Networked Information Gathering and Fusion of PMU Data

Future Grid Initiative White Paper

Power Systems Engineering Research Center

Empowering Minds to Engineer
the Future Electric Energy System

Networked Information Gathering and Fusion of PMU Data

A Broad Analysis Prepared for the Project
"The Future Grid to Enable Sustainable Energy Systems"
Funded by the U.S. Department of Energy

White Paper Team

Junshan Zhang and Vijay Vittal
Arizona State University

Peter Sauer
University of Illinois at Urbana Champaign

PSERC Publication 12-07

May 2012

For information about this white paper contact:

Junshan Zhang
Professor, School of Electrical, Computer and Energy Engineering
Arizona State University
Tempe, Arizona 85287-7206
Office: Goldwater Center 411D
E-mail: junshan.zhang@asu.edu
Phone: (480) 727-7389
Fax: (480) 965-8325

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-965-0745

Notice Concerning Copyright Material

Acknowledgements

This white paper was developed as one of nine broad analysis white papers in the project "The Future Grid to Enable Sustainable Energy Systems: An Initiative of the Power Systems Engineering Research Center." This project is funded by the U.S. Department of Energy. More information about the Future Grid Initiative is available at the website of the Power Systems Engineering Research Center (PSERC), www.pserc.org. This white paper is in the broad analysis area "The Information Hierarchy for the Future Grid."

We also recognize the staff, faculty, and students of the Power Systems Engineering Research Center for their efforts in developing the vision that led to the "Future Grid" project under which this white paper falls. Finally, we express deep appreciation to the several reviewers who significantly contributed to the quality of this white paper, as listed below. Their identification here does not constitute endorsement regarding any of the contents of this report.

- Gilbert Bindewald, U.S. Department of Energy
- Floyd Galvan, Program Manager, Research & Development, Entergy
- Naim Logic, Senior Electrical Engineer, Salt River Project (SRP)
- Shimo Wang, Southern California Edison

Executive Summary

The nation's power grid is perhaps the most dynamic and heterogeneous man-made network, and its modernization involves not only the physical-system, but also its cyber-infrastructure. The smart grid in the making is envisaged to integrate a considerable amount of renewable energy resources, which are highly variable. To meet these challenges, a key step is to develop real-time, lightweight and adaptive algorithms for three core functions, namely measurement, fusion, and communication, which will be responsive to the dynamics of the grid and support various applications with diverse requirements. However, the existing supervisory control and data acquisition (SCADA) systems provide only the static states or the quasi-static states of the power grid.

The synchrophasor technology is emerging as an enabling technology to facilitate both information interaction as well as energy interaction between providers and customers, and help revolutionize the power system. In particular, it is critical to ensure reliable and secure communication systems for synchrophasor data. In this report, we identify a few important problems in this fundamental building block in the smart grid as follows.

- What data processing, calibrating, and filtering algorithms are needed to ensure that quality data is stored and distributed for use in energy management systems at the proper time scale?
- What are the suitable criteria for designing the communication systems for synchrophasor data, and how would the off-the-shelf communication technologies perform?
- What fusion mechanisms would work efficiently to extract useful information from synchrophasor data?
- How robust are interdependent cyber-physical systems (particularly the interconnected power grid and communications system) to cascading failures, and how can we improve the robustness of the overall system?

A primary objective of this white paper is to provide an overview of major challenges in gathering and data fusion of PMU measurements, and to discuss potential solutions to the aforementioned problems.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

A phasor measurement unit (PMU) is a device that is capable of measuring the time-stamped values of voltage and current (fundamental-frequency) phasors in power grids at a rate of up to one per fundamental cycle. PMUs' integration of global positioning satellite (GPS), together with a common time reference provided by GPS, allows the measurements from widely dispersed locations of power grids to be gathered in a synchronized fashion. Compared to the measurements in traditional SCADA systems, *synchrophasor data* (which refer to the time-aligned measurements collected by PMUs), can provide the *real-time* measurements of system states, including the voltage and current phase angles, at a higher precision than estimated states in SCADA systems. Further, in synchrophasor data, the measurements are taken at a much finer timescale (PMU measurements can be collected up to once per fundamental cycle), which allow synchrophasor data to capture reasonably fast dynamics of power systems. These salient features have made PMUs powerful monitoring instruments and widely deployed in power grids. The benefits of synchrophasor data to other power system applications have also been well recognized [1]. Generally, synchrophasor-based applications can be classified into three categories:

- *Wide-area monitoring*: visualization, state measurement and estimation, load model synthesis;
- *Wide-area protection and control*: e.g., dynamic security assessment (DSA), voltage stability detection and correction, islanding control;
- *Post-event analysis and research*: e.g., fault detection and localization, model validation.

1.2 White Paper Organization

This white paper is organized into four chapters. Following the Introduction in Chapter 1, Chapter 2 addresses networked communications of synchrophasor data, where diverse quality-of-service (QoS) requirements of synchrophasor data communications and potential networking technologies are discussed. Chapter 3 is focused on network fusion of synchrophasor data. Building on our recent studies on two important synchrophasor-based applications, we demonstrate how data fusion could be effectively performed. Finally, in Chapter 4, we discuss the impact of the inherent interdependence between the communication network and the power grid, from a robust architecture perspective.

2 Networked Communications of Synchrophasor Data

2.1 System Architecture

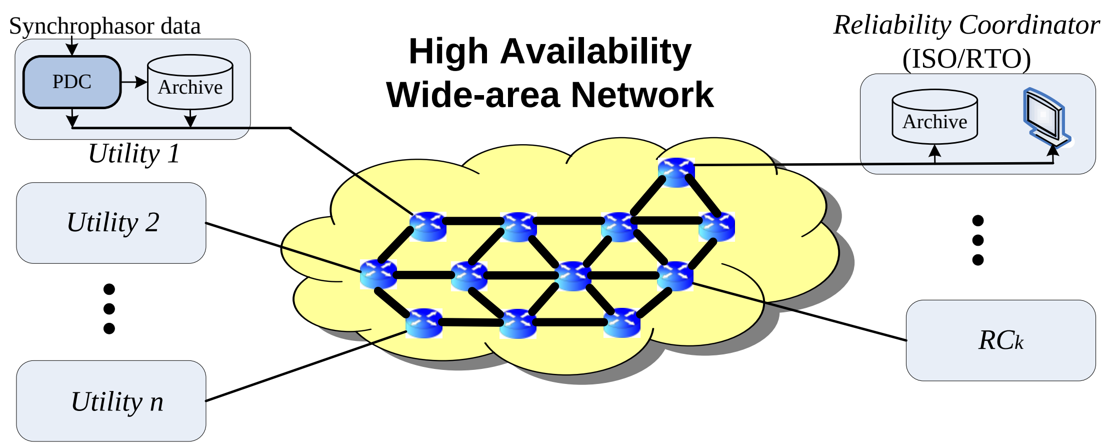


Figure 1: An Open-Access Information Architecture for Power Grids

Modern electric power grids are highly interconnected systems. Recently, the deregulation of the power industry has moved the operations of power grids from vertically integrated-centralized ones to coordinated-decentralized ones [2]. Specifically, in North America electric power grids, utilities are committed to balancing the load and generation in real-time in a given area; and these balancing authorities (BAs) and reliability coordinators (RCs), such as independent system operators (ISOs) or regional transmission organizations (RTOs), are responsible for overseeing the reliable operations of the grid and providing coordination over a wide area as needed.

Figure 1 depicts an information architecture that supports the aforementioned coordinated operations. Within this architecture, the communication system consists of two levels - the intra-utility level and inter-utility level. At the intra-utility level, phasor data concentrators (PDCs) gather the synchrophasor data from phasor measurement units (PMUs), process them (e.g., time-align the data), and then submit them to the utility control center for various applications and archiving. At the inter-utility level, a wide area network (WAN) ensures the high availability of synchrophasor data (real-time data and archived historical data) to proper applications at various utilities and RCs, so that wide-area monitoring, protection and controls can be carried out in a timely manner. The North American SynchroPhasor Initiative (NASPI) is coordinating a significant effort in this area (https://www.naspi.org/).

2.2 Enabling Technologies for High Availability of Synchrophasor Data

As synchrophasor data becomes more important to the monitoring and operations of the power grids, there is a need to architect and design communication systems that ensure a high level of availability of high-quality synchrophasor data. Specifically, high availability of synchrophasor data at the intra-utility level means that the measurements at substations should be consistently accessible to local utilities. And at the intra-utility level, the delivery of synchrophasor data to RCs and other utilities has to be completed within a critical timeline, since outdated or erroneous measurements neither contribute to enhancing the real-time situational awareness nor contain valuable information for protection and controls. Those two aspects of high availability of synchrophasor data should be respected when we design the intra-utility level and inter-utility level communication systems.

2.2.1 Redundance Configuration of Intra-Utility Level Communication Systems

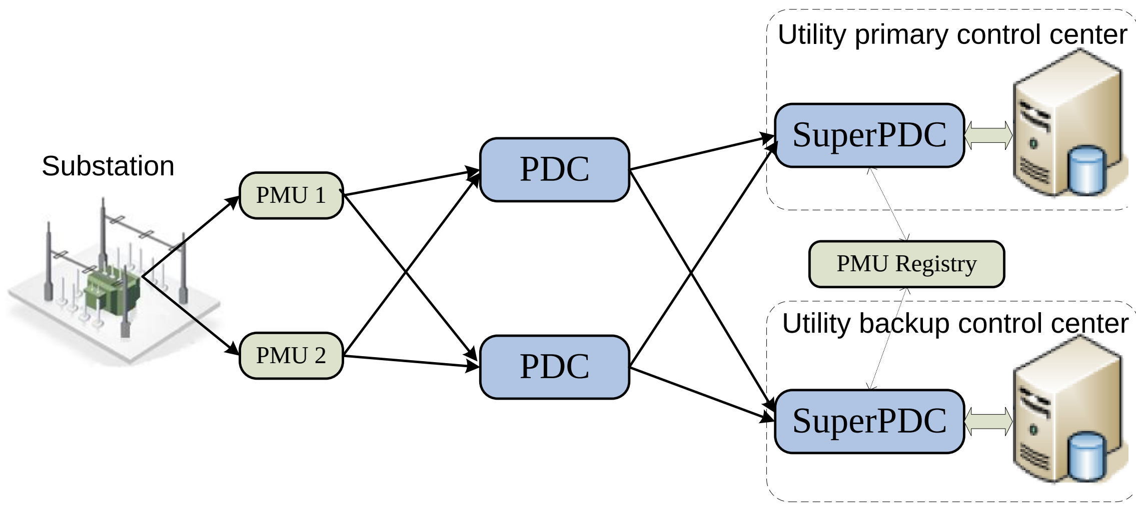


Figure 2: A Redundance Configuration of the Intra-Utility Communication System

One practical criterion for the design of the intra-utility level communication systems is that a utility should still be able to access all the local measurements under the “N-1” events, i.e., when one of the devices (e.g., PMUs or PDCs) or a communication link is down.

In order to satisfy this criterion, consistent accessibility of synchrophasor data could be assured by increasing the redundancy throughout the intra-utility level communication systems [3]. Specifically, redundant communication links could be deployed, to mitigate the failures of communication links. Further, PMUs and PDCs could be implemented in a redundant pair in substations and control centers, respectively, to assure that there are no disruption of availability, in case that these devices experience unexpected failure or scheduled maintenance.

Figure 2 illustrates a redundancy configuration, which results in an intra-utility level communication system complying with the “N-1" criterion, from redundant PMUs, to redundant communications links, and to redundant SuperPDCs at both primary and backup control centers. With interactions with the PMU registry [4], a name server which maps synchrophasor data to the information on where and how the measurements are taken, the SuperPDCs at the utility control centers ensure that only one copy of the redundant measurements is submitted.

2.2.2 Deadline-Driven Data Delivery for Inter-Utility Level Communications

Early efforts on the delivery of synchrophasor data at the inter-utility level have focused on using off-the-shelf networking technologies, including transmission control protocol (TCP) and user datagram protocol (UDP) (e.g., in [5]), and bandwidth reservation mechanisms (e.g., in [6]). However, by exploring the diverse QoS requirements of transmitting synchrophasor data, we observe that existing off-the-shelf networking technologies are subject to noticeable deficiencies, when used for the delivery of synchrophasor data. Specifically, UDP is not a good choice for the communications of synchrophasor data, since it provides no guarantees for delivery. TCP can provide delivery guarantee, but is not deadline-aware. In a nutshell, existing reservation mechanisms lack flexibility when handling short data flows, and thus may result in inefficient communications.

In what follows, we first discuss the diverse QoS requirements for synchrophasor data communications. Then, we give a brief introduction to the aforementioned off-the-shelf networking technologies and discuss their deficiencies for the delivery of synchrophasor data. Finally, we discuss several key techniques for designing a new deadline-driven flexible data delivery scheme which was proposed for meeting deadlines in data center communications [7].

QoS Requirements of Synchrophasor Data Communications

Power system dynamic phenomena are complex multi-timescale events. A variety of wide-area sensing and control actions have been designed to take place on time scales ranging from 10^{-6} to 10^{4} seconds. The multi-timescale nature of monitoring and control applications implies that the delivery of synchrophasor data could have different requirements in terms of latency and update frequency. Further, for critical applications, the synchrophasor data should have commensurate priorities in the delivery. In short, synchrophasor data communications for different applications can have different QoS requirements.

Based on [8], some requirements of communications for the three categories of synchrophasor-based applications considered here are summarized in Table I. Generally,

- Synchrophasor data for wide-area monitoring, protection and control applications have stringent latency requirements and higher priorities;
- Synchrophasor data for other applications, which may contain a large amount of historical measurements, have relatively larger deadlines and lower priorities;
- Synchrophasor data flows for wide-area monitoring, protection and control applications, which correspond to real-time updates on the measurements, are mostly very short. These characteristics and the diverse QoS requirements of synchrophasor data flows should be taken into account when designing the inter-utility level communication systems.

Table I: QoS Requirements of Synchrophasor Data Communications

	Monitoring	**Protection and Control**	**Post-event Analysis and Research**
Latency	≤ 1000 ms	≤ 5 ms	10^4-10^6 ms
Updating frequency	1-120 Hz	30-120 Hz	≤ 1 Hz
Priority	medium - high	High	Low

Off-the-Shelf Networking Technologies: Deficiencies

This section starts with a brief overview of TCP and bandwidth reservation mechanisms, followed by a discussion of their possible pitfalls when used for the delivery of synchrophasor data.

TCP

TCP is layered above the Internet Protocol (IP), which is a best-effort delivery scheme, in the sense that packets sent via IP are not guaranteed to be delivered to the destination. In TCP, successful transmissions are verified through the acknowledgements (ACKs), and reliability is provided by retransmitting the packets that are identified as lost, until their arrival at the receiver are confirmed through ACKs.

TCP uses a flow-control mechanism, where the receiver advertises the size of the available buffer space, so that the transmitter will not overwhelm the receiver's capacity to process the received packets. Usually, in large-scale networks, such as the Internet, there are a large number of TCP transmitters/receivers with sufficient buffer spaces, which may transmit more packets than the network can handle. This could lead to a situation of *congestion*, which could degrade the network throughput dramatically. To handle this situation, TCP includes a congestion control mechanism.

Specifically, TCP uses a number of mechanisms to mitigate/avoid congestion to acheive high network throughput. These mechanisms control the rate of data entering the network, and try to maintain the data flow below a rate that would trigger collapse otherwise. One congestion avoidance algorithm used by TCP is the *additive increase/multiplicative decrease (AIMD)* scheme, with other schemes such as *slow-start* in order to achieve *congestion avoidance*. In *slow start*, TCP begins by transmitting just one packet at a time. When each successful transmission is confirmed by an ACK, the number of packets that can be transmitted is doubled. This exponential increase of the number of packets to transmit in the *slow start* phase continues until a threshold is reached. Then, the additive increase mechanism is invoked, and the transmitter increases the transmission rate by a fixed amount every round trip time (RTT). When congestion is detected, the transmitter decreases the transmission rate by a multiplicative factor (e.g., 1/2, i.e., to reduce the transmission rate by half). An AIMD mechanism requires a signal of network congestion. Usually, it is assumed by the TCP that the loss of a packet is an indicator of network congestion. In summary, TCP's flow control and congestion control mechanisms can result in very high network utilization that can be shared in a fair manner between the TCP connections in the network. These advantages have made TCP widely used in Internet applications.

Despite these advantages, TCP was not designed for applications with diverse deadline requirements, and it has some undesirable pitfalls that many result in deficiencies in the delivery of synchrophasor data. Specifically, TCP is oblivious of data deadlines, and can incur relatively long delays (in the order of seconds) while waiting for the out-of-order packets or re-transmitting the lost packets. As a result, it can severely impact the application performance, e.g., for synchrophasor data that are required to be delivered with latency no greater than 5ms. TCP also lacks the provisioning for priorities. For the queue at the transmitters and routers, messages are delivered in a strict first-in first-out (FIFO) order. When network traffic intensity is high, it would be difficult for time-critical and high-priority synchrophasor data to initiate new TCP connections or to override existing connections. Further, TCP's tightly integrated congestion control mechanism could interfere with time-critical transmissions. For example, the slow-start phase can make high-priority data undergo unnecessary delays.

Bandwidth Reservation Technologies

The inherent lack of mechanisms for prioritizing the data flows makes TCP vulnerable to the new situation, where synchrophasor data have diverse priorities and QoS requirements. Bandwidth reservation mechanisms can be used for network operators to reserve bandwidth for data with different priorities and to help mitigate this vulnerability. For example, synchrophasor data flows can travel over reserved channels through the multiprotocol label switching (MPLS) services [6], which allow the bandwidth reserved within routers so that high-priority synchrophasor data flow will be guaranteed to be allocated of proper resources regardless of the other traffic types in the network. And in GridStat [9], QoS brokers are responsible for routing and creating bandwidth reservations over communication links.

One potential drawback of the bandwidth reservation technique noted above is its inflexibility. For MPLS, if the bandwidth required by the synchrophasor data is above the reserved value, the extra data will be delivered in a best-effort manner. Complication may arise when disturbance or other system events happen. It is very likely that future protection and corrective control schemes would depend on synchrophasor data with very high updating rates (e.g., up to 720Hz [8]). However, in a bandwidth reservation environment, these high volumes of synchrophasor data are likely to be transmitted in a best-effort manner. The same issues can also be troublesome with GridStat. GridStat can provide QoS guarantees during the stable state, but when unexpected system events happen followed by a surge of high-priority synchrophasor data, the global adaptation of the broker-based systems would be necessary, i.e., the QoS brokers negotiate and routers wait for QoS brokers' instructions until a new routing and bandwidth reservations are agreed by QoS brokers and setup. The time interval between the request for global adaptation and the accomplishment of broker negotiation is non-negligible [6], and further, within this interval the delivery of synchrophasor data flows with their latency requirements are not guaranteed, i.e., not meeting the deadlines or packet loss may occur.

In summary, the bandwidth reservation mechanisms noted above lack flexibility when handling synchrophasor date communications during unexpected system events, and may result in low network utilization, especially given the characteristics of synchrophasor date flows, i.e., most high-priority synchrophasor date flows are very short.

Towards a Deadline-Driven Flexible Delivery of Synchrophasor Data

In order to mitigate the deficiencies of off-the-shelf networking technologies, one need to design a new scheme catering to the diverse characteristics and the QoS requirements of synchrophasor data communications. Next, we elaborate on a deadline-driven scheme that consists of a queue management model, a dynamical rate allocation mechanism, and a flow quenching mechanism.

Queue Management

Each router in the WAN (as depicted in Figure 1) maintains three queues, corresponding to the three categories of synchrophasor-based applications. The priority of the queues used for rate allocation by the router is the same as those of the applications.

Dynamic Rate Allocation

With dynamic rate allocation, each transmitter makes rate requests on a slot basis (e.g., a slot may span one RTT). The desired rate of a flow is set by the transmitter, and carried in the packet header, to traverse the routers along the path to the destination. For example, given a flow with size s and deadline d, the desired rate can be set as:

$$r = s / d .$$

For each of the outgoing interfaces, routers receive rate requests from flows with different deadlines and priorities. Specifically, the rate allocation problem for a router is defined as follows: given the rate requests of outgoing data flows, what rates should be allocated to flows, so that (based on their priorities) the number of flows which satisfy their deadlines is maximized and the network capacity is most utilized. Clearly, the solution to this multi-objective problem is non-trivial, especially in a dynamic setting.

After all rates are allocated to data flows and fed back to the transmitter through the ACKs on the reverse path, the transmitter thus can determine its sending rate, i.e., the minimum of all rates allocated by the routers the data flow traverses. The transmitter then sends data at this rate during the current slot, while piggybacking a rate request for the next slot. It is worth noting that, different from those in bandwidth reservation mechanisms, the data flows are not assigned with a reserved bandwidth throughout its duration. The rate that a router allocates to a data flow varies all the time, and each transmitter must periodically make request for a new allocation. Since the actual rate allocated by routers may not be exactly what is needed, the desired rate of a data flow should also be re-computed as the deadline and the remaining flow size change.

Flow quenching

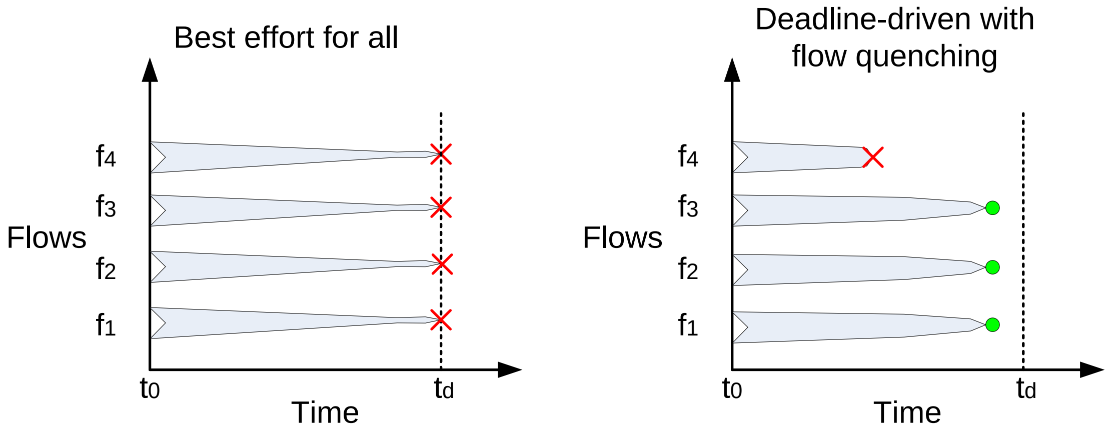


Figure 3: Multiple Flows: Conventional Best Effort vs. Flow Quenching

The deadline-driven scheme can also utilize "flow quenching" to cope with severe congestions. Under congestion, it might be better to shed some loads, and spare the resources for the rest of data flows to meet their deadlines, rather than to make all flows to compete, under which scenario many of the data flows may probably miss their deadlines. Given the information on the deadlines of synchrophasor data flows, it is possible to determine or predict when the network is congested and to quench some flows at the proper time, so that the remaining flows can meet their deadlines. For example, in Figure 3, multiple flows have the same deadline t_d. As the network becomes congested, the rate allocated to each flow decreases; and if all the flows proceed, then none could meet the deadline. However, quenching one flow ensures that the others finish before the deadline.

A flow quenching algorithm will make use of the deadline information of existing data flows, to decide when and which flows to quench. This has to be accomplished in a dynamic setting, and the stochastic models of data flows can be helpful in formulating and solving this problem.

3 Networked Computation and Fusion of Synchrophasor Data Towards a Secure Smart Grid

3.1 Synchrophasor Data Fusion for Online DSA

Dynamic security assessment (DSA) is an analysis tool that can provide system operators with important information such as voltage, thermal, and transient stability under various probable contingencies. With the real-time or near real-time synchrophasor data collected by PMUs, online DSA can produce prompt decisions for current or impending operating conditions (OCs). Recently, several efforts have been directed towards cost-effective online DSA schemes using synchrophasor data [10, 11]. However, it remains a challenging task, due to the computational complexity incurred by the large size of the contingency list and the massive scale of power systems. First, the combinatorial possibilities of $N-k$ contingencies make it intractable to perform detailed analysis (e.g., power flow analysis and time domain simulations) for all contingencies. In practice, contingency screening schemes (see [11] and the references therein) are used to select the active contingencies that are likely to cause instability, and detailed analysis is performed on only those "active" contingencies. However, the number of active contingencies can still be very large (possibly over thousands for a regional power system [11]). Another challenge for online DSA is the high computational complexity of detailed analysis in processing the high-dimensional measurement data.

3.1.1 A Data-Mining Framework for Online DSA

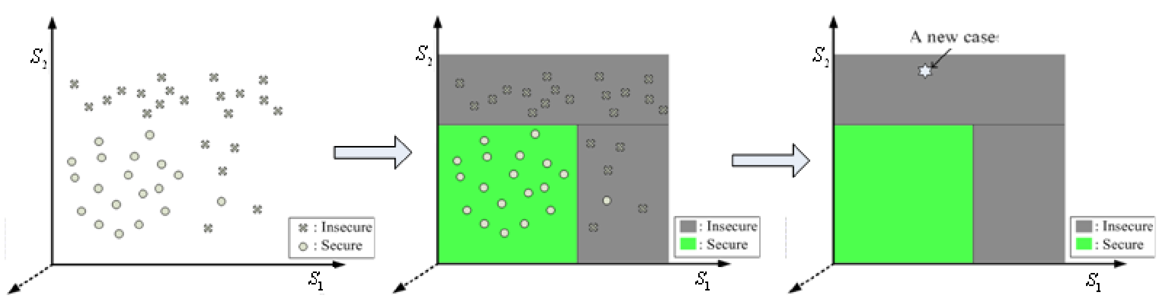


Figure 4: Characterizing the Decision Regions through Data Mining

As illustrated in Figure 4., a cost-effective online DSA scheme developed in [12], characterizes the decision regions through a data mining process, instead of performing detailed analysis for each OC. Specifically, a knowledge base is first prepared through offline exhaustive studies. A classifier is then trained from the knowledge base, and the decision regions are characterized by the classifier. Finally, online DSA simply boils down to mapping the new case into a specific decision region.

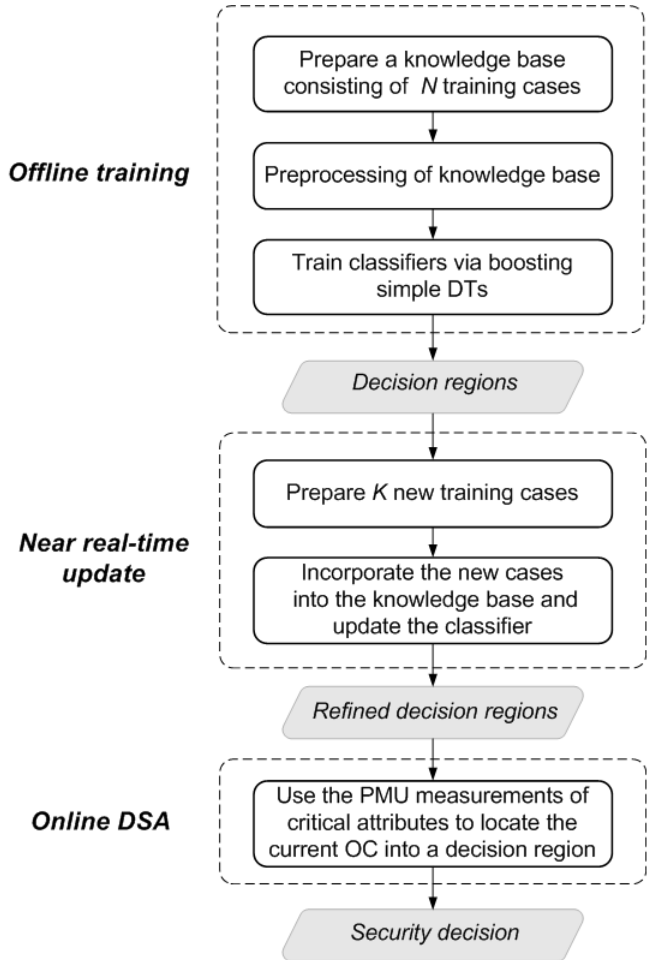


Figure 5: A Data Mining Framework for Online DSA

As depicted in Figure 5. in the offline training stage, a group of N_{OC} predicted OCs are generated for each period T_1 in the next day, based on load forecast and generation schedules. Then, through offline studies on the predicted OCs for a given contingency list C, a knowledge base, consisting of N ($N = N_{OC} \times N_C$) training cases $\{\mathbf{s}_n, d_n\}_{n=1}^{N}$, is used to train the classifiers, where N_C is the number of active contingencies, $\mathbf{s}$ is the attribute vector (the contingency index and the PMU measurements) and d is the security decision. In the near real-time update stage, new data are incorporated into the classifier to refine the decision regions as needed, e.g., when the day-ahead prediction turns out to be inaccurate and new stressed conditions are expected to occur. These new data are created by using past and anticipated OCs, together with new active contingencies. Through the previous two stages, the decision regions for the OCs of the T_1 period can be accurately characterized by the classifiers. In the online DSA stage, the PMU measurements of the critical attributes are collected for each T_2 period, and security decisions are obtained by locating current OC to a decision region. Generally, T_1 is at the scale of hours, and the timescale of T_2 can be on the same order as that of PMU measurements.

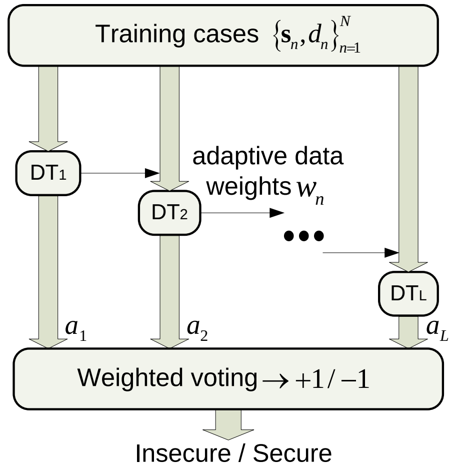


Figure 6: Classifier via Boosting Simple DTs

In the proposed scheme, the classifier for online DSA is obtained via boosting simple DTs, where "boosting" [13] refers to the process of training multiple simple DTs sequentially using adaptive data weights, and combining the simple DTs with proper voting weights to boost the accuracy of the classifier. And simple decision trees are defined as a class of DTs H with a small height J (e.g., J =3). Generally, an individual simple DT might have relatively lower prediction accuracy, but can be less prone to overfitting compared to a fully-grown DT [14]. Further, the classifiers obtained from boosting algorithms are shown to be quite resistant to overfitting. Therefore, boosting simple DTs can produce more accurate classifiers than the approaches which utilize a single DT [10].

Offline Training

The primary objective of offline training is to find a function $F_L : S \rightarrow R$ as weighted voting of L simple DTs, i.e.,

$$F_L(\boldsymbol{s}) = \sum_{l=1}^{L} a_l h_l(\boldsymbol{s}), \boldsymbol{s} \in S$$

where $a_l \in R^+$ is the voting weight of simple DT $h_l \in H$, $l = 1, 2, \cdots L$, and the corresponding binary classifier $F : S \rightarrow \pm 1$, obtained by:

$$F(\boldsymbol{s}) = sign\left(F_L(\boldsymbol{s})\right) = sign\left(\sum_{l=1}^{L} a_l h_l(\boldsymbol{s})\right), \boldsymbol{s} \in S$$

so that the classifier F could fit the given training data.

In order to quantify the performance of the classifier in fitting the training cases $\{\boldsymbol{s}_n, d_n\}_{n=1}^N$, first define the cost function of F_L as follows:

$$C_N(F_L) = \frac{1}{N}\sum_{n=1}^{N}\log_2\left(1+e^{-d_n F_L(\boldsymbol{s}_n)}\right). \quad (1)$$

Then, the offline training problem is formulated as follows:

$$P_F : \min_{\substack{h_1,\cdots,h_L \in H \\ a_1,\cdots,a_L \in R^+}} C_N\left(\sum_{l=1}^{L} a_l h_l\right),$$

It can be seen from (1) that the cost function is convex and lower-bounded. This fact motivates the use of a multi-stage optimization strategy, similar to the line search approach [15]. Specifically, initially with F_0 as a zero function, a simple DT $h_l \in H$ is identified together with a voting weight $a_l \in R^+$, and added to F_{l-1}, i.e., $F_l = F_{l-1} + a_l h_l$ iteratively for $l = 1, 2, \cdots L$. As a result, the classifier via boosting L simple DTs is obtained by solving $P_{DT}^{(l)}$ as follows:

$$P_{DT}^{(l)} : \min_{h_l \in H} \frac{1}{N}\sum_{n=1}^{N}\omega_n^{(l)} 1_{\{d_n \neq h_l(\boldsymbol{s}_n)\}},$$

and the data weights and voting weights are given by:

$$\begin{cases} \omega_n^{(l)} = \dfrac{1}{1+e^{-d_n F_{l-1}(\boldsymbol{s}_n)}} & n = 1,\cdots,N \\ a_l = \arg\min\limits_{a_l \in R^+} g_l(a) & l = 1,\cdots,L \end{cases} \quad (2)$$

According to (2), the cases with smaller margins $d_n F_L(\boldsymbol{s}_n)$ are reassigned with higher weights when used for training the simple DT h_l. Therefore, the simple DT h_l is trained so that correct decisions could be obtained for those cases which are misclassified by previous simple DTs. And for the classifier, by choosing a proper voting weight a_l for h_l, it tries to reduce the overall classification error. In this sense, the classifier generated by the boosting process can fit the training data better as more simple DTs are used. Boosting simple DTs algorithm relates to the multiple optimal DTs algorithm in [10]. Both algorithms aim to enhance the accuracy by using multiple DTs. The major differences are: 1) for boosting, the simple DTs are trained sequentially, in a gradient descent manner, while DTs are usually trained independently in [10]; 2) weighted voting is adopted in boosting and the voting weights are chosen so that the cost function is minimized, while a majority voting is used in [10]. Therefore, the proposed algorithm can guarantee the accuracy of the classifier.

Near Real-Time Update

Suppose that L simple DTs are obtained based on the training data, and in the near real-time update stage, K new case are used to update the classifier one at a time. The algorithm for updating the classifier can be developed in a similar way to the offline training. Specifically, for the k th new case $\{\boldsymbol{s}_{N+k}, d_{N+k}\}$ ($k = 1, \cdots, K$), the classifier is updated by incorporating $\{\boldsymbol{s}_{N+k}, d_{N+k}\}$ with weight $\omega_{N+k}^{(l)}$ into the simple DT h_l by using an incremental tree induction algorithms [16], computing the new voting weight a_l, and then adding it to the classifier.

Summarizing, the classifier via boosting simple DTs can deliver high accuracy on security decision, even when the training data are noisy. Moreover, the low-complexity algorithm for updating the classifier guarantees that the proposed scheme works smoothly in an online environment. More technical details, along with numerical testing on a practical power system, can be found in [12].

3.1.2 Online DSA with Missing PMU Data

In the proposed online DSA scheme, the attributes used in DTs are usually the measurements from multiple locations of the power grid. It is clear that the feasibility of online DSA using DTs would depend on the availability of the synchrophasor data relating to those attributes. In online DSA, however, some synchrophasor data could be unavailable, due to the failures of PMUs and PDCs. Further, the delivery of synchrophasor data can also experience large latency when the communication network is heavily congested, which could also makes synchrophasor data unavailable when online DSA is performed. Therefore, towards a robust online DSA scheme, the issue of missing values has to be taken into account.

In CART [17], missing values are usually handled by using *surrogate splits*. A surrogate split of a decision node of CART is the one which use a different attribute and splitting rule, and "mimics" the original split of the decision node best, i.e., gives the most similar splitting on the set of training data. Specifically, the basic idea of using surrogate splits to handle missing values is to find a surrogate split for each decision node in the tree. Then, in each decision node, if the value of the attribute used in the original split is missing, the corresponding surrogate split is used instead to give a decision. The advantage of using surrogate splits is the subtree corresponding to a decision node could still be used, when the attribute of the original split has a missing value. Obviously, the accuracy of the surrogate split depends on how well it resembles the original split. In online DSA, it is possible that, in a decision node of simple DTs, all the other attributes could not mimic the original split well. In this case, using surrogate split would result in considerable degradation in the accuracy of corresponding subtree.

Motivated by the aforementioned problems, we study online DSA with missing PMU data, in a different line from the surrogate split approach of CART. We observe that, for each original split of DTs, there could be many competitive splits that have comparable accuracy as the original one. Based on this observation, a potential approach to handle missing PMU data in online DSA is: 1) in offline training, multiple subsets of attributes are randomly chosen, and one simple DT is trained by using each of the subsets of attributes; 2) in online DSA, according to the availability of synchrophasor data, the simple DTs without missing values are used to obtain a classifier via boosting.

In the above approach, randomized attribute subsets are used for two purposes: 1) to reduce the impact of missing PMU data (i.e., most simple DTs could still be used, when only several measurements are missing), and 2) to reduce the complexity of training simple DTs. Despite these advantages of this approach, several specific issues need to be addressed carefully, including the choice of size of the subset of attributes, the number of simple DTs trained offline, and the complexity therein.

3.1.3 Modeless Assessment

Another possible approach to security assessment is the "modeless" approach where Thevenin Equivalents as seen by key lines are computed from PMU data and used to estimate margins to security violations using fundamental criteria such as thermal, voltage and angles across the system. With PMU data being the primary source of creating the equivalents, the assessment would be able to track network and load changes almost instantaneously and thereby track margins to critical values in real time.

3.2 Synchrophasor Data Fusion for Fault Detection and Localization

One of the primary concerns on the reliability of power systems has been the issue of large-scale fault events and their impacts on the overall stability of the power grid. However, today's power systems are not equipped with sufficient fault diagnosis mechanisms against various malicious attacks and natural physical events [18]. Thus, there is an urgent need for quickly assessing the impact of fault events so that corrective actions can be taken promptly to avoid cascading events.

It is known that fault diagnosis of transmission lines is challenging [18], due to the massive scales, complex system uncertainty and inevitable measurement errors, and deterministic approaches would not work well in some practical scenarios due to many stochastic events in power systems. In light of the stochastic nature of power systems, the bus injections and branch flows could be volatile across various time scales, which would be especially true in the smart grid which is supposed to integrate a large number of distributed generations. With this insight, we propose to use probabilistic graphical models for modeling the spatially correlated data from PMUs, and use statistical hypothesis testing for the task of fault diagnosis.

3.2.1 A GMRF Model for Synchrophasor Data

It follows from the DC power flow model that the phasor angle at bus *i* could be represented as:

$$\theta_i = \sum_{j \neq i} c_{ij} \theta_j + \frac{1}{\sum_{j \neq i} b_{ij}} P_i , \tag{3}$$

where θ_i and θ_j denote the phasor angles at bus i and j, respectively, P_i denotes the flow injection to bus i, b_{ij} denotes the inverse of line inductive reactance, and:

$$c_{ij} = b_{ij} / \sum_{j \neq i} b_{ij} .$$

Following the probabilistic power flow approaches [19], we observe that the phasor angles at non-slack buses could be approximately modeled as Gaussian random variables. Let $\boldsymbol{\theta}_{-i}$ denote the sites except θ_i, then by (3) the conditional distribution of θ_i could be specified in the form of conditional auto-regression (CAR) model [20]:

$$\theta_i \mid \boldsymbol{\theta}_{-i} \sim N\left(u_i + \sum_{j \neq i} r_{ij}\left(\theta_j - u_j\right), 1\right)$$

It is shown in [20] that under mild conditions, the joint distribution of the GMRF $\boldsymbol{\theta}$ follows $N\left(\boldsymbol{u}_i, \boldsymbol{J}^{-1}\right)$, with the information matrix $\boldsymbol{J} = \boldsymbol{I} - \boldsymbol{R}$ and $\boldsymbol{R} = \left[r_{ij}\right]$ as the matrix consisting of partial correlation coefficients. Note that for each θ_i its partial correlation coefficients $\left\{r_{ij}, j \neq i\right\}$ are proportional to $\left\{c_{ij}, j \neq i\right\}$. We have a few key observations in order. 1) The dependency graph of phasor angles agrees with the topology of power systems; 2) As the susceptance matrix $\boldsymbol{B} = \left[b_{ij}\right]$, the partial correlation matrix $\boldsymbol{R}$ also reflects the electrical distance between buses. Intuitively, the reduction in the electrical connectivity of buses would result in less partially-correlated phasor angles; further, r_{ij} vanishes if a line outage takes place between buses i and j.

3.2.2 Decentralized Network Inference Using Synchrophasor Data

Let E' be the edge set, Σ' be the covariance matrix of GMRF, and $\boldsymbol{R}'$ be the partial correlation matrix when the power system is under normal conditions. When fault events take place, some edges might fail and the partial correlation matrix of GMRF would change. Mathematically, the proposed fault detection and localization approach boils down to hypothesis testing on the changes of partial correlations, with null hypothesis given by $H_0 : \left\{\text{there is no change in } r_{ij}, \forall \left\{i, j\right\} \in E'\right\}$.

One main difficulty in performing the above hypothesis testing originates from the fact that the observations of $\left\{\theta_i, \theta_j\right\}$ could only lead to the knowledge of the correlation coefficient ρ_{ij} between θ_i and θ_j, rather than the change of r_{ij}. Accordingly, it is necessary to obtain a complete estimate of $\boldsymbol{J}$. Another challenge is the requirement on the sparsity of $\hat{\boldsymbol{J}}$, the estimate of $\boldsymbol{J}$. Since the inverse of sample covariance matrix $\hat{\Sigma}^{-1}$ might not have the same sparsity as $\boldsymbol{J}$, due to noisy observations or a small number of samples, thus it is critical for $\hat{\boldsymbol{J}}$ to have desired sparsity.

In related work, the estimation of the information matrix $\boldsymbol{J}$ of GMRF is often treated as a constrained optimization problem which maximizes the likelihood [21]:

$$\begin{aligned} &\text{maximize} \quad \log|\hat{\boldsymbol{J}}| - tr\left(\hat{\boldsymbol{J}}\hat{\Sigma}\right) \\ &\text{subject to} \quad \hat{\boldsymbol{J}}_{ij} = 0, (i, j) \notin E' \end{aligned}$$

The solution to the above problem often requires centralized computation and global observations. As noted in [21], the computational complexity could be very high for large-scale problems but existing algorithms are not scalable. Worth noting is that the estimation of $\boldsymbol{J}$ generally requires the number of observations at least comparable to the size of $\boldsymbol{\theta}$.

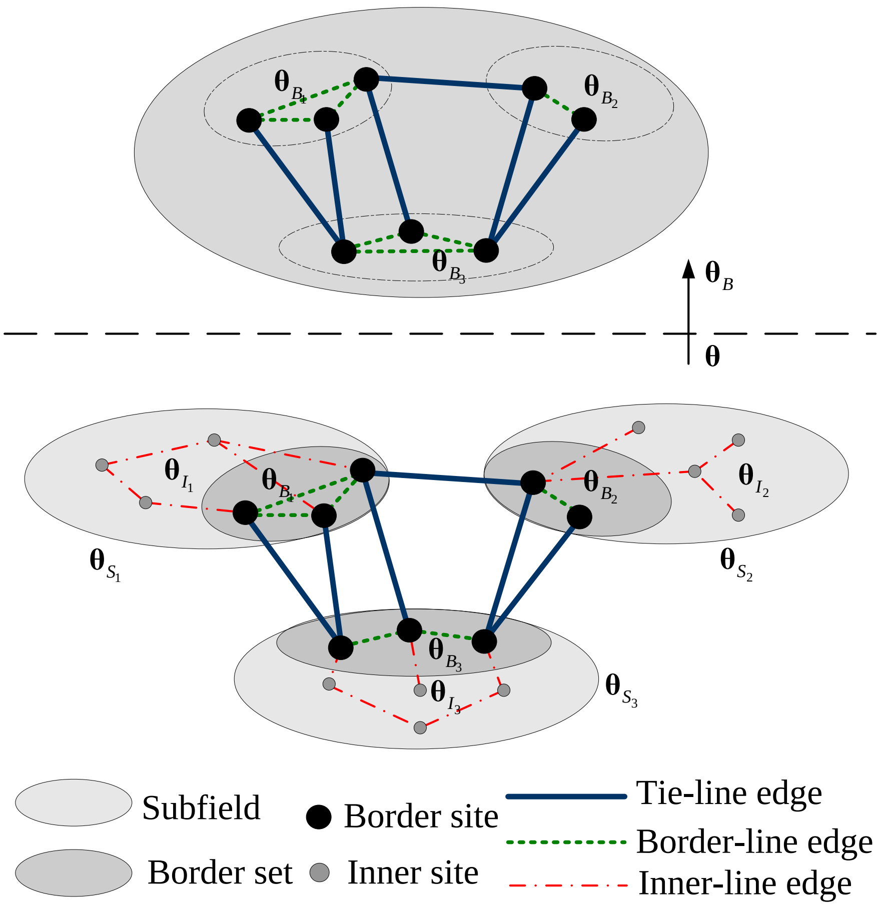


Figure 7: Two-Scale Decomposition of GMRF

To tackle the aforementioned challenges, we devise a scheme of multi-resolution transform on GMRF which could reconstruct the GMRF from the subfields, and propose a multi-scale message-passing procedure to find a global solution for fault localization. Specifically, for a power system consisting of several sub-systems, we decompose the hypothesis testing problem into multiple sub-problems, in which the inference can be carried out based on local observations. We note that a direct decomposition of the GMRF, by grouping the sites into K disjoint sub-fields, would not completely capture the dependence structure across the subfields. With this insight, we construct an additional sub-field for each level, as illustrated in Figure 7. In a nutshell, the sites are grouped into border sites and inner sites, of which the latter are not connected to the other sub-fields. Furthermore, there are three classes of edges: tie-line edges which connect different sub-fields, border-line edges which connect border sites of the same sub-field, and inner-line edges which have at least one end as inner site.

Algorithm 1: Network inference of $\boldsymbol{J}$ via fusion of synchrophasor data

Local estimation: Estimate the information matrices of all the sub-fields based on local measurements, by solving the sub-problem using the dependency graph of $\boldsymbol{\theta}_{S_k^{(l)}}$.

Down-top message passing: For $l = 1, 2, \cdots, L-1$, the inference centers of $\boldsymbol{\theta}_{S_f^{(l)}}$, $f \in F(k,l)$, submit $\hat{\boldsymbol{J}}_f^{(l)}$ to that of $\boldsymbol{\theta}_{S_k^{(l)}}$.

Top-down reconstruction: For $l = L-1, L-2, \cdots, 1$, the inference center of $\boldsymbol{\theta}_{S_f^{(L)}}$ reconstruct $\hat{\boldsymbol{J}}^{(l)}$ from $\hat{\boldsymbol{J}}^{(l+1)}$ and $\hat{\boldsymbol{J}}_k^{(l)}$, $k = 1, 2, \cdots, K^{(l)}$.

Top-down message passing: The inference center of $\boldsymbol{\theta}_{S^{(L)}}$ broadcasts $\hat{\boldsymbol{J}}$, i.e., $\hat{\boldsymbol{J}}^{(1)}$ to the inference centers of all the sub-fields.

In solving the $K+1$ sub-problems, a key challenge is that the graphs of subfields no longer agree with the system topology. Indeed, as discussed in [22], the decomposition on MRF would introduce new edges into the graphs of the subfields. For GMRF, the information matrices of the subfields would have different sparse patterns and non-zero entries from the corresponding diagonal blocks of $\boldsymbol{J}$. Therefore, the knowledge of local information matrix is not sufficient to identify all of the faults in the concerned sub-problem. To tackle the above challenge, we first rigorously prove that for the information matrices of the subfields and $\boldsymbol{J}$, the entries corresponding to the inner sites, inner-line edges, and tie-line edges remain the same. Then, we propose to employ message-passing between subfields to "recover" the information about the border sites and border-line edges, lost due to decomposition. We first studied two-scale decomposition of GMRF, and proved that $\boldsymbol{J}$ could be reconstructed from the information matrices of the subfields through message-passing. Further, we show that this two-scale decomposition can be extended to the multi-scale decomposition of GMRF.

Based on the above idea, we present the decentralized network inference algorithm. Suppose all the buses of the power system are observable, we first perform a multiscale decomposition on $\boldsymbol{\theta}$ based on the hierarchical topology of the power system. Once the estimates of the information matrices of sub-fields are obtained, a complete $\hat{\boldsymbol{J}}$ could be reconstructed from the estimated information matrices of sub-fields. For each scale l ($l = 1, 2, \cdots, L$), we assume that there is an inference center at each sub-field $\boldsymbol{\theta}_{S_k^{(l)}}$ ($k = 1, 2, \cdots, K$). Let $F(k,l)$ be the collection of the indices of the sub-fields that are located at the lower scale of $\boldsymbol{\theta}_{S_k^{(l+1)}}$. Then, the procedure of the decentralized estimation of the information matrix is summarized in Algorithm 1.

3.2.3 Conclusion

In summary, the proposed network inference approach could effectively detect and localize the faulted transmission lines, and the decentralized algorithm can achieve comparable performance with a centralized one. Moreover, the proposed multi-scale data fusion scheme could effectively address the following potential issues in practice: 1) global data is not available in some cases, i.e., when the utilities cannot share the synchrophasor data due to confidentiality constraint, or when the synchrophasor data formats are incompatible (e.g., different sample rates); 2) the gathering or processing of the global data cannot be accomplished in a timely manner, due to the constraints on communication bandwidth and computation capacity.

4 Robust Architecture for Smart Grids: Cascading Failures and Interdependence between Communication Network and Power Grids

The power grid and the synchrophasor communication system depend on another to provide proper functionality. This interdependence has motivated us to study the cascading phenomena between the two systems, i.e., in the event of cyber/physical attacks, node failures in the communication/power system may result in a cascade of failures, which can be devastating since they can trigger the failures of many more components in both systems and cumulatively progress into the potential collapse of the entire system.

Under this framework, we have explored mechanisms to improve the robustness of interdependent systems against cascading failures [23]. Specifically, in [23], we exploited the topology information to improve the robustness of the entire system against cascading failures, and developed a “regular” allocation strategy that allots inter-network links uniformly across all nodes. Our findings reveal that from a network resilience perspective, the proposed regular allocation strategy yields a significant gain compared to conventional random allocation strategy. We expect that our findings can help understanding and designing the topology of the entire system.

4.1 Regular Allocation of Inter-Edges

We consider a cyber-physical system consisting of two interacting networks, namely network A and network B, and assume that they are of the same size N, with vertex sets denoted by $\{v_1, \cdots, v_N\}$ and $\{v'_1, \cdots, v'_N\}$, respectively. We refer the edges connecting nodes within the same network as *intra-edges*, and those connecting nodes from two different networks as inter-edges. Figure 8 illustrates the regular allocation of inter-edges, i.e., each node in A and B has exactly k inter-edges.

We are particularly interested in understanding the network robustness in a cascade of failures. Specifically, in the dynamics of cascading failures, we assume that a node is “functioning” at stage $t+1$ if the following two conditions are satisfied simultaneously:

- The node belongs to the giant component of its own network;
- The node has at least one inter-edge from nodes functioning at Stage t of the cascading failures, in the other network.

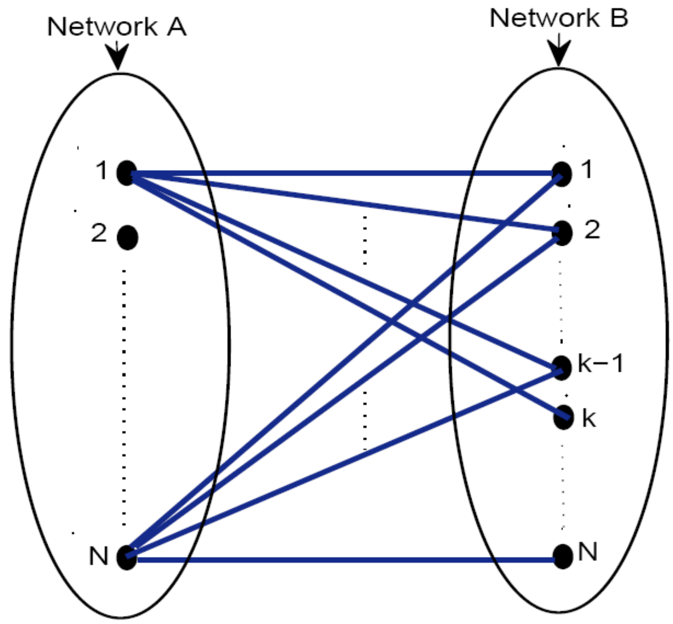


Figure 8: The Set-Up of the Inter-Network Connections

Moreover, we call the giant component composed of functioning nodes a *functioning giant component*. With this setup, we are infested in analyzing the dynamics of cascading failures in two interacting networks with the regular allocation of inter-edges.

4.2 Analysis of Cascading Failures

Suppose after the first stage of failures, a fraction $1-p$ of the nodes in network A stop functioning. Due to the interdependence, this "shrinking" phenomenon of functioning nodes in network A would trigger node failures in network B, which is called the second stage of failures. This propagation of cascading failures continues in such a recursive manner, which eventually leads to either 1) *a mutually connected functioning giant component* or 2) *complete dysfunctioning of the entire system consisting of two networks*.

A principal objective of this study is to characterize the ultimate fractions of the giant components, denoted by $|A_\infty|$ and $|B_\infty|$, and the critical threshold p_c, which serves as an important measurement of the system robustness. To that end, we will use the technique of generating functions [24] to quantify the sizes of functioning giant components in two networks at each stage i, denoted as $|A_i|$ and $|B_i|$. The key notations in the calculation can be found in Table II.

Table II: Key Notation in the Analysis of Cascading Failures

A_i , B_i	The functioning giant component in A (resp. B) at stage i
p_{A_i} , p_{B_i}	The fractions of functioning giant components at stage i , $\lvert A_i \rvert = p_{A_i} N$, $\lvert B_i \rvert = p_{B_i} N$
p'_{A_i} , p'_{B_i}	The equivalent remaining fraction of A (resp. B) at stage i
$\bar{A}_i$, $\bar{B}_i$	The remaining fraction of nodes in A (resp. B) with at least one inter-edge at stage i .

Along the process of recursive "shrinkage", one can construct the sequence of the functioning giant components at different stages of the cascading failures: $A_1 \supset A_3 \supset \cdots \supset A_{2m+1}$ and $B_2 \supset B_4 \supset \cdots \supset B_{2m}$. In particular, it is easy to check that $p'_{A_1} = p$ and the fraction of giant components can be obtained by recursive relations:

$$
\begin{aligned}
p_{B_{2i}} &= p'_{2i} P_B\left(p'_{B_{2i}}\right), \\
p'_{B_{2i}} &= 1-\left(1-pP_A\left(p'_{A_{2i-1}}\right)\right)^k, \\
p_{A_{2i+1}} &= p'_{A_{2i+1}} P_A\left(p'_{A_{2i+1}}\right), \\
p'_{A_{2i+1}} &= p\left(1-\left(1-P_B\left(p'_{B_{2i}}\right)^k\right)\right),
\end{aligned}
$$

for each $i = 1, 2, \cdots, m$, where $P_A(p)$ denotes the fraction of the giant component in a random subgraph that occupies p fraction of the nodes in network A, and $P_B(p)$ denotes that for network B.

This recursive process stops at the "equilibrium point", where we have $p'_{B_{2m-2}} = p'_{B_{2m}}$ and $p'_{A_{2m-1}} = p'_{A_{2m+1}}$, so that neither network A nor network B will fragment further. By setting $s = p'_{A_{2m+1}}$ and $t = p'_{B_{2m}}$, we obtain the set of equations:

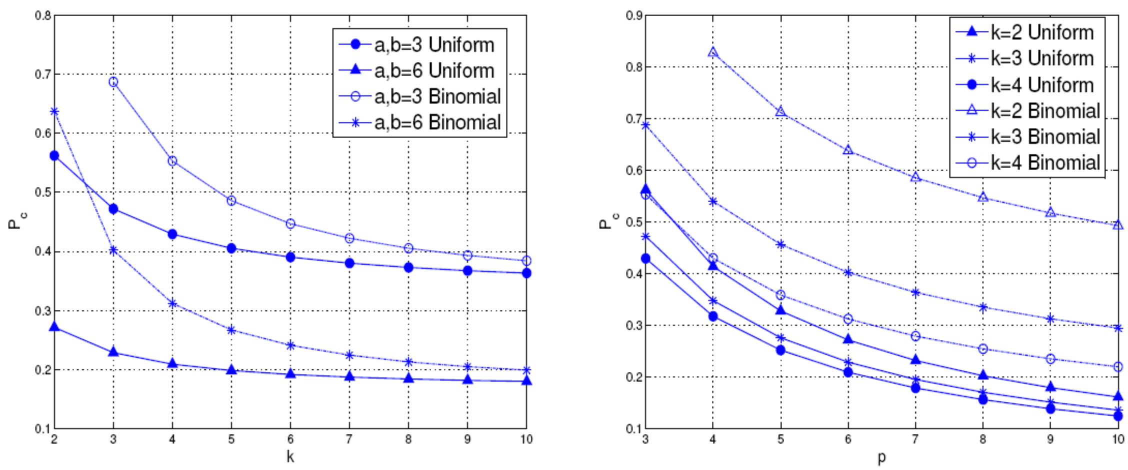
$$
s = p\left(1-\left(1-P_B(t)\right)^k\right) \quad t = 1-\left(1-P_A(s)\right)^k
$$

Furthermore, the fractions of nodes that appear in the giant components are given by $P_{A_\infty} = sP_A(s)$ and $P_{B_\infty} = sP_B(s)$, which holds for networks with arbitrary intra-degree distributions.

4.3 Regular Allocation vs. Random Allocation

We then compare the robustness performance corresponding to the proposed regular allocation with that under random allocation [25], in term of critical threshold p_c. We consider two Erdos-Renyi networks of the same size N, and assume that their average intra-degrees are a and b, respectively. For fair comparison, the inter-degree at each node in the random allocation follows an i.i.d. Binomial distribution with mean k, and on the other hand, the number of inter-edges per node is fixed at k in the regular allocation.

The values of critical threshold p_c corresponding to both allocation schemes are compared under a variety of conditions. Figure 9(a) depicts p_c as a function of mean inter-degree k, for various values of $a = b$, whereas Figure 9(b) depicts the variation of p_c with respect to $a = b$ for different k values. It can be seen that the regular allocation yields a much smaller p_c than the random allocation. Therefore, a more robust system can be obtained by regular allocation.



(a) (b)

Figure 9: The Critical Values of p_c with Average Inter-Degree Equal to k

We believe that the drastic improvement in robustness against cascading failures can be attributed to the following two reasons. 1) In the random allocation, there always exists a non-negligible fraction of nodes with no inter-edge support from the other network. Clearly, the regular allocation scheme promises a guaranteed support in terms of inter-edges, for all nodes in both networks. 2) In the absence of the intra-degree distribution information, the regular allocation allots the inter-edges uniformly in a deterministic manner and can always yields smaller p_c than any unequal allocation strategy, which corresponds to a specific realization of random allocation.

In summary, assuming no information of network intra-degree distributions is available, our study reveals that compared to random allocation, the proposed regular allocation of inter-edges yields a significant gain in terms of network robustness. We expect that the topology information can be exploited to improve further the robustness of cyber-physical systems against cascading failures.

We believe that the studies we initiated here on robust interconnecting architecture for smart grids scratch only the tip of the iceberg. There are still many questions remaining open to design cyber-physical systems, in a manner with tight conjoining and coordination.

References

[1] GE Power Systems Engineering. *Assessment of applications and benefits of phasor measurement technology in power systems*. EPRI Final Report, April 1997.

[2] Ilic, M. *From hierarchical to open access electric power systems*. Proceedings of the IEEE, vol. 95, no. 5, pgs. 1060–1084, May 2007.

[3] Xie, Z.; G. Manimaran, V. Vittal, A. Phadke, and V. Centeno. *An information architecture for future power systems and its reliability analysis*. IEEE Transactions on Power Systems, vol. 17, no. 3, pgs. 857–863, August 2002.

[4] NASPI PMU Registry. Available at: https://naspi.tva.com/pmuregistry.

[5] Fahid, K.; P. Gopalakrishnan, and S. *Cherian. Phasornet a high performance network communications architecture for synchrophasor data transfer in wide area monitoring, protection and control applications*. Bulk Power System Dynamics and Control - VII. Revitalizing Operational Reliability, 2007 iREP Symposium, pgs. 1–4, August 2007.

[6] Hopkinson, K.; G. Roberts, X. Wang, and J. Thorp. *Quality of service considerations in utility communication networks*. IEEE Transactions on Power Delivery, vol.24, no.3, pgs.1465-1474, July 2009.

[7] Wilson, C.; H. Ballani, T. Karagiannis, and A. Rowstron. *Better Never than Late, Meeting Deadlines in Datacenter Networks*. Proceedings of Sigcomm, 2011.

[8] Bakken, D.; A. Bose, C. Hauser, D. Whitehead, and G. Zweigle. *Smart generation and transmission with coherent, real-time data*. Proceedings of the IEEE, vol. 99, no. 6, pgs. 928–951, June 2011.

[9] Gjermundrod, H.; H. Gjermundrod, D. Bakken, C. Hauser, and A. Bose. *Gridstat: A flexible qos-managed data dissemination framework for the power grid*. IEEE Transactions on Power Delivery, vol. 24, no. 1, pgs. 136–143, January 2009.

[10] Diao, R.; K. Sun, V. Vittal, R. O'Keefe, M. Richardson, N. Bhatt, D. Stradford, and S. Sarawgi. *Decision tree-based online voltage security assessment using PMU measurements*. IEEE Transactions on Power Systems, vol. 24, no. 2, pgs. 832–839, May 2009.

[11] Chiang, H. D., J. Tong, and Y. Tada. *On-line transient stability screening of 14,000-bus models using TEPCOBCU: Evaluations and methods*. in Power and Energy Society General Meeting, 2010 IEEE, pgs. 1–8, July 2010.

[12] He, M.; J. Zhang, and V. Vittal. *A data mining framework for online dynamic security assessment: Decision trees, boosting, and complexity analysis*. to appear, IEEE PES Conference on Innovative Smart Grid Technologies 2012.

[13] Mason, L.; J. Baxter, P. L. Bartlett, and M. R. Frean. *Boosting algorithms as gradient descent*. Neural Information Processing Systems, pgs. 512–518, 1999.

[14] Freund, Y.; and R. E. Schapire. *A decision-theoretic generalization of on-line learning and an application to boosting*. Journal of Computer and System Sciences, vol. 55, pgs. 119–139, 1997.

[15] Box, M. J.; D. Davies, and W. H. Swann. *Non-Linear optimisation Techniques*. Oliver and Boyd, 1969.

[16] Utgoff, P. E.; N. C. Berkman, and J. A. Clouse. *Decision tree induction based on efficient tree restructuring*. Mach. Learn., vol. 29, pgs. 5–44, October 1997.

[17] Breiman, L.; J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*, 1984.

[18] Leon, R.; V. Vittal, and G. Manimaran. *Application of sensor network for secure electric energy infrastructure*. IEEE Transactions Power Delivery, vol. 22, no. 2, pgs. 1021–1028, April 2007.

[19] Heydt, Gerald T.; and Peter W. Sauer. *Probabilistic Methods for Planning and Operational Analysis*. The Electric Power Engineering Handbook second edition – Power Systems (Leo Grigsby Editor), CRC Press - Taylor and Francis group, Chapter 20, pgs. 20-1 to 20-10, 2007.

[20] Rue, H.; and L. Held. *Gaussian Markov Random Fields: Theory and Applications*. Chapman & Hall/CRC, 2005.

[21] Dahl, J.; V. Roychowdhury, and L. Vandenberghe. *Covariance selection for nonchordal graphs via chordal embedding*. Optimization Methods and Software, vol. 23, no. 4, pgs. 501–520, 2005.

[22] Perez, P.; and F. Heitz. *Restriction of a Markov random field on a graph and multiresolution statistical image modeling*. IEEE Transactions Information Theory, vol. 42, no. 1, pgs. 180–190, January 1996.

[23] Yagan, O.; D. Qian, J. Zhang, and D. Cochran. *On allocating interconnecting links against cascading failures in cyber-physical networks.* Computer Communications Workshops (INFOCOM Workshops), 2011 IEEE Conference on, pgs. 930–935, April 2011.

[24] Newman, M. E. J. *Spread of epidemic disease on networks*. Physical Review E, vol. 66, 2002.

[25] Shao, J.; S. V. Buldyrev, S. Havlin, and H. E. Stanley. *Cascade of failures in coupled network systems with multiple support-dependent relations*. Computing Research Repository, vol. abs/1011.0, 2010.