

A New Modeling Approach for Power Grid Online Analysis

Mike Zhou* DongHao Feng**

*China EPRI, Beijing 100192, China (mike.zhou@interpss.org)

** State Grid EPRI, Beijing 100192, China (dh.feng@foxmail.com)

Abstract: A new modeling approach for power grid online analysis is proposed to support the realization of a new online analysis system architecture. The model-driven software development, automatic code generation, and in-memory computing techniques are employed in the modeling approach. Data source adapters are developed for the integration of the model with the existing EMS system. A large-scale power grid online network data model (~40K-bus) is used for the model performance testing in a Lab environment. The proposed modeling approach was used to develop a new online analysis application, which was deployed and running in a provincial grid dispatching control center. The Lab and field performance measurement shows that the modeling approach-based application can achieve second-order end-to-end responsiveness.

Keywords: Grid Online Analysis, Online Analysis Model, Physical Data Model, Simulation Data Model, Digital Twin.

This work was supported by the National Natural Science Foundation of China under the Grant U1766214.

NOMENCLATURE

SCADA	Supervisory Control and Data Acquisition
SE	State Estimation
RTU	Remote Terminal Unit
EMS	Energy Management System
IMC	In-Memory Computing
MFP	Model-Focused Pattern
AFP	Algorithm Focused Pattern
DSA	Dynamic Security Assessment
DT	Digital Twin
EMF	Eclipse Modeling Framework
NN	Neural Network
CCT	Critical Clearance Time

1. INTRODUCTION

China has built the world's largest UHV AC/DC interconnected power grid. Because of the interconnection, the Chinese state grid currently becomes a strongly interconnected power grid. The disturbance caused by an outage in one part of the grid is much more likely to be spread to other parts of the grid. As a consequence, the grid is facing unprecedented operation complexity and security risk. Therefore, there is an urgent need to develop a new generation of online analysis tools to help the operator to mitigate complexity and risks. [Zhou(2018a)]

In the context of power grid online analysis, real-time refers to the response speed of data processing and computation in the order of seconds (sec-order). After a decade of development, power grid online analysis tools have been widely used in the dispatching control centers in China. The state grid currently is modeled using a ~40K-bus network data model. A round online analysis trip, including SCADA, State Estimation (SE), and online analysis, currently takes about 10 minutes to complete. Or in other words, the current online analysis system has min-order end-to-end responsive-

ness. [Zhou(2018a)] Online analysis is currently performed in the dispatching control centers in China periodically at an interval of 15 minutes. One of the key requirements for the new generation online analysis system is the sec-order real-time responsiveness. A real-time online analysis system could help the grid operator to perform the online analysis and support their decision-making process in real-time, while the system dynamic process is developing, rather than the current min-order after-the-fact analysis.

Conceptually, in a grid online analysis process, a Loadflow snapshot is first created, through SE, to be used as the basecase or initial state for the downstream online analysis applications. In the Digital Twin's (DT) view [Zhou(2019a)], the Loadflow snapshot could be considered as a virtual model representation to mirror the physical grid for the online analysis purpose. However, in the current online analysis systems used in the dispatching control centers in China, the Loadflow snapshot (virtual model) normally has a time delay of ~2 min (min-order) with respect to the RTU measurement info captured in the SCADA system. As stated in Ref. [Chen(2012)], with the conventional approaches in the dispatching control centers, SE and contingency analysis are normally computed and updated in an interval of minutes, which is not fast enough to recognize and anticipate system status if there is an emergency. In the upcoming new generation of online analysis systems, the virtual model is required to have a sec-order time delay in order to achieve the so-called sec-order end-to-end responsiveness. This paper presents a grid online modeling approach and associated software implementation techniques to support the realization of such a real-time (sec-order) virtual model.

In the current online analysis systems, there is usually a data integration module/layer to assemble the online data stored in the virtual model with other off-line simulation data sources,

to feed the downstream online analysis applications. [Wang(2006a), Wang(2006b), Yan(2007), Li(2006)] To our knowledge, there is no formally defined computer in-memory data model per se in the current online analysis systems in production worldwide.

A new online analysis system architecture was proposed in Ref. [Zhou(2018a)], in which the Model-Focused Pattern (MFP) approach is used for the grid online analysis model (virtual model) implementation. The virtual model is hosted in a data grid to mirror the power grid operation state with a sec-order delay. The IMC approach is used to allow the simulation algorithms to be applied to the Loadflow snapshot stored in the virtual model directly through in-memory access to perform "what-if" simulation. This new modeling approach and the associated implementation techniques are presented in this paper. Discussions of the general modeling concept and supporting implementation techniques are first presented in Section-II. The architecture and implementation details of the modeling approach are presented in Section-III. Application of the modeling approach to an online application and some performance measurement results are presented in Section-IV.

2. MODELING CONCEPT AND TECHNIQUE

Today, virtual models used in the real-world production environment are complex software systems, developed using modern software development methodologies and technologies. Modeling concepts and relevant software development techniques are reviewed in this section.

2.1 Model Concept

In general, there are three types of models: 1) physical model; 2) conceptual model; 3) concrete implementation model. [Zhou(2013)] The online analysis model discussed in this paper belongs to the concrete implementation model type category. The model is, in fact, an in-memory software implementation in the form of a Java library.

Almost all modern power grid simulation algorithms are formulated based on the grid network [Y]-matrix, which essentially is a conceptual model representing the power network for the simulation purpose. The [Y]-matrix model has been commonly implemented in Fortran using either one or two- dimensional arrays as the foundation data structure. This Fortran-style software implementation of the [Y]-matrix model has been used by almost all existing commercial power system simulation software.

It is very important to point out that a conceptual model could be implemented in multiple ways. This has been demonstrated by an object-oriented (OO) representation of the very same power network [Y]-matrix model in Ref. [Zhou(1996)]. In this OO representation, a set of power network modeling vocabulary (Network, Bus, and Branch) and ontology (the relationship between the entities), as shown in Fig. 1, are defined for modeling power network for the simulation purpose.

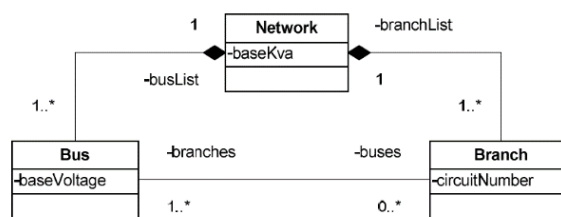


Fig. 1. Power Grid Simulation Data Model

2.2 Network Data Model

Power grid operation has various situation awareness analysis requirements. Therefore, an online analysis model usually consists of multiple types of data models. In general, there are three main types of grid online analysis data models: 1) physical (Node/Breaker) data model; 2) Bus/Breaker data model; 3) simulation (Bus/Branch) data model. For example, in the current EMS systems used in the dispatching control centers in China, the online Loadflow snapshot info after SE, stored in a file, called QS File in the CIM/E format, contains a Bus/breaker data model and a simulation data model.

In the recent PowSyBl project from the Linux Foundation Energy open-source initiative [LFenergy(2019)], IIDM (iTesla Internal Data Model) [Leclerc(2018)] provides an object-oriented model implementation in Java. PowSyBl/ IIDM formally establishes both physical data model and Bus/Breaker data model, and provides additional physical, Bus/Breaker and simulation client user data model views (interfaces). The physical data model and the Bus/Breaker data model do not contain a direct representation of the buses, branches, and their connectivity information, and therefore, could not be used efficiently to form the network [Y]-matrix for the online simulation algorithms. Therefore, IIDM could not be used directly for the IMC implementation of the online simulation algorithms, because of the absence of simulation (Bus/Branch) data model.

2.3 Data Processing Pattern

A power grid analysis application, in general, is intended to process some input info by applying some simulation algorithm(s) to produce some output results. Broadly speaking, there are two application data processing patterns: 1) Algorithm-Focused Pattern (AFP); 2) Model-Focused Pattern (MFP), as shown in Fig. 2. [Zhou(2017)]

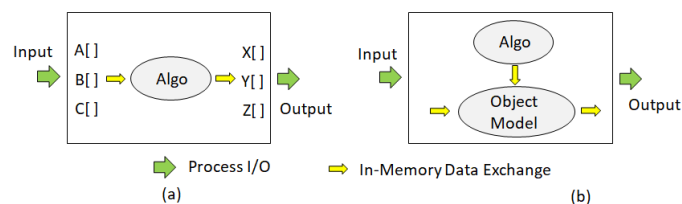


Fig. 2. Data Processing Patterns (a) AFP; (b) MFP.

In the AFP approach, info describing a physical entity(s), for example, a transformer, needs to be analyzed and then

transformed by the application user to a set of numbers, confirming a particular input data format, to create an input data file. Since the input data format could be defined to specially fit the simulation algorithm software implementation, it could be optimized to provide the highest input data processing efficiency. However, it often poses serious challenges for the maintainability and the extensibility of the algorithm software implementation.

In the MFP approach, info describing a physical entity(s) is preserved as an object(s) when creating its digital representation to be input to create the object model (the virtual model in DT's term). The data object stored in the object model in the simulation process needs to be transformed by some internal logic to fit the simulation algorithm software implementation, which might have some processing performance impact. Although there might be some efficiency penalty, this approach will benefit in terms of the maintainability and the extensibility of the algorithm software implementation.

The grid online analysis model presented in this paper is based on the MFP approach, while its simulation data model on the InterPSS object-oriented data model [Zhou (2017)].

2.4 Model Persistence

Grid online analysis data models are commonly persisted in EMS in-memory databases in the relational data structure. [Zhou(2018a)] In the conventional online analysis approach, where the AFP approach is employed, the virtual model data is exported from the data source, often to a set of exchange files, to feed the downstream online applications. On the other hand, when the MFP approach is employed, the virtual model could be hosted in a data grid in the object data structure. This object-oriented type of model persistence mechanism could support the IMC implementation, in which the data and algorithm are encapsulated inside the virtual model. [Zhou(2018b)] When a simulation request is sent to the virtual model, the algorithm is applied directly to the object model to process the data stored inside the model. The main IMC idea is "moving algorithm instead of moving data", which was originally proposed to solve the data movement challenges/issues in the Big Data processing situations.

2.5 EMF Modeling Technique

The model-driven development approach has been employed in the online analysis model development process. Particularly, Eclipse Modeling Framework (EMF) [Steinberg (2008)] based automated code generation technique has been employed. EMF is a modeling framework and code generation facility for building tools and applications based on structured data models. From a model specification, EMF provides tools and runtime support to produce a set of Java classes for the concrete implementation of the model. The EMF-based virtual model has been extended to be hosted in Hazelcast in-memory data grid [Hazelcast(2019)], which provides out-of-the-box support for the IMC implementation.

3. MODEL ARCHITECTURE AND IMPLEMENTATION

The online analysis modeling approach, including the model architecture and model implementation details, are presented in this section.

3.1 Modeling Approach Overview

In the online analysis process, the data stored in the virtual model, also called Loadflow snapshot or basecase, representing the current grid operation state, is fed as the input info to the downstream online applications. The current online analysis systems use the AFP approach, where the simulation algorithms are usually implemented and maintained separately and often independently from the virtual model. The virtual model and algorithms are usually hosted on different locations (computation nodes) in the EMS systems. The basecase data stored in the virtual model is often passed to the algorithm nodes using a set of exchange files or through remote memory mapping. There are at least two major challenges with this conventional approach: 1) There might be a large amount of data movement during the online analysis process, from the virtual model storage location to the simulation algorithm execution location, which often results in significant impact on the overall system performance; 2) There are intrinsic complications to maintain AFP based online applications, which were explained in Sec. 2.3.

The online analysis process has been re-architected using the MFP pattern as the foundation. In the MFP approach, the online analysis simulation algorithms are implemented on top of the virtual model. In other words, simulation algorithms become an integral part of the virtual model. In this way, the virtual model and the algorithms could be optionally hosted on the same computing node to implement the IMC strategy [Zhou(2018b)]. The simulation algorithms could have direct in-memory access to the basecase data stored in the virtual model to minimize the data movement in the online analysis process. Also, the MFP approach will force the simulation algorithm development to work closely with the virtual model development to reduce the software maintenance complexity and enhance the software extensibility.

3.2 Online Analysis Model Architecture

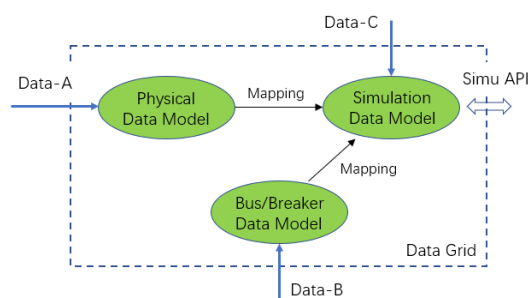


Fig. 3. Online Analysis Model Architecture

The online analysis model, as shown in Fig. 3, consists of three data models: 1) physical data model; 2) Bus/Breaker data model; 3) simulation model, and the mapping

relationship from the physical data model and the Bus/Breaker data model to the simulation data model. There are three data input channels or interfaces (Data-A~C in Fig. 3) for loading data into the data models from outside data sources, and a simulation API (Simu API in Fig. 3) to provide simulation service. Since the purpose of the online analysis model is to provide grid analysis/simulation service, therefore, there is only one simulation API for the outside world to interact with the model. Also, the simulation API, together with the hosting data grid, supports the IMC implementation of the simulation algorithm. Typically, the physical device measurement info from RTU together with their connectivity relationship and device parameters are stored in EMS's in-memory database. This physical device data is loaded into the online analysis model to establish the physical data model in the data grid, and also create the corresponding simulation data model through the topology analysis. The mapping relationship between the physical data model and the simulation data model is cached inside the data grid so that the topology analysis is only performed when there is a network configuration change caused by some switching event(s). Furthermore, the topology analysis is only performed to those substations affected by the switching event. Since the data models are hosted in the data grid, it can be updated partially and efficiently, i.e., only updating the changed parts to track the grid operation state change.

The online analysis model is implemented using the Java programming language. It could be used as a library in an application. However, in most cases, it is hosted in a data grid. The open-source in-memory data grid software Hazelcast [Hazelcast(2019)] is used in the current online analysis model implementation. Furthermore, the simulation data model implementation is based on the extension of InterPSS object model [Zhou(2017)]. The data input channel Data-A (Fig. 3) has been integrated with the D5000 EMS system, which currently is the standard EMS system used widely in China, to update the physical data model in real-time. The online analysis model, which hosted in the data grid, could be updated in two ways: 1) Batch processing mode in a fixed interval to refresh the model; 2) Event-driven mode in real-time for stream processing of the RTU measurement info.

3.3 Physical Data Model

Typically, the primary physical data model info from the RTU measurement info is stored in EMS's in-memory database. Each EMS vendor has its own proprietary in-memory database implementation. In our physical data model development, based on the D5000 EMS database structure configuration file (*.h), a set of physical data model record definitions (Java class) and the associated parsing/mapping logics are generated by an automatic code generation utility tool. In the runtime, based on the Java API of the D5000 EMS system, the data stored in the in-memory database is retrieved. The data is imported into the physical data model of the online analysis model through a D5000 EMS adapter (see Sec. 3.5). Then the simulation data model is created or updated through topology analysis. In the online analysis

model, the mapping relationship from the physical data model to the simulation data model is cached in the data grid, since the mapping relationship does not change frequently.

3.4 Simulation Data Model

In the power grid simulation, the grid can be represented by an object model, as shown in Fig. 1. In practice, the grid simulation problems can be divided into several abstract network layers, such as topology network (the base type), Loadflow calculation network, short-circuit current calculation network, transient stability simulation network, etc. Through inheritance and extension of these layers of network type definitions, a broader set of grid analysis/simulation problem solutions can be formulated [Zhou(2017)].

The online analysis model uses a key-value storage mechanism to store data in the data grid, as shown in Fig. 3. The data models could be stored as a large object directly inside the data grid in the form of a value object. Data grid software, in general, has out-of-the-box support for dynamic scalability, high availability, and support for object data model implementation. The data model objects stored in the data grid could be easily and efficiently cloned. By using the data affinity feature of data grid software, the cloned data model copies could be efficiently moved to other computation nodes in a distributed computing environment. Various online analysis application algorithms, such as N-1 contingency analysis, short-circuit current calculation and Neural Network (NN) model-based prediction could be applied to the copies to implement distributed parallel computation. The data affinity feature can also guarantee that a simulation algorithm shall be executed in the same place where the data model is located so that the data movement across computer nodes during the analysis/simulation process could be avoided. In this way, the so-called "moving algorithm instead of moving data" IMC processing strategy can be implemented in a distributed computing environment.

3.5 Data Source Adapter

In software engineering, the adapter pattern is a software design pattern that allows the interface of an existing data source to be used as another adapted interface. Therefore, the adapter design pattern allows otherwise incompatible data sources to work together by converting the interface of one data source into an interface expected by another data source. Based on the adaptation architecture, as shown in Fig.4, a set of adapters were developed for the communication of the data models (See Fig. 3) with various outside data sources, including the D5000 EMS system data sources.

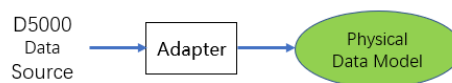


Fig. 4. D5000 EMS Data Source Adaptation Architecture

The D5000 EMS system uses the C programming language as its main programming language, while the online analysis model uses the Java programming language. The D5000 has

a high-speed Data Bus, which is used for the event-driven communication bridge between the D5000 subsystems (in C) and the online analysis model modules (in Java). A D5000 Data Bus based adapter was developed to facilitate the real-time event-driven data communication between the D5000 system and the online analysis model.

3.6 Model Performance Testing

Performance testing was carried out in a Lab environment. The Lab currently serves as the technical support center for the D5000 EMS system in the state grid dispatching control center. The state grid is currently modeled by a ~40K-bus online network data model. The Lab can reproduce actual scenarios happened in the state grid EMS system. The performance testing input data (RTU) were created by the replay of the RTU messages recorded in the field. The state grid SCADA subsystem currently has ~440K Point measurement points and ~450K Analog measurement points.

The performance testing steps include: 1) Replaying the RTU messages to the SCADA; 2) SCADA publishes Point/Analog messages on to the Data Bus; 3) The online analysis model subscribes the Point/Analog messages and updates the physical data model; 4) The simulation data model is updated accordingly through in-memory data mapping.

The following are the main Point/Analog message processing logic:

- 1) Point Message: The message is parsed, and the physical data model updated first. Then through topology analysis of the affected substation(s), the corresponding Bus/Branch connection relationship in the simulation data model is updated.
- 2) Analog Message: The message is parsed, and the physical data model updated first. Then the corresponding Bus/Branch state (P, Q, V, I) in the simulation data model are updated.

Table 1. Model Performance Testing Results

RTU Message Type	Model Throughput (per sec)	Actual Field Msg. Rate (per sec)	SCADA Throughput (per sec)
Point	1.7K	< 1	6.5K
Analog	290K	~1K	20.5K

The performance testing results to update the online analysis model are shown in Table 1, where the online analysis model RTU message processing throughput is shown, together with the actual field RTU message rate and the current SCADA RTU message processing throughput. As can be seen, the online analysis model can efficiently process the RTU measurement info and update its data models to achieve tracking of the changes in the grid operation status in real-time. The Analog/Point message processing processes have a sub-second delay when dealing with the large-scale grid network data model (~40K-bus).

4. ONLINE ANALYSIS APPLICATION

The modeling approach presented in the paper has been used in a new online analysis application, which was deployed and running in a provincial dispatching control center in China for the demonstration purpose. This section presents a summary of the application system, application scenarios and some of the performance measurement results.

4.1 Application System

The provincial dispatching center currently has, as shown in Fig. 5, a D5000 EMS system with an existing online analysis application (DSA in the upper part), running periodically with a 15-min interval. The D5000 EMS system has three Security Regions (I~III), where Region-I has the highest security restriction. The existing DSA is running in Region-I. The new online analysis application (the lower part in Fig. 5), because of the security policy constraint, was deployed in Region-III, since at this stage it was only for the demonstration purpose. The RTU measurement info was routed from SCADA in Region-I to Region-III with approximately one-min delay.

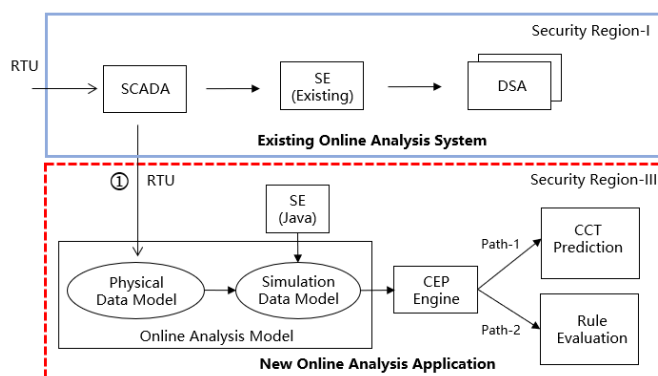


Fig. 5. Process Flow of the Online Analysis Application

4.2 Application Scenario

As shown in Fig. 5, the new online analysis process consists of three main steps: 1) The online analysis model subscribes the RTU message from SCADA and updates the physical data model and then the simulation data model to track the grid operation state; 2) IMC based SE (SE (Java) in the figure) is performed on the simulation model to create a converged Loadflow snapshot or basecase; 3) The CEP engine [Zhou (2019b)] drives the NN-model based CCT prediction application and the grid operation rule evaluation application. Data mapping and topology analysis (when necessary), are performed in Step-1 to map the physical data model changes to the simulation data model to keep them in synchronization. The NN models for the CCT prediction were trained off-line using the Machine Learning approach.

The provincial power grid has approximately 1000 buses. The new online analysis application is performed periodically at an interval of 20 seconds. When performing the online analysis, the IMC based SE algorithm is applied to the simulation data model to produce a basecase for the two

online analysis applications, i.e., the CCT prediction, and the Rule Evaluation (see Fig. 5).

4.3 Performance Measurement

Performance measurement results of the new online analysis application are listed in Table 2, where the processing time in each of the three steps is shown. As can be seen, using the RTU data entrance point (⊙ in Fig.5) as a starting point, the end-to-end online analysis process, including physical data model update, simulation data model update, SE, the NN-model based security assessment and the operation rule evaluation, could be completed in less than 300 ms. The SE computation could be completed in approximately 50 ms with a Qualification Measurement Rate (QMR) of ~98%. The QMR difference between the new IMC based SE and the existing production SE in the D5000 EMS system is less than 0.5%. In-depth discussion of the IMC based SE implementation is beyond the scope of this paper.

Table 2. Performance Observation Results

Processing Task	Task Details	Processing Time (sec)
Step-1	Synch of physical data model with SCADA	0.1
Step-2	Simulation data model update through topology analysis, then perform State Estimation	0.1
Step-3, Path-1	NN-model based CCT prediction	< 0.1
Step-3, Path-2	Operator dispatching rule evaluation	< 0.1

It should be pointed out that the new online analysis application currently only has two data-driven online analysis applications (see Fig. 5) for the demonstration purpose. Conventional online applications, such as N-1 contingency analysis, time-domain transient stability simulation, will be added and integrated in the near future.

5. CONCLUSIONS

A new modeling approach for power grid online analysis has been developed and presented, for supporting the realization of a new online analysis system architecture. The model-driven software development, automatic code generation, and in-memory computing techniques were employed in the modeling approach. Data source adapters were developed for the integration of the model with the D5000 EMS system, including its SCADA in-memory database and the SE result in the CIM/E QS File format.

A large-scale power grid online network data model (~40K-bus) was used for the model performance testing in a Lab environment. The online analysis model could efficiently process the RTU measurement info and update its data models to achieve tracking of the change of the grid operation status in real-time. The Analog/Point message processing processes had a sub-second delay when dealing with the large-scale network data model.

The modeling approach was also been used in a new online analysis application, which was deployed and running in a provincial grid dispatching control center. The field performance measurement results show that the new online analysis application can achieve sub-second end-to-end responsiveness. The online analysis process, including physical data model update, simulation data model update, SE, the NN-model based security assessment and the operation rule evaluation, could be completed within 300 ms.

REFERENCES

- Chen, Y.S., Huang, Z.Y., et al. (2012). Computational challenges for power system operation. *45th Hawaii International Conference on System Science*, Hawaii, USA, 2141-2150.
- Hazelcast, (2019). Hazelcast in-memory data grid. hazelcast.com.
- Leclerc, S., (2018). PowSyBl Basics IIDM. *iPST/ PowSyBl day*, <http://www.itesla-pst.org/>
- LF Energy, (2019). Linux Foundation Energy open source initiatives, <https://www.lfenergy.org>.
- Li, Y.F. , et al, (2006) . System Data Conformity on Dispatching Application of Electric Power System. *Jinlin Electric Power*, Vol.34, No.5, 24-25.
- Steinberg, D., el at, (2008). *EMF: Eclipse Modeling Framework*, 2nd Edition, Addison-Wesley Professional, 2008.
- Wang, J.Y., (2006a). An integration Solution of Power Information Resources based on Grid Technology . *Automation of Electric Power Systems*, Vol.30, No.20, 84-87.
- Wang, J.Y., Zhang, C.J., (2006b). Survey on Power Information Resources Integration Methods. *Power System Technology* , Vol.30, No.9, 83-87.
- Yan, Y.Q., et al, (2007). Research and Practice of Integration of Information in Power Dispatching Center. *Relay*, Vol.35, No.17, 37-40.
- Zhou, E. (1996). Object-oriented Programming, C++ and Power System Simulation. *IEEE Transactions on Power Systems*, Vol.11, No.1.
- Zhou, M., (2013). Distributed parallel power system simulation. *High Performance Computing in Power and Energy Systems*, pp.71-100, Berlin Heidelberg, Springer-Verlag.
- Zhou, M., Huang, Q.H. (2017). InterPSS: A New Generation Power System Simulation Engine, <https://arxiv.org/abs/1711.10875>.
- Zhou, M., Yan, J.F. (2018a). A New Solution Architecture for Online Power System Analysis. *CSEE Journal of Power and Energy System*, Vol.4, No.2, 49-55
- Zhou, M., Feng, D.H. (2018b). Application of In-Memory Computing to Online Power Grid Analysis, *The 10th IFAC Symposium on Control of Power and Energy Systems*, Tokyo, Japan.
- Zhou, M., Yan, J.F. and Feng, D.H. (2019a). Digital Twin Framework and its Application to Power Grid Online Analysis. *CSEE Journal of Power and Energy System*, Vol.5, No.3, 391-398.
- Zhou, M., Yan, J.F. and Feng, D.H. (2019b). Development of Complex Event Processing (CEP) Engine for Application to Power Grid Online Analysis. *Automation of Electrical Power Systems*, Vol.43, No.20:105-113.