitoring installation methods and cost constraints to get the best point for smart grid monitoring and reliability.

N. REAL-TIME PRICING

Real-time pricing method on the design of main power is based on demand response as the basis, take the large data of time and space technology, effective analysis of electricity and water load on the real-time electricity price changes before and after the implementation of the overall situation, residents and stand in the electric power company and industrial users of dimension analysis, compare the peak valley price and fixed price mechanism, from the power grid operation, peak peel, profit maximization and so on many Angle of the emphasis on real-time electricity price advantage. At present, smart devices can provide users with convenient data collection infrastructure. Others like Samadi *et al.* [107] have proposed a smart power terminal that allows several users to share the same energy. Each smart terminal is connected not only to the power grid, but also to the communication infrastructure, forming mobile edge nodes. It can also focus on the interaction between smart terminals and suppliers, exchanging information such as users' electricity bills and real-time prices.

VI. CHALLENGES AND RESEARCH DIRECTIONS

Spatio-temporal big data is a rapidly developing research field in academia and industry in recent years, including intelligent transportation, intelligent city, intelligent medical treatment and intelligent recommendation. However, the application of spatio-temporal big data technology in the power system based on mobile computing is still in its infancy and still faces many technical problems.

1) integration and storage of multi-source spatio-temporal data. The data format of the power system is highly heterogeneous, and there is an information barrier for data sharing, which makes it more difficult for multi-source power big data to be applied across systems and platforms. Therefore, it is necessary to generate standardized metadata and data normalization methods to solve the problem of heterogeneity in the multi-source integration process of power big data. The power industry basically uses centralized database to store data, which brings the problems of small storage capacity, difficult expansion and low loading efficiency, and cannot be applied to the mobile computing environment. Therefore, it is very important to study the lightweight data storage technology suitable for moving edge computing.

2) real-time data processing technology. In the power system based on mobile computing, real-time detection and analysis can be carried out on the equipment based on the historical operation of the equipment. In case of equipment deterioration, operation and maintenance personnel can be timely reminded of possible problems of the equipment and maintenance of the equipment in advance, so as to reduce the risk of sudden shutdown and economic loss. Especially in emergency applications, such as fault detection and transient oscillation detection, the response time requirements are up to the millisecond level. At present, the fast computing service of the cloud system is delayed by network congestion, and the research on real-time processing technology based on mobile computing will be the key to solve this problem.

3) reduction of power big data. Due to the large terminal resources consumed by mobile computing, the efficiency is low and insufficient. And the operation of the power system has remarkable parallelism, monitoring information, the traditional reduction method of non parallel due to its single mode of processing, the processing data set must be one-time into memory, the memory bottleneck problems caused by the reduction efficiency is low, does not apply to solve based on mobile computing power, the spatio-temporal data pretreatment needs to be introduced to adapt to the parallel processing method.

4) spatio-temporal data visualization technology. With the online access processing of real-time dynamic data such as power sensor networks, internet of things and social networks, the analysis based on mobile computing will become a typical feature of power multimode spatio-temporal data processing. However, the current analysis is mainly descriptive and diagnostic, while the future will be more based on spatio-temporal historical data, predictive and forward-looking real-time analysis, to predict the development trend and provide a

basis for decision-making. For example, how to effectively discover and express correlations or trends between multiple sources of data is a huge challenge. Other challenges include visualization algorithms, information extraction and rendering, and image synthesis techniques.

5) data privacy and security. In the power of mobile computing scenario, a new wireless network as the core of the mobile computing support technology promote the development of various new applications, however, the openness of the wireless network transmission medium resource limitation, mobile terminal and mobile network topology rapidly changing sex to power mobile computing to bring greater security risks. The growing use of smart meters for household energy consumption, for example, has created a growing amount of personal information. Because data is Shared between different entities, a private data leak can be a disaster and cause cascading problems. Security and privacy protection is more prominent, and have gradually become an important bottleneck restricting the development of mobile computing in the emerging scene, and become an important challenge in this field.

VII. CONCLUSION

With the vigorous development of mobile communication technology, it has brought unprecedented opportunities to various applications in the mobile computing environment. In this context, the study of power spatio-temporal big data (PSTBD) based on mobile computing has become a promising development direction of smart grid.

In general, this paper conducts research based on relevant literature, analyzes the research status of PSTBD technology, and introduces the PSTBD platform architecture related to mobile computing. We focus on PSTBD algorithms analysis including big data modeling, anomaly detection, power GIS temporalization, spatio-temporal big data, efficient calculation, heterogeneous multi-data source data fusion, data processing, related to PSTBD technologies. In addition, we investigate the application of PSTBD in real-time energy management, transmission operation, state estimation, network physical system, network security defense, real-time price forecasting and so on. At the same time, the future research challenges and research directions are presented. This review can be used as the basis for future research on spatio-temporal big data of power based on mobile computing.

REFERENCES

[1] E. Patti, A. L. A. Syrri, M. Jahn, P. Mancarella, A. Acquaviva, and E. Macii, "Distributed software infrastructure for general purpose services in smart grid," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 1156–1163, Mar. 2016.

[2] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, "Mobile edge computing: A survey," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 450–465, Feb. 2018.

[3] A. Asrari, S. Lotfifard, and M. Ansari, "Reconfiguration of smart distribution systems with time varying loads using parallel computing," *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2713–2723, Nov. 2016.

[4] S. Yin and O. Kaynak, "Big data for modern industry: Challenges and trends [point of view]," *Proc. IEEE*, vol. 103, no. 2, pp. 143–146, Feb. 2015.

[5] H. Hu, Y. Wen, T.-S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," *IEEE Access*, vol. 2, pp. 652–687, 2014.

[6] J. Zhan, J. Huang, L. Niu, X. Peng, D. Deng, and S. Cheng, "Study of the key technologies of electric power big data and its application prospects in smart grid," in *Proc. IEEE PES Asia–Pacific Power Energy Eng. Conf.*, Dec. 2014, pp. 1–4.

[7] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016.

[8] P. Mach and Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1628–1656, 3rd Quart., 2017.

[9] Y. Ma, C. Huang, Y. Sun, G. Zhao, and Y. Lei, "Review of power spatio-temporal big data technologies, applications, and challenges," in *Proc. Int. Conf. Secur., Privacy Anonymity Comput., Commun. Storage*, 2019, pp. 197–206.

[10] Y. Liu, Y. Peng, B. Wang, S. Yao, and Z. Liu, "Review on cyber-physical systems," *IEEE/CAA J. Automatica Sinica*, vol. 4, no. 1, pp. 27–40, Jan. 2017.

[11] A. Primadianto and C.-N. Lu, "A review on distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3875–3883, Sep. 2016.

[12] Y. Liu, Y. Zhang, J. Guo, and D. Zhou, *Wide Area Monitoring Through Synchrophasor Measurement*. Hoboken, NJ, USA: Wiley, 2016.

[13] A. G. Phadke and T. Bi, "Phasor measurement units, WAMS, and their applications in protection and control of power systems," *J. Mod. Power Syst. Clean Energy*, vol. 6, no. 4, pp. 619–629, 2018.

[14] H. Zhang, F. Shi, Y. Liu, and V. Terzija, "Adaptive Online disturbance location considering anisotropy of frequency propagation speeds," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 931–941, Mar. 2016.

[15] A. Xue, S. Leng, Y. Li, F. Xu, K. E. Martin, and J. Xu, "A novel method for screening the PMU phase angle difference data based on hyperplane clustering," *IEEE Access*, vol. 7, pp. 97177–97186, 2019.

[16] V. C. Gungor and M. K. Korkmaz, "Wireless link-quality estimation in smart grid environments," *Int. J. Distrib. Sensor Netw.*, vol. 8, no. 2, Feb. 2012, Art. no. 214068.

[17] Z. Liang, Q. Xue, and Y. Du, "Communication management system design of the smart grid based on the GIS," in *Proc. 4th Int. Conf. Machinery, Mater. Comput. Technol.*, 2016, pp. 318–320.

[18] K. K. Radhakrishnan, J. Moirangthem, S. K. Panda, and G. Amaratunga, "GIS integrated automation of a near real-time power-flow service for electrical grids," *IEEE Trans. Ind. Appl.*, vol. 54, no. 6, pp. 5661–5670, Nov./Dec. 2018.

[19] A. Slingsby, R. Beecham, and J. Wood, "Visual analysis of social networks in space and time using smartphone logs," *Pervasive Mobile Comput.*, vol. 9, no. 6, pp. 848–864, Dec. 2013.

[20] H. Daki, A. El Hannani, A. Aqqal, A. Haidine, and A. Dahbi, "Big Data management in smart grid: Concepts, requirements and implementation," *J. Big Data*, vol. 4, no. 1, Apr. 2017, Art. no. 13.

[21] M. Lu, E. Pebesma, A. Sanchez, and J. Verbesselt, "Spatio-temporal change detection from multidimensional arrays: Detecting deforestation from MODIS time series," *ISPRS J. Photogram. Remote Sens.*, vol. 117, pp. 227–236, Jul. 2016.

[22] P. Compieta, S. Di Martino, M. Bertolotto, F. Ferrucci, and T. Kechadi, "Exploratory spatio-temporal data mining and visualization," *J. Vis. Lang. Comput.*, vol. 18, no. 3, pp. 255–279, Jun. 2007.

[23] I. Idehen, B. Wang, K. Shetye, T. Overbye, and J. Weber, "Visualization of large-scale electric grid oscillation modes," in *Proc. IEEE North Amer. Power Symp. (NAPS)*, Sep. 2018, pp. 1–6.

[24] Y. Li, Z. Wang, and Y. Hao, "A hierarchical visualization analysis model of power big data," *IOP Conf., Earth Environ. Sci.*, vol. 108, no. 5, Jan. 2018, Art. no. 052064.

[25] C. N. Pitas, C. E. Tsirakis, E. T. Zotou, and A. D. Panagopoulos, "Emerging communication technologies and security challenges in a smart grid wireless ecosystem," *Int. J. Wireless Mobile Comput.*, vol. 7, no. 3, pp. 1–14, Jan. 2014.

[26] X. Liu, T. Qiu, and T. Wang, "Load-balanced data dissemination for wireless sensor networks: A nature-Inspired approach," *IEEE Internet Things J.*, to be published, doi: 10.1109/JIOT.2019.2900763.

[27] B. Sadiq, "A spatio-temporal multimedia big data framework for a large crowd," in *Proc. IEEE Int. Conf. Big Data*, Oct. 2015, pp. 2742–2751.

[28] H. Cai, B. Xu, L. Jiang, and A. V. Vasilakos, "Iot-based big data storage systems in cloud computing: Perspectives and challenges," *IEEE Internet Things J.*, vol. 4, no. 1, pp. 75–87, Jan. 2017.
[29] Y. Ma, G. Luo, X. Zeng, and A. Chen, "Transfer learning for cross-company software defect prediction," *Inf. Softw. Technol.*, vol. 54, no. 3, pp. 248–256, 2012.
[30] J. Gou, L. Wang, B. Hou, J. Lv, Y. Yuan, and Q. Mao, "Two-phase probabilistic collaborative representation-based classification," *Expert Syst. Appl.*, vol. 133, pp. 9–20, Nov. 2019.
[31] J. Gou, B. Hou, Y. Yuan, W. Ou, and S. Zeng, "A new discriminative collaborative representation-based classification method via l_2 regularizations," *Neural Comput. Appl.*, to be published, doi: 10.1007/s00521-019-04460-x.
[32] R. Y. Zhong, S. T. Newman, G. Q. Huang, and S. Lan, "Big Data for supply chain management in the service and manufacturing sectors: Challenges, opportunities, and future perspectives," *Comput. Ind. Eng.*, vol. 101, pp. 572–591, Nov. 2016.
[33] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, "Digital twin-driven product design, manufacturing and service with big data," *Int. J. Adv. Manuf. Technol.*, vol. 94, nos. 9–12, pp. 3563–3576, Feb. 2018.
[34] X. He, Q. Ai, R. C. Qiu, W. Huang, L. Piao, and H. Liu, "A big data architecture design for smart grids based on random matrix theory," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 674–686, Mar. 2017.
[35] C. Xiao, D. Han, Y. Ma, and Z. Qin, "CsiGAN: Robust channel state information-based activity recognition with GANs," *IEEE Internet Things J.*, to be published, doi: 10.1109/JIOT.2019.2936580.
[36] L. Hu, A. Liu, M. Xie, and T. Wang, "UAVs joint vehicles as data mules for fast codes dissemination for edge networking in smart city," *Peer-to-Peer Netw. Appl.*, vol. 12, no. 6, pp. 1550–1574, Nov. 2019, doi: 10.1007/s12083-019-00752-0.
[37] E. Hossain, I. Khan, F. Un-Noor, S. S. Sikander, and M. S. H. Sunny, "Application of big data and machine learning in smart grid, and associated security concerns: A review," *IEEE Access*, vol. 7, pp. 13960–13988, Jan. 2019.
[38] A. Thusoo, J. S. Sarma, N. Jain, Z. Shao, P. Chakka, S. Anthony, H. Liu, P. Wyckoff, and R. Murthy, "Hive: A warehousing solution over a map-reduce framework," *Proc. VLDB Endowment*, vol. 2, no. 2, pp. 1626–1629, Aug. 2009.
[39] C. Olston, B. Reed, U. Srivastava, R. Kumar, and A. Tomkins, "Pig latin: A not-so-foreign language for data processing," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, Vancouver, BC, Canada, Jun. 2008, pp. 1099–1110.
[40] P. Rob, "Interpreting the data: Parallel analysis with Sawzall," *Sci. Program.*, vol. 13, no. 4, pp. 277–298, 2005.
[41] W. Peng, M. Dan, Z. Jianfeng, and T. Bibo11, "Review of programming models for data-intensive computing," *J. Comput. Res. Develop.*, vol. 47, no. 11, pp. 1993–2002, 2010.
[42] X. Liu, T. Qiu, X. Zhou, T. Wang, L. Yang, and V. Chang, "Latency-aware anchor-point deployment for disconnected sensor networks with mobile sinks," *IEEE Trans. Ind. Informat.*, to be published, doi: 10.1109/TII.2019.2916300.
[43] Q. Li, A. Liu, and T. Wang, "Pipeline slot based fast rerouting scheme for delay optimization in duty cycle based M2M communications," *Peer-to-Peer Netw. Appl.*, vol. 12, no. 6, pp. 1673–1704, Nov. 2019, doi: 10.1007/s12083-019-00753-z.
[44] X. Liu and P. Zhang, "Data drainage: A novel load balancing strategy for wireless sensor networks," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 125–128, Jan. 2018.
[45] X. Liu, M. Zhao, A. Liu, and K. K. L. Wong, "Adjusting forwarder nodes and duty cycle using packet aggregation routing for body sensor networks," *Inf. Fusion*, vol. 53, pp. 183–195, Jan. 2020.
[46] A. Prahlad, P. Gokhale, and R. Kottomtharayil, "Data mining systems and methods for heterogeneous data sources," U.S. Patent 9 405 632. Aug. 2, 2016.
[47] C. Kong, M. Gao, and C. Xu, "Entity matching across multiple heterogeneous data sources," in *Proc. Int. Conf. Database Syst. Adv. Appl.* Cham, Switzerland: Springer, 2016, pp. 133–146.
[48] E. M. Eisman, M. Navarro, and J. L. Castro, "A multi-agent conversational system with heterogeneous data sources access," *Expert Syst. Appl.*, vol. 53, pp. 172–191, Jul. 2016.
[49] Y. Wang, Q. Chen, C. Kang, and Q. Xia, "Clustering of electricity consumption behavior dynamics toward big data applications," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2437–2447, Sep. 2016.
[50] M. Kiran, P. Murphy, I. Monga, J. Dugan, and S. S. Baveja, "Lambda architecture for cost-effective batch and speed big data processing," in *Proc. IEEE Int. Conf. Big Data*, Oct./Nov. 2015, pp. 2785–2792.
[51] B. R. Prasad and S. Agarwal, "Handling big data stream analytics using SAMOA framework-a practical experience," *Int. J. Database Theory Applicat*, vol. 7, no. 4, pp. 197–208, Aug. 2014.
[52] P. D. Diamantoulakis, V. M. Kapinas, and G. K. Karagiannidis, "Big data analytics for dynamic energy management in smart grids," *Big Data Res.*, vol. 2, pp. 94–101, Sep. 2015.
[53] Y. Cao, H. Song, O. Kaiwartya, B. Zhou, Y. Zhuang, Y. Cao, and X. Zhang, "Mobile edge computing for big-data-enabled electric vehicle charging," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 150–156, Mar. 2018.
[54] A. Yassine, S. Singh, and A. Alamri, "Mining human activity patterns from smart home big data for health care applications," *IEEE Access*, vol. 5, pp. 13131–13141, 2017.
[55] G. Derakhshan, H. A. Shayanfar, and A. Kazemi, "The optimization of demand response programs in smart grids," *Energy Policy*, vol. 94, pp. 295–306, Jul. 2016.
[56] S. Maharjan, Q. Zhu, Y. Zhang, S. Gjessing, and T. Basar, "Dependable demand response management in the smart grid: A Stackelberg game approach," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 120–132, Mar. 2013.
[57] M. Ghorbanian, S. H. Dolatabadi, and P. Siano, "Big data issues in smart grids: A survey," *IEEE Syst. J.*, vol. 13, no. 4, pp. 4158–4168, Dec. 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8809368, doi: 10.1109/JSYST.2019.2931879.
[58] J. Kim and Y. Dvorkin, "Enhancing distribution system resilience with mobile energy storage and microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 4996–5006, Sep. 2018.
[59] A. Ahadi, N. Ghadimi, and D. Mirabbasi, "An analytical methodology for assessment of smart monitoring impact on future electric power distribution system reliability," *Complexity*, vol. 21, no. 1, pp. 99–113, Sep./Oct. 2015.
[60] Z. Ai, Y. Liu, F. Song, and H. Zhang, "A smart collaborative charging algorithm for mobile power distribution in 5G networks," *IEEE Access*, no. 6, pp. 28668–28679, 2018.
[61] P. Zhang, F. Li, and N. Bhatt, "Next-generation monitoring, analysis, and control for the future smart control center," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 186–192, Sep. 2010.
[62] P. Järventausta, S. Repo, A. Rautiainen, and J. Partanen, "Smart grid power system control in distributed generation environment," *Annu. Rev. Control*, vol. 34, no. 2, pp. 277–286, Dec. 2010.
[63] C. Chelmis, J. Kolte, and V. K. Prasanna, "Big data analytics for demand response: Clustering over space and time," in *Proc. IEEE Int. Conf. Big Data*, Oct./Nov. 2015, pp. 2223–2232.
[64] N. Yu, S. Shah, R. Johnson, R. Sherick, M. Hong, and K. Loparo, "Big data analytics in power distribution systems," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Feb. 2015, pp. 1–5.
[65] A. Jindal, N. Kumar, and M. Singh, "A unified framework for big data acquisition, storage, and analytics for demand response management in smart cities," *Future Gener. Comput. Syst.*, to be published. [Online]. Available: https://sciencedirect.xilesou.top/science/article/pii/S0167739X17324780, doi: 10.1016/j.future.2018.02.039.
[66] H. Sun, Z. Wang, and J. Wang, Z. Huang, N. Carrington, and J. Liao, "Data-driven power outage detection by social sensors," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2516–2524, Sep. 2016.
[67] P. C. Chen, V. Malbasa, Y. Dong, and M. Kezunovic, "Sensitivity analysis of voltage sag based fault location with distributed generation," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2098–2106, Jul. 2015.
[68] M. Kezunovic, L. Xie, and S. Grijalva, "The role of big data in improving power system operation and protection," in *Proc. IEEE IREP Symp. Bulk Power Syst. Dyn. Control-IX Optim., Secur. Control Emerg. Power Grid*, Aug. 2013, pp. 1–9.
[69] J. Zhao, G. Zhang, K. Das, G. N. Korres, N. M. Manousakis, A. K. Sinha, and Z. He, "Power system real-time monitoring by using PMU-based robust state estimation method," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 300–309, Jan. 2016.
[70] M. U. Usman and M. O. Faruque, "Validation of a PMU-based fault location identification method for smart distribution network with photovoltaics using real-time data," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 21, pp. 5824–5833, Sep. 2018.
[71] H. Jiang, X. Dai, D. Gao, J. Zhang, Y. Zhang, and E. Muljadi, "Spatial-temporal synchrophasor data characterization and analytics in smart grid fault detection, identification, and impact causal analysis," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2525–2536, Sep. 2016.

[72] Y. Zhou, R. Arghandeh, and C. J. Spanos, "Partial knowledge data-driven event detection for power distribution networks," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5152–5162, Aug. 2018.
[73] R. Moghaddass and J. Wang, "A hierarchical framework for smart grid anomaly detection using large-scale smart meter data," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5820–5830, Nov. 2018.
[74] J. Bai, C. Gao, Y. Wang, W. Zhao, and J. Li, "Research and application of wide-area intelligent protection and control system based on the shared data source," *Power Syst. Protection Control*, vol. 44, no. 18, pp. 157–162, 2016. [Online]. Available: http://www.wanfangdata.com.cn/details/detail.do?_type=perio&id=jdq201618024
[75] H. Ni, G. T. Heydt, and L. Mili, "Power system stability agents using robust wide area control," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1123–1131, Nov. 2002.
[76] B. Chaudhuri, R. Majumder, and B. C. Pal, "Wide-area measurement-based stabilizing control of power system considering signal transmission delay," *IEEE Trans. Power Syst.*, vol. 19, no. 4, pp. 1971–1979, Nov. 2004.
[77] J. W. Stahlhut, T. J. Browne, G. T. Heydt, and V. Vittal, "Latency viewed as a stochastic process and its impact on wide area power system control signals," *IEEE Trans. Power Syst.*, vol. 23, no. 1, pp. 84–91, Feb. 2008.
[78] A. Heniche and I. Kamwa, "Assessment of two methods to select wide-area signals for power system damping control," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 572–581, May 2008.
[79] J. Zhao, G. Zhang, J. Y. Dong, and K. P. W. Davoudi, "Forecasting-aided imperfect false data injection attacks against power system nonlinear state estimation," *IEEE Trans. Smart Grid.*, vol. 7, no. 1, pp. 6–8, Jan. 2016.
[80] G. Wang, G. B. Giannakis, and J. Chen, "Robust and scalable power system state estimation via composite optimization," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6137–6147, Nov. 2019.
[81] E. Ghahremani and I. Kamwa, "Local and wide-area PMU-based decentralized dynamic state estimation in multi-machine power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 547–562, Jan. 2016.
[82] K. Sun, J. Qi, and W. Kang, "Power system observability and dynamic state estimation for stability monitoring using synchrophasor measurements," *Control Eng. Pract.*, vol. 53, pp. 160–172, Aug. 2016.
[83] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1630–1638, Jul. 2017.
[84] J. Kwac and R. Rajagopal, "Demand response targeting using big data analytics," in *Proc. IEEE Int. Conf. Big Data*, Oct. 2013, pp. 683–690.
[85] A. Agarwal, J. Balance, B. Bhargava, J. Dyer, K. Martin, and J. Ma, "Real time dynamics monitoring system (RTDMS) for use with synchrophasor technology in power systems," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2011, pp. 1–8.
[86] F. Capitanescu, J. L. M. Ramos, P. Panciatici, D. Kirschen, A. M. Marcolini, P. Platbrood, and L. Wehenkel, "State-of-the-art, challenges, and future trends in security constrained optimal power flow," *Electr. Power Syst. Res.*, vol. 81, no. 8, pp. 1731–1741, Aug. 2011.
[87] J. Hu and A. V. Vasilakos, "Energy big data analytics and security: Challenges and opportunities," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2423–2436, Sep. 2016.
[88] A. J. Ardakani and F. Bouffard, "Identification of Umbrella Constraints in DC-Based Security-Constrained Optimal Power Flow," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3924–3934, Nov. 2013.
[89] J. Peppanen, M. J. Reno, R. J. Broderick, and S. Grijalva, "Distribution system model calibration with big data from AMI and PV inverters," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2497–2506, Sep. 2016.
[90] H. Shaker, H. Zareipour, and D. Wood, "Estimating power generation of invisible solar sites using publicly available data," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2456–2465, Sep. 2016.
[91] H. Shaker, H. Zareipour, and D. Wood, "A data-driven approach for estimating the power generation of invisible solar sites," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2466–2476, Sep. 2016.
[92] X. Zhang and S. Grijalva, "A data-driven approach for detection and estimation of residential PV installations," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2477–2485, Sep. 2016.
[93] L. Chu, R. Qiu, and X. He, Z. Ling, and Y. Liu, "Massive streaming PMU data modelling and analytics in smart grid state evaluation based on multiple high-dimensional covariance test," *IEEE Trans. Big Data*, vol. 4, no. 1, pp. 55–64, Mar. 2018.
[94] K. J. Ross, K. M. Hopkinson, and M. Pachter, "Using a distributed agent-based communication enabled special protection system to enhance smart grid security," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 1216–1224, Jun. 2013.
[95] C. Vellaithurai, A. Srivastava, S. Zonouz, and R. Berthier, "CPIndex: Cyber-physical vulnerability assessment for power-grid infrastructures," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 566–575, Mar. 2015.
[96] A. Ahmed, V. V. G. Krishnan, S. A. Foroutan, M. Touhiduzzaman, C. Rublein, A. Srivastava, Y. Wu, A. Hahn, and S. Suresh, "Cyber physical security analytics for anomalies in transmission protection systems," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting (IAS)*, Sep. 2018, pp. 1–8.
[97] M. Touhiduzzaman, A. Hahn, and A. Srivastava, "ARCADES: Analysis of risk from cyberattack against defensive strategies for the power grid," *IET Cyber-Phys. Syst., Theor. Appl.*, vol. 3, no. 3, pp. 119–128, Sep. 2018.
[98] L. Wei, A. I. Sarwat, and W. Saad, "Risk assessment of coordinated cyber-physical attacks against power grids: A stochastic game approach," in *Proc. Ind. Appl. Soc. Meeting*, Oct. 2016, pp. 1–7.
[99] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, and N. A. Khan, "Load forecasting, dynamic pricing and DSM in smart grid: A review," *Renew. Sustain. Energy Rev.*, vol. 54, pp. 1311–1322, Feb. 2016.
[100] A. K. Singh, I. Ibraheem, S. Khatoon, M. Muazzam, and D. K. Chaturvedi, "Load forecasting techniques and methodologies: A review," in *Proc. IEEE 2nd Int. Conf. Power, Control Embedded Syst.*, Dec. 2012, pp. 1–10.
[101] J. Kang and S. Lee, "Data-driven prediction of load curtailment in incentive-based demand response system," *Energies*, vol. 11.11, p. 2905, Oct. 2018.
[102] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE Trans. Ind. Informat.*, vol. 7, no. 3, pp. 381–388, Aug. 2011.
[103] A. J. Conejo, J. M. Morales, and L. Baringo, "Real-time demand response model," *IEEE Trans. Smart Grid*, vol. 1, no. 3, pp. 236–242, Dec. 2010.
[104] S. Datchanamoorthy, S. Kumar, Y. Ozturk, and G. Lee, "Optimal time-of-use pricing for residential load control," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Oct. 2011, pp. 375–380.
[105] K. Zhou, C. Fu, and S. Yang, "Big data driven smart energy management: From big data to big insights," *Renew. Sustain. Energy Rev.*, vol. 56, pp. 215–225, Apr. 2016.
[106] S. Zhang, "Cloud computing-based analysis on residential electricity consumption behavior," *Power Syst. Technol.*, vol. 37, no. 6, pp. 1542–1546, 2013.
[107] P. Samadi, A.-H. Mohsenian-Rad, R. Schober, V. W. S. Wong, and J. Jatskevich, "Optimal real-time pricing algorithm based on utility maximization for smart grid," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Oct. 2010, pp. 415–420.
[108] A. Alsalemi, C. Sardianos, F. Bensaali, I. Varlamis, A. Amira, and G. Dimitrakopoulos, "The role of micro-moments: A survey of habitual behavior change and recommender systems for energy saving," *IEEE Syst. J.*, vol. 13, no. 3, pp. 3376–3387, Sep. 2019.
[109] B. Chen, K. L. Butler-Purry, and D. Kundur, "Impact analysis of transient stability due to cyber attack on FACTS devices," in *Proc. North Amer. Power Symp. (NAPS)*, Sep. 2013, pp. 1–6.
[110] A. Bifet and G. De F. Morales, "Big data stream learning with SAMOA," in *Proc. IEEE Int. Conf. Data Mining Workshop*, Dec. 2014, pp. 1199–1202.
[111] L. M. Camarinha-Matos, "Collaborative smart grids—A survey on trends," *Renew. Sustain. Energy Rev.*, vol. 65, pp. 283–294, Nov. 2016.

YING MA received the Ph.D. degree in computer science from the University of Electronic Science and Technology of China, in 2012. Since 2015, he has been an Associate Professor of software engineering with the Xiamen University of Technology. He is currently the Deputy Director of the Key Laboratory of Data Mining and Intelligent Recommendation, Fujian Province University. He has published over 30 articles in top conferences and journals, and has been awarded several grants, including the National Nature Science Foundation of China. He received the Distinguished Young Scholar Career Awards from Fujian Province. His article has been ranked as the third most cited articles in information and software technology articles list (Elsevier). He serves as the Reviewer for the IEEE TRANSACTIONS ON SOFTWARE ENGINEERING and *Information Sciences*.

CHAO HUANG graduated from the Department of Computer Science and Technology, Fuzhou University. He has worked as the Deputy Manager of market development of Fujian Yili Power Technology Company Ltd., the Assistant General Manager of State Grid Xintong Yili Technology Company Ltd., and the Deputy General Manager of Nanjing Nanrui Group Company Ltd. He is currently the General Manager of Xiamen Great Power Geo Information Technology Company Ltd. He holds five authorized patents and has published dozens of articles. It received the first prize of 2018 China energy enterprise information-based product technology innovation award—siji map issued by China information association.

YU SUN received the master's degree from the University of Exeter. She is currently pursuing the Ph.D. degree in education and science from National Tsing Hua University. She has been a Visiting Student with UCLA and has published more than a dozen articles.

GUANG ZHAO graduated from Wuhan University majoring in cartography and geographic information system. He has been the System Architect, the Technical Director, and the Deputy General Manager of Xiamen Great Power Geo Information Technology Company Ltd., where he is currently the Chief Engineer. He obtained five authorized patents and published more than a dozen articles. His independently developed geographic information service platform based on cloud computing received the first prize of Xiamen Science and Technology Progress Award, the second prize of Fujian Science and Technology Award, and the first prize of China Electric Power Science and Technology Award. He led the research and development project of key technology and application of spatio-temporal information service platform for complex large power grid, which received the second prize of National Science and Technology Progress Award.

YUNJIE LEI received the bachelor's degree in software engineering from the Anyang University of Technology, in 2016. She is currently pursuing the master's degree in computer science with the Xiamen University of Technology. She was a Big Data Mining Engineer with Zfusion Tech Company Ltd. Her main research interests include pattern recognition and data mining.