Merging the Temporal and Spatial Aspects of Data and Information for Improved Power System Monitoring Applications

MLADEN KEZUNOVIC, FELLOW, IEEE, AND ALI ABUR, FELLOW, IEEE

Invited Paper

Many power system applications, particularly those related to monitoring, control, and protection, depend on the use of field data recorded by intelligent electronic devices (IEDs) in substations. With the introduction of IEDs in the last 20 years, the amount and type of data collected in substations have dramatically increased. This paper explores the benefits of data integration and information extraction achieved by merging temporal and spatial considerations. In the new approach, time and space issues are treated consistently across all IEDs, and as a result new applications with improved performance characteristics can be defined. As an illustration, two new monitoring functions are discussed in detail: 1) automated fault analysis and 2) hierarchical state and topology estimation. Discussion ends with a description of an implementation framework that will allow development of other functions that will benefit from the merger of time and space considerations.

Keywords*—*Communication protocols, fault diagnosis, spatial data structures, state estimation, time measurement.

I. Introduction

This paper focuses on the issues of time and space as they relate to data integration and information extraction. While the issues of time and space have been well known for many years and have been dealt with in different power system applications, they may not yet have been fully utilized and explored. The reason for their consideration in a new data integration and information exchange approach is their importance and practical consequence in a modern monitoring, control, and protection system. Two events have created the need to explore the new approach: the development of intelligent electronic devices (IEDs) with improved data acquisition and communication capabilities and the advent of computational algorithms that can be hierarchically distributed across substations. The core of the new approach involves a combination of increased data acquisition capability and enhanced computational performance of the data-processing equipment used in substations. This allows implementation of a hierarchically structured, distributed processing oriented power system monitoring, control, and protection system that will be capable of performing its tasks in a much more accurate, reliable, and efficient way than the existing system.

In the early days, the protection concepts were established as a decentralized automation that has been focused on local operation, which is based on local measurements and local actions [1]. In the 1960s, modern power system monitoring and control solutions were introduced with centralized databases for supervisory control and data acquisition (SCADA) systems [2]. The two concepts—decentralized protection and centralized monitoring and control—were implemented using separate equipment and communication facilities. This approach was based on a paradigm that separated the power system states into "normal," "alert," "emergency," and "restorative." Hence the corresponding equipment was designed to deal with a specific power system operating state [3]. This paradigm was further explored by several researchers and practicing engineers, and led to the conclusion that separation of equipment and algorithms is justified since the various operating states of the power system needed different temporal and spatial approaches to data integration and information extraction [4], [5].

This paper offers a different view based on a recent study of how the new equipment, communication capabilities, and algorithmic developments may enhance the monitoring, control, and protection requirements and implementations [6]. As a result, the researchers recognized that system-wide protection as well as decentralized monitoring and control may

Manuscript received October 1, 2004; revised June 1, 2005.

The authors are with the Department of Electrical Engineering, Texas A&M University, College Station, TX 77843-3128 USA (e-mail: kezunov@ee.tamu.edu; abur@ee.tamu.edu).

Digital Object Identifier 10.1109/JPROC.2005.857497

be needed to address dynamic changes in system loading and other "normal" contingencies as well as emergencies caused by faults and cascading events. The core of this new approach is a unique way of considering the temporal and spatial aspects of data integration and information extraction, as well as its uses for monitoring, control, and protection applications.

The paper starts with simple explanations of the temporal and spatial considerations in the context of monitoring, control, and protection. It then illustrates the new approach to temporal and spatial considerations when implementing an automated fault analysis and hierarchical state and topology estimation. Finally, the paper discusses how a new approach based on a merged temporal and spatial consideration of data and information may affect the new generation of monitoring, control, and protection solutions.

II. Temporal and Spatial Considerations

A. Temporal Considerations

As a background, we will review now a number of considerations related to the temporal aspects of power systems.

1) Time as a Reference for Correlating Power System Events: Monitoring, control, and protection functions require knowledge of the instant of time when a given event has occurred. Examples of such events are occurrence of the fault or opening of a breaker. In some instances, a *relative time* is all that is needed. For example, if a fault has caused a relay to issue a trip command to a breaker, it is important to understand the time sequence between the incident, the relay action, and the breaker operation. Besides, an *absolute* time plays a role, as well. We need to know when various disturbances have happened in terms of actual time so that we can place many different events in the same and adjacent systems (which may be associated with the initial event) in a time sequence.

2) Time as a Reference for Signal Sampling: Various measurements in the power system are performed by IEDs. They convert the measurements to samples by performing analog-to-digital (A/D) conversion at the time the measurement is taken. The samples are taken by a sampling and hold (S/H) circuit, and then the A/D converts sample into a computer word, known as data. The clock signal used for initiating the S/H circuit can be applied simultaneously (synchronously) for all the measured channels or sequentially as each channel is measured (scanned). Recovery of the information from data samples depends heavily on whether the signals were sampled synchronously or scanned. For example, it is possible to recover the phase angle between different phases in a three-phase circuit if synchronous sampling was performed, but it may not be possible to recover it if the signal was scanned.

3) Time as a Reference for Waveform Representation: Many of the analysis functions for monitoring, control, and protection require that the analog waveforms of current and/or voltage be analyzed either as *time-domain functions* or *phasors*. The time domain representation is important when waveforms experience transient behavior while the phasor representation is sufficient for steady state conditions. In both instances, how time is represented is important, which leads to either an accurate representation of a waveform at any instant in time or an approximation of the waveform with a phasor.

4) Time as a Reference for a Control Action: An ac power system acts as one big machine that couples many generators and loads in *a synchronous* operation. The control actions that are taken to disconnect or connect various components of the system from or to the overall system may have to be performed at a given instant of time so that the synchronized operation is not disturbed. Connecting transmission lines or generators to the rest of the system at the instants that will cause a major departure from synchronous conditions may cause serious equipment damage. Thus, we must carefully consider the control sequence.

B. Spatial Considerations

It is necessary to understand the various aspects of spatial considerations, as well.

1) Space as a Reference for Power Apparatus Locations: Power systems typically are spread across wide areas. They consist of transmission lines connecting substations, generators providing power, and load centers consuming power. As we implement monitoring, control, and protection functions, we target *specific areas* of a power system such as the individual power apparatus, substations, selected regions, the entire system, or the areas surrounding the original system. For various events and corresponding operating conditions, only specific power system components involved in the monitoring, control, or protection actions are related to the events.

2) Space as a Reference for Location of Decision-Making Equipment: It is well recognized that the equipment may be placed in a variety of locations, such as a substation switchyard, a control house in a substation, a power plant, a load center, a regional center, the main control center for the entire power system, the independent system operator (ISO) center, or the regional transmission operator (RTO) center. The location of the equipment and the type of data and information to be provided depends on the communication infrastructure built to connect various parts of the decision-making monitoring, control, and protection system. This translates into an outcome in which some of the equipment is being distributed and some centralized.

3) Space as a Reference for Data Processing and Information Extraction: The computers that control today's power system may be placed at various locations throughout the system. Data collected from substations and power plants may be processed locally ("decentralized processing") to extract the required information, or it may be sent to a centralized location to perform such processing. The algorithms for monitoring, control, and protection may be located on different computers, which often are spread across several physical locations. Hence, the main options are to have either localized (often denoted as decentralized) or centralized processing. Distributed processing is a variation in which the data processing and information extraction algorithms are performed at several locations while the decision-making function is either localized or centralized.

4) Space as a Reference for Execution of a Command: A control command can be executed automatically by equipment or manually by an operator. The exact location at which the command is executed influences both the speed of execution and the total area of the power system that is affected. In the case of protective relaying, the area in which the command is executed is most often the same area in which the data is collected. In some rare cases, the control action is initiated at the same place the data is collected, but the execution takes place in an adjacent substation (transfer trip action). In either case, the control action is executed automatically. The manual initiation of a command may be local or remote with respect to the point where the actual control action takes place.

III. Data Integration and Information Extraction

This section gives an explanation of how temporal and spatial considerations may affect data integration and information extraction. To contrast the traditional approaches to the new approach proposed by the authors, two cases of data integration and information extraction are discussed: substation-wide and system-wide.

A. Substation-Wide Data Integration and Information Extraction

Several aspects of the temporal and spatial consideration at the substation level are important. We will focus on the two most commonly discussed aspects: signal sampling and feature extraction.

1) Substation-Wide Signal Sampling: Traditionally, the vendor independently selects and implements signal sampling in commonly used IEDs. This means that the sampling rates, synchronization of sampling among different signal channels, and A/D conversion characteristics may be quite different from one IED to the next. This creates a major problem if the data from different IEDs is to be integrated since the samples are not "aligned" in time. The datat processing required to correct the sample "alignment" may be quite elaborate. Another approach, advocated in this paper and utilized to enhance the existing applications, suggests that temporal characteristics of the signal sampling be uniformly applied across substation IEDs. This translates into a common approach to signal sampling across a given substation as well as between adjacent substations. Besides the selection of a common sampling rate, the sampling synchronization should be controlled by a common reference such as a GPS time synchronization signal. This makes data integration much easier, since all the signals are being sampled at the same instant. The temporal relationship among different signals makes the spatial correlation much easier, since all the signals taken at the same instant of time can be easily utilized for various applications irrespective of the distance between measurement points.

2) Substation-Wide Feature Extraction: Two types of feature extraction are mentioned here with the goal of distinguishing current practice from the new approach proposed by the authors: signal feature extraction and substation topology extraction. A typical approach to signal feature extraction in current IEDs is to recover the signal phasor using samples taken during a full signal cycle or samples taken at certain time intervals during a number of signal cycles. While the phasors extracted this way may be used for many applications, taking samples synchronously across all the IEDs is more beneficial, especially if the feature includes not only a phasor but also a vector of signal samples containing an arbitrary number of samples. Current substation applications do not require knowledge of the substation topology. If this information were extracted from IEDs, it would greatly enhance substation-wide applications.

B. System-Wide Data Integration and Information Extraction

We will now discuss how temporal and spatial considerations affect the present practice and future possibilities regarding signal sampling and feature extraction for system-wide applications.

1) System-Wide Signal Sampling: Traditionally, signal sampling for system-wide applications has not had any temporal requirements for "alignment" of samples or phasors. The most common example of this approach is implementation of the current SCADA system implementation. The remote terminal units (RTUs) of a common SCADA implementation today will not have synchronized sampling, but will be based on signal scanning. As a result, phasors recovered from such samples are not on the same temporal scale. Hence we cannot establish a time correlation between phasors taken at different spatial locations. Monitoring and control of large systems containing multiple areas can be accomplished provided that boundary buses connecting individual areas are sufficiently metered. Such metering will no longer be required if all internal area measurements are synchronized system wide using a GPS receiver. This will enable the decentralized treatment of individual area monitoring and control problems while accounting for system-wide dependencies. This important consideration in multiarea system representation for monitoring and control applications is directly affected by temporal considerations. Spatial treatment of the areas in the multiarea systems is affected at the same time.

2) System-Wide Feature Extraction: The features of the measured signals and topology of the entire system are important. The feature extraction of analog signals at the system level is almost always associated with the calculation of phasors. Most advanced IED designs today are capable of measuring phasors across the entire power network using synchronously taken samples leading to an accurate "alignment" of phasors. For that reason, new IEDs specifically developed for synchronous phasor measurements are being widely deployed and new networks of such IEDs are created to integrate the data and extract relevant information. This paper discusses how such measurements may be utilized to enhance the existing applications. Finally, the extraction of information related to the system topology is also worth discussing. Today's system-wide applications do not focus on close tracking of the system topology. This is manifested through the use of reduced system models and occasional processing of the bus/breaker information to obtain the reduced one-line diagrams. This paper illustrates one possible implementation of a topology tracking system that exploits the capability of IEDs to provide detailed breaker level measurements at the substations. This is another example of the potential benefits provided by a substation-wide data integration and information extraction system.

IV. Automated Fault Analysis

The importance of the temporal and spatial considerations will be discussed using new automated fault analysis application [7]. This will lead to better understanding of how the same application may be implemented differently depending on how the temporal and spatial requirements are selected.

A. Traditional Characterization of Transmission Line Faults

This section illustrates how the first step of the analysis—namely, determining that a fault has occurred, what is the fault type, and where the fault is located—depends on specific temporal and spatial considerations.

1) Fault Detection and Classification: Traditionally, fault detection and classification are performed using phasors of current and voltage taken from one side of a transmission line. Since a transmission line has three phases, to determine fault type one must sample the three voltages and currents simultaneously so that proper phase angles between different phases can be established, leading to a proper classification of the fault type. Typically the data taken at one location, closest to the fault, are sufficient to detect and classify a fault. One remaining issue needs to be resolved, however: how can we determine the location of the fault? Since it is impossible to predict where the fault will occur, it is impossible to predetermine which end of the transmission line is closest to the fault. Hence, it becomes important to establish a metric for spatial determination of a fault location with respect to the measurement point. This leads to the next essential portion of the fault analysis, namely, fault location.

2) Fault Location: Traditionally, fault location is determined based on impedance measurements, which in turn are based on phasor measurements of voltage and current, obtained from one end of a transmission line [8]. Since such measurements may not be accurate, voltage and current phasors from two ends of a transmission line are used to improve accuracy. The sampling of the respective signals using a different sampling rate at two ends of a line does not allow temporal alignment of samples, and hence phasors may be misaligned. In the spatial sense, the transmission line ends are the locations of the measurements and the rest of the system is irrelevant. This approach still creates two problems: 1) if measurements are not available at each line end (which may still be the case in many utilities) and if the adjacent lines are branching out, measurements may be inaccurate, and 2) if the measurements are available at both ends, the calculation of phasors may still create accuracy problems if the length of the recorded waveforms is insufficient due to a fast interruption of the breaker currents.

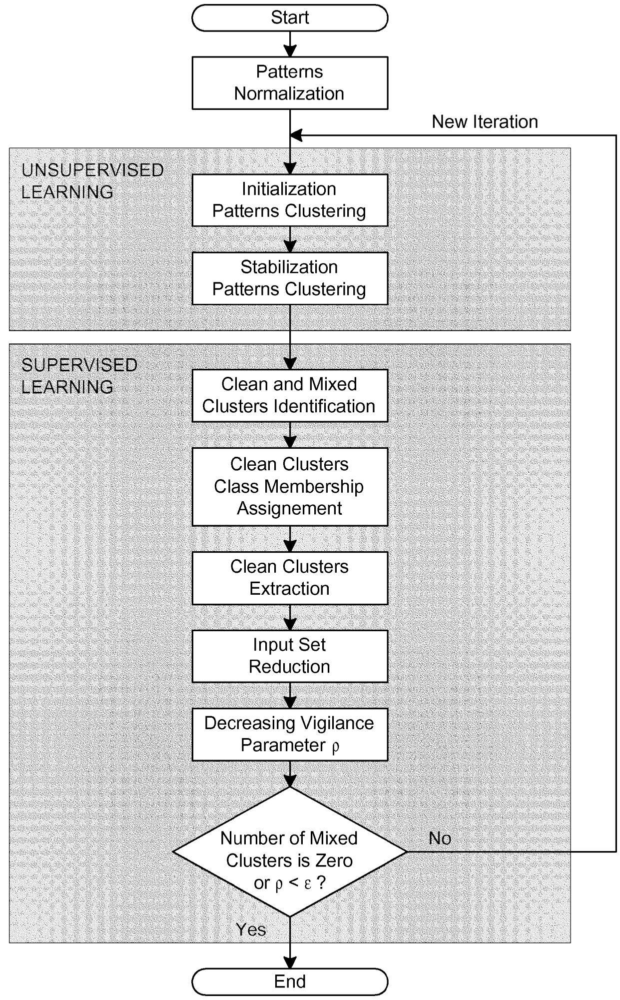


Fig. 1. Neural network training.

B. Novel Characterization of Transmission Line Faults

The novelty is related to a different view of the temporal and spatial consideration than what was discussed above.

1) Fault Detection and Classification Using Neural Nets: In the new approach, the temporal consideration is reduced to a selected length of the waveform used as a time vector of samples that may or may not coincide with a half or full cycle typically used for phasor reconstruction, which is the basis for traditional algorithms. The following discussion gives a brief summary of an application in which neural networks and fuzzy logic are deployed to achieve the fault detection, classification, and verification tasks [9].

The used neural network combines unsupervised and supervised learning techniques to give the best performance. Neural networks first use unsupervised learning with unlabeled data (the time vectors of samples) to form internal clusters. Labels are then assigned to the clusters during the supervised learning stage. The neural network training usually consists of a few hundreds of iterations with consecutively alternating unsupervised and supervised learning phases until prototypes of typical events (patterns) are established (Fig. 1).

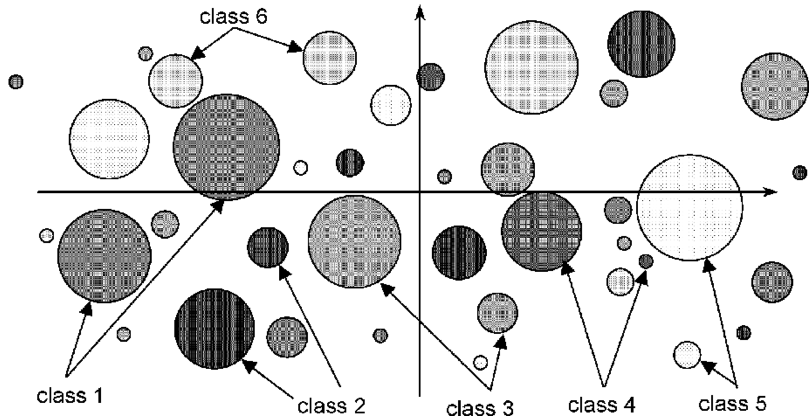


Fig. 2. The structure of clean clusters.

Using the cluster structure established during training performs classification of test patterns. The class labels of a selected number (usually a small odd number) of the nearest clusters are established. During classification, this classifier assigns to a test pattern the class label of the majority of class labels of the nearest prototypes in the neighborhood. Thus, output of this neural network is in a discrete form that reflects different types of faults common in protective relaying.

Input into the neural network is in the form of a moving data window (the time vector of samples) containing samples of phase current and voltage. Phase current and voltage measurements are filtered by an analog filter and sampled with desired sampling frequency. Each pattern is extracted from the samples in a desired time length of the moving data window, normalized and arranged together to form a common input vector with feature components. This is quite a different temporal situation than using phasors, since the selection of the time (length) of the waveform can be arbitrary.

One illustrative example of a reference set of clusters related to the fault analysis requirements is shown in Fig. 2 [9]. It relates to classification of the fault type and allocation of the fault location to the zone of relay protection. It is significantly simplified and given in two dimensions only.

In the training procedure used so far, incrementally established clusters tend to take positions in which they mutually overlap. Classifying test patterns located in overlapped regions, therefore, may be erroneous. Suitable training procedures that do not allow an overlap among the clusters should be applied, and classification results should be compared with an existing case when the clusters do overlap.

The test patterns might be heterogeneous and quite different from the training patterns, since there are many operating states and possible events in the power network. Test patterns are classified according to their similarity to prototypes adopted during training. Classification is performed by applying the K-nearest neighbor classifier (KNN) to the cluster structure established during neural network training procedure [10]. The main advantage of the KNN classifier is its computational simplicity, but its substantial disadvantage is that each of the neighboring clusters is considered equally important in detecting the class membership of the pattern being classified, regardless of their size and distances to that pattern.

To solve the problem the theory of fuzzy sets is introduced into the KNN technique to develop a fuzzy version of the classifier [9]. The first important extension of KNN is based on taking into account distances between patterns and selected number K of nearest clusters. The idea is that the closer neighbors should exert more influence on the class membership for the test pattern being labeled. The distance is generally selected to be a weighted Euclidean distance between a pattern and a prototype (cluster center). The fuzzy variable is introduced to determine how heavily the distance is weighted when calculating each neighbor's contribution to the class membership of a test pattern. The second extension is the introduction of a fuzzy membership value as a measure of a cluster belonging to its own class. The idea has been interpreted in an original way, considering a special cluster structure generated by the used neural network [9]. The measure of cluster membership in its own class is selected to be proportional to the cluster size. The outcome is that the larger clusters have more influence than the smaller ones.

Consequently, test patterns are classified based on the weighted distances to K nearest clusters, as well as on relative size and class labels of these clusters. Fuzzy K-nearest neighbor classifier calculates a vector of membership values of an input pattern to all classes present in K nearest prototypes. When membership values for all K neighbors have been calculated, pattern is classified as belonging to the class with the highest membership degree. Introduced fuzzyfication is a nonlinear interpolation technique used to help classify a test pattern dissimilar to all patterns presented during training process, and that pattern is classified based on the level of similarity to the neighboring training patterns.

This advanced approach also offers more realistic spatial consideration, since the training patterns can be selected at a variety of locations in the system, which makes the training amenable to particular configurations of the network and a variety of events and their locations. In the traditional solutions, the spatial determination of the event's location is driven by the impedance calculation, which may be an inaccurate measure of distance due to difficulty in defining apparent impedance in the case several lines are branching out.

2) Fault Location Using Synchronized Sampling: Traditional phasor-based fault location techniques are accurate if

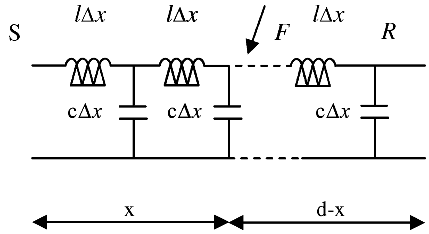


Fig. 3. Unfaulted long transmission line.

the phasors are correctly calculated. In some special cases, such as high-speed tripping or time-varying fault resistance, phasor-based techniques may experience large errors, while the synchronized sampling techniques are inherently transparent to such conditions. The proposed fault location method is based on a discrete form of Bergeron's traveling wave equations or lumped parameter line equations [11]. To derive these equations, we can consider an unfaulted long transmission line, shown in Fig. 3.

A transmission line longer than 150 mi can be represented as an L–C circuit, since the contribution of the resistance and conductance to the series impedance and shunt admittance can be neglected. The length of the line is d. The l and c are the series inductance and shunt capacitance per unit length. The voltage and current at the point F, at distance x from the sending end S, are given by

$$\begin{aligned} v_F(t) =& \frac{z}{2}\left[i_S(t-\tau_x) - i_S(t+\tau_x)\right] \\ &+ \frac{1}{2}\left[v_S(t-\tau_x) - v_S(t+\tau_x)\right] \qquad (1) \\ i_F(t) =& -\frac{1}{2}\left[i_S(t-\tau_x) + i_S(t+\tau_x)\right] \\ &- \frac{1}{2z}\left[v_S(t-\tau_x) - v_S(t+\tau_x)\right]. \qquad (2) \end{aligned}$$

These equations follow directly from Bergeron's traveling wave equations. Here, z is the characteristic impedance of the line and τ_x is the travel time from end S to point F

$$z = \sqrt{\frac{l}{c}}, \quad \tau_x = x\sqrt{lc}. \qquad (3)$$

The voltage and current also can be written in terms of the R end voltage and current by replacing the subscript S with R and changing the travel time τ_x to τ_{d-x}, which is the travel time from end R to F. If a fault occurs at F, then the voltage at point F due to the end S voltage and current will be the same as the voltage at F due to the end R voltages and currents. Thus the fault location equation becomes

$$\begin{aligned} &\frac{z}{2}\left[i_S(t-\tau_x) - i_S(t+\tau_x)\right. \\ &\quad \left. -i_R(t-\tau_{d-x}) + i_R(t+\tau_{d-x})\right] \\ &+ \frac{1}{2}\left[v_S(t-\tau_x) + v_S(t+\tau_x)\right. \\ &\quad \left. -v_R(t-\tau_{d-x}) - v_R(t+\tau_{d-x})\right] = 0. \qquad (4) \end{aligned}$$

The distance to the fault does not appear explicitly in the equation. When the equation is written in a discrete form, the travel times to point F from either end will no longer be exact. The right-hand side of (4) will have a finite nonzero value. Based on the sampling time step, the line now can be divided into a number of discrete points, and (4) can be used to compute the error voltage at each of those discrete points. The point that yields the minimum error value is the estimate of the fault.

This approach emphasizes the importance of both the time (synchronization to GPS) and spatial (data from all line ends) aspects of an accurate fault location calculation. The transient voltage and current are processed in real time. In this case, the temporal consideration requires synchronous sampling and the spatial consideration requires data synchronously sampled at both ends, which is different from the traditional use of phasors.

V. Hierarchical State and Topology Estimation

This section addresses two important issues, namely, the spatial and temporal considerations in the context of multiarea state estimation and topology tracking respectively.

A. *Multiarea State Estimation*

Ever increasing sizes of power system models used in power system applications present a significant computational challenge. One of the commonly used applications is the state estimation function. Several researchers studied the idea of decomposing the problem into several smaller ones with manageable sizes and solving it through a hierarchical method [12]–[15]. Their main focus, however, remained on the reduction in computing time and memory requirements. It was assumed that data and information exchange between decomposed parts of the overall system would somehow be possible, which implies nontraditional spatial consideration of data availability.

State estimators process measured values of power flows, power injection, and bus voltage magnitudes to determine the corresponding operating state of the system. Ideally, the estimated state should correspond to a single snapshot of the overall system. However, an inevitable time skew between scanned measurements will make this only an approximation. Advances in communication and processing capabilities in the substations will minimize this time skew, as well as any bias due to that in the estimated state. Furthermore, the overall system can be broken down into nonoverlapping areas, each of which has its own state estimator. The results can be centrally coordinated if individual area solutions as well as the measurements at the boundaries of these areas are available. This is a bilevel hierarchical scheme: individual areas represent the lower level and the coordinator represent the top level.

One such estimator is developed and tested using simulated measurements [16]. In this setup, each area will solve the following optimization problem:

$$\text{Minimize } J_i = r_i^T R_i^{-1} r_i \qquad (5)$$

$$\text{Subject to } z_i = h_i(x_i) + r_i \qquad (6)$$

where

- z_i — the vector of available measurements in area i, having m_i elements. They include not only all the internal measurements but also the injection and flow measurements incident at the boundary buses and the area tie-lines.
- r_i — the residual of measurement z_i.
- R_i — the measurement error covariance matrix for area i.
- $h_i(x_i)$ — the measurement function for area i measurements.

Once individual area solutions are available, these will be passed on to the central coordinator, which will solve the following problem:

$$\text{Minimize } J_S = r_S^T R_S^{-1} r_S = [z_S - h_S(x_S)]^T R_S^{-1} [z_S - h_S(x_S)] \quad (7)$$

$$\text{Subject to } z_S = h_S(x_S) + r_S \quad (8)$$

where

- $z_S = [z_u^T, z_{\mathrm{ps}}^T, \hat{x}^{b^T}, \hat{x}^{ext^T}]^T$ — all data and measurements available to the coordinator;
- z_u — measurement vector including the tie-line flows and injections incident at all boundary buses;
- z_{ps} — GPS synchronized phasor measurements vector;
- r_S — the residual vector of measurement z_S;
- $\hat{x}^b = [\hat{x}_1^{b^T}, \hat{x}_2^{b^T}, \cdots, \hat{x}_n^{b^T}]^T$ — boundary state variables estimated by individual areas and used as pseudomeasurements by the coordinator.

$\hat{x}^{ext} = [\hat{x}_1^{ext^T}, \hat{x}_2^{ext^T}, \cdots, \hat{x}_n^{ext^T}]^T$, similar to $\hat{x}^b$, except defined for the external buses of each area.

The measurement model will then be given as follows:

$$z_S = h_S(x_S) + e_S$$

where

- $x_S^T = [x^{b^T}, u^T]$ — coordination state vector with dimension n_S;
- e_S — measurement error with a zero mean normal distribution and covariance $R_S = E(e_S e_S^T)$;
- h_S — nonlinear measurement function.

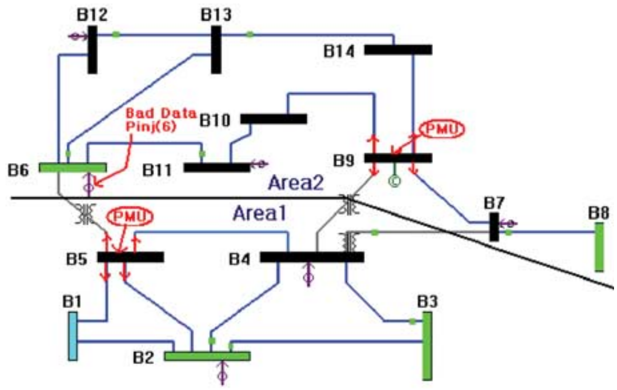


Fig. 4. Areas and measurements for the 14-bus system.

In the above formulation, each area is assumed to solve its own state estimation problem by ignoring some of the boundary measurements, whose processing is not possible without knowing the estimated states from the neighboring areas. As a result, it is possible for individual areas to fail to detect bad data appearing at the boundary buses. This data, however, will be detected and identified at the second level by the coordinator. Notable measurements used by the coordinator are the GPS synchronized phasors. These measurements allow spatial coordination without hampering individual estimation processes. Such measurements will facilitate detection and identification of bad boundary measurements, which would otherwise be missed due to low redundancy around these buses.

A tutorial example will be given below in order to illustrate the procedure outlined above. This also may serve as an illustration of how spatial data exchange and specially designed algorithms can facilitate large-scale solution to an EMS problem.

Fig. 4 shows the system diagram, area designations and available measurements for a 14-bus power system. State estimation results obtained using the integrated system are given in Table 1. The table also provides individual area estimates independently obtained by Area 1 and Area 2 estimators, as well as their coordinated solution which is obtained by the top level estimator. Note that the Area 2 reference bus is selected as bus 14. Since no bad data are present, top level and integrated system solutions match closely.

Next, boundary injection at bus 6 is modified to simulate a bad measurement. As evident from the sorted normalized residuals for Area 2 estimation given in Table 2, bad measurement cannot be detected by Area 2 estimator, since it is critical in the absence of external measurements from Area 1.

Bad data problems will be resolved by the top-level estimator, which will make use of all boundary measurements and PMU measurements as well as the estimated states by individual areas. Normalized residual tests at the top level identify the voltage magnitude and phase-angle estimates of bus 5 by the Area 2 estimator, as well as the bad real power injection at bus 6 as bad data. Upon their removal, an unbiased estimate is obtained for the overall system. The sorted normalized residuals and corresponding measurements are

Table 1
Results of Integrated, Individual Area, and Coordination Level Estimation for the 14-Bus System

Bus no.	Estimated State: V, θ					
	Integrated system		Individual area estimates		Top level SE	
1	1.059	0.00	1.0591	0.00	1.0591	0.00
2	1.044	-4.98	1.0441	-4.98	1.0441	-4.98
3	1.010	-12.73	1.0094	-12.72	1.0094	-12.72
4	1.018	-10.33	1.0178	-10.33	1.0178	-10.32
5	1.019	-8.79	1.0195	-8.79	1.0201	-8.76
6	1.069	-14.25	1.0692	1.99	1.0696	-14.22
7	1.061	-13.36	1.0594	2.69	1.0606	-13.36
8	1.088	-13.31	1.0862	2.73	1.0862	-13.32
9	1.055	-14.95	1.0538	1.10	1.0549	-14.95
10	1.050	-15.10	1.0489	0.95	1.0489	-15.10
11	1.057	-14.82	1.0564	1.33	1.0564	-14.72
12	1.054	-15.00	1.0545	1.17	1.0545	-14.88
13	1.048	-15.24	1.0491	0.93	1.0491	-15.12
14	1.034	-16.05	1.0331	0.00	1.0331	-16.05

Table 2
Results of Area 2 State Estimation When Pinj at Bus 6 Contains Gross Error

Measurement	Normalized Residual
Pflow 12-13	2.01
Pinj 12	1.87
Pflow 6-13	1.67
Decision:	No Bad data

Table 3
Bad Data Identification Cycles for the Top Level Estimator

Cycle	1st	2nd	3rd	4th
Meas. : R_n	$\hat{\theta}_5^2$: 129.1	Pinj 6: 118.2	$\hat{V}_5^2$: 15.1	$\hat{\theta}_{14}^2$: 2.6
	Pinj 6: 114.03	$\hat{\theta}_6^1$: 96.5	$\hat{\theta}_6^1$: 1.4	$\hat{\theta}_1^1$: 2.6
	$\hat{\theta}_6^1$: 90.4	Qinj 6: 37.5	$\hat{\theta}_1^1$: 1.3	Pinj 7: 2.0
Eliminated meas.	$\hat{\theta}_5^2$	Pinj 6	$\hat{V}_5^2$	No More Bad data

shown for all the bad data identification cycles in Table 3. Superscripts indicate the area that performs the estimation, and quantities with hats are estimated variables used as pseudomeasurements in the top-level estimation, and R_n indicates the normalized residual value.

B. *Topology Tracking*

The system model used by the state estimator is composed of buses and branches. This model is built by the topology processor, which uses the information available on the status of breakers and switches at the substations to convert the detailed breaker-level model into the much simpler bus-branch model of the system. This conversion may not produce the true system model when the information on the breaker status is wrong. Such topology errors are difficult to detect and correct using the current state estimators.

Several recent investigations provide possible solutions to topology error detection and identification problems [17]–[20]. These approaches are based on state estimation, which is reformulated to incorporate substation models whenever topology errors are suspected. However, the problem can be handled with less effort if it is solved at the substation before it is allowed to corrupt the central state estimator database. Implementing a topology tracking system at the substation and giving the central state estimator access to tracking system outputs when necessary can accomplish this. The proposed tracking system will collect data and measurements within the substation and perform various consistency checks based on circuit laws and connectivity logic. The result will be an error-free substation topology that is dynamically updated at each measurement scan based on the real-time measurements. A prototype of such a tracking system is described in [21]. The substation topology tracking system can be coupled with a two-stage state estimator as shown in [22].

Regardless of the solution method used, the conventional state estimation formulation is based on the bus/branch model obtained from the topology processor. The circuit breakers will not appear in the model. Estimation of the power flows through circuit breakers is first suggested for data validation at the substation by Irving and Sterling [23]. This requires the detailed topology of the substation, including the circuit breakers, to appear in the system model. Circuit breakers are modeled as zero impedance branches and their flows are treated as additional state variables [18]. Here, a weighted least absolute value estimator (WLAV) will be reformulated so that the substation models can be incorporated. The choice of WLAV method facilitates automatic rejection of bad measurements and incorrect constraints for breakers of unknown status. If a substation is to be modeled in detail, representing the individual circuit breakers and their configuration, then the linearized measurement equations will take the following form:

$$\Delta z = H \cdot \Delta x + M \cdot f + e \qquad (9)$$

where H is the measurement Jacobian, Δx is the state update, and e is the measurement error.

$[M]$ is a $(m \times l)$ measurement to circuit breaker incidence matrix defined as follows:

if the measurement i is an injection:

$$M_{ij} = \begin{cases} 1 & \text{if; the injection is at the to-end of the breaker j} \\ -1 & \text{if; the injection is at the from-end of the breaker j} \\ 0 & \text{otherwise;} \end{cases}$$

if the measurement i is a line flow:

$$M_{ij} = \begin{cases} -1 & \text{if; the flow is at the to-end of the breaker j} \\ 1 & \text{if; the flow is at the from-end of the breaker j} \\ 0 & \text{otherwise;} \end{cases}$$

where l is the number of the circuit breakers and f is a $(l \times 1)$ vector of power flows through the circuit breakers.

A new vector is defined to designate the state vector augmented by the circuit breaker power flows

$$\Delta y = [\Delta x^T f^T]^T. \tag{10}$$

WLAV state estimation problem can now be stated as the following LP problem:

$$\min J(x) = \sum_{i=1}^{m} \omega_i(u_i + v_i) \tag{11}$$

$$\text{subject to } \Delta z^k = H(x^k) \cdot \Delta y^k + u - v. \tag{12}$$

Additional constraints are appended to the LP problem in the form of zero voltages across closed circuit breakers. Since the status of the breakers is not known *a priori*, such constraints are made soft by introducing a pair of slack variables so that the constraints will be disregarded if the breakers are actually open. For a circuit breaker between buses j and k, the following equation will be appended to (12):

$$x_j - x_k + u_{m+1} - v_{m+1} = 0 \tag{13}$$

where $u_{m+1} v_{m+1}$ are the nonnegative slack variables for the newly added pseudomeasurement.

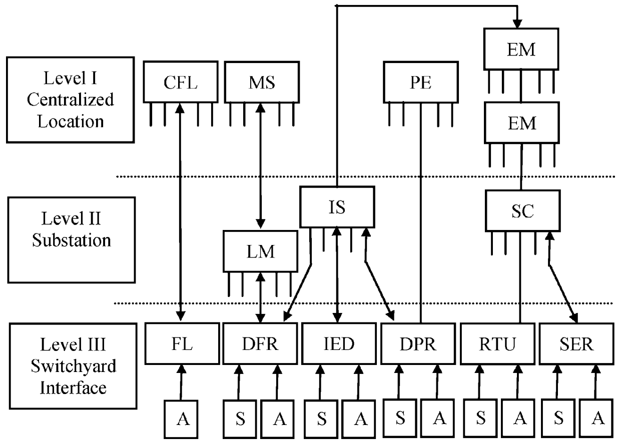
Depending on the column rank of the matrix $[H|M]$, some or all entries in f will be observable. The WLAV estimator can identify the unobservable states during the initial phase of the solution. Including all substations in full detail is not efficient due to the exponential increase in the size of matrix M. Instead, the state estimator can ignore topology errors unless the residual analysis yields a warning and identifies one or more suspect substations of having such errors. Subsequently, these suspect substations will be polled and their most recent topology tracking outputs will be telemetered to the state estimator. A second stage estimation will then be executed as shown in (11) and (12), this time using the detailed models of the polled substations and true topology of the system will be uncovered based on the redundant set of measurements.

VI. Future Integrated Solutions

This section discusses how the temporal and spatial considerations may be merged in an integrated monitoring, control, and protection solution. The first section summarizes how the solutions are implemented today and the second discusses future solutions.

A. Existing Data Acquisition and Information Extraction Paradigm

A typical equipment infrastructure outline for the legacy solutions is given in Fig. 5. It is organized to follow the individual temporal and spatial needs of traditional functions, which are defined by the traditional power system operating states.



A- Analogue inputs S- Status (contact) inputs
SC-substation computer MS-master station
CFL-centralized fault location EM-energy management
PE-protection engineer IS- integrated systems

Fig. 5. Legacy infrastructure for monitoring and control.

Table 4
Typical Examples of the Temporal Consideration

Time Scale	Application	IED Used	Synchronization
Microseconds	Phasor Measurements	PMUs	GPS receiver
Milliseconds	Relaying	DPRs	Local clock
100s of ms	Fault location	FLs	Two line ends
Seconds	State Estimation	RTUs	Local/Central time stamp

Substation recording IEDs can be quite versatile [7]. The choice may depend on many factors, including the history of the substation construction and upgrades, utility operating practices, the strategic importance of the substation, etc. The different recording IED types typically used in modern substations are:

- digital protective relays (DPRs);
- digital fault recorders (DFRs);
- sequence-of-event recorders (SERs);
- RTUs of a SCADA system;
- IEDs used for variety of monitoring and control applications;
- fault locators (FLs) developed for stand-alone high-accuracy fault locating.

Observing the legacy infrastructure and the properties of the data acquisition parts of each of the infrastructures, one can quickly conclude that SCADA gives a broad spatial but limited temporal view of the system dynamics, while the DPRs and DFRs give much better time resolution of the signal and status changes. Their view, however, is local. The temporal issue is best illustrated with different sampling requirements indicated in Table 4. The limited SCADA view may be enhanced with a view available from other IED infrastructures, but this is not done today.

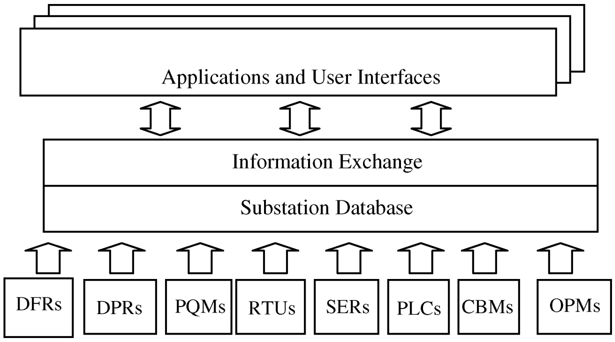


Fig. 6. New infrastructure for monitoring and control.

B. *Future Data Integration and Information Exchange Paradigm*

The new data integration and information exchange approach proposed in the mentioned EPRI study [6] is based on a different infrastructure shown in Fig. 6. In this infrastructure, all the IEDs are synchronized to the GPS clock, which in turn can be tied to the absolute time.

In the temporal sense, all the data collected by IEDs would be recorded using the same clock for S/H circuits on each of the measurement channels. This will make all the different time considerations mentioned at the beginning of the paper far less complex and yet significantly more effective for the following reasons.

- All the events would be correlated to the same time reference and could be interpreted using an absolute time.
- All the waveforms would be sampled synchronously, allowing more accurate applications based on synchronized samples and synchronized phasors.
- All the waveform representation can be in the time domain, allowing an accurate assessment of the disturbances, while the phasors could be extracted if needed.
- All the commands could be synchronized using the GPS clock.

In the spatial sense, any space allocation mentioned at the beginning of the paper could be accommodated for: power apparatus coverage, decision making, data processing and information extraction, and command execution. As an example, the option for implementing monitoring, control, and protection application could be expanded to the following major cases:

- decentralized (localized) data processing and decision making;
- centralized (EMS) data processing and decision making;
- distributed (integrated substation system) data processing and decision making with a coordination capability.

VII. Conclusion

The following major conclusions may be drawn.

- Performance of current monitoring functions is constrained by the time and space allocation to specific IEDs in customized separated infrastructures.
- Performance of future monitoring functions can benefit from transparent allocation of time and space across many IEDs in a common infrastructure.
- Data integration and information exchange are needed for the time and space transparency.

Acknowledgment

The authors would like to thank D. J. Sobajic, formerly with the Electric Power Research Institute (EPRI), and A. Edris, currently with EPRI, for their comments during an EPRI study related to many of the issues reported in this paper.

References

[1] M. Kezunović, "Fundamentals of power system protection," in *The Electrical Engineering Handbook,* W.-K. Chen, Ed.. New York: Academic, 2005, ch. 9, pp. 787–804.

[2] T. E. Dy Liacco, "The adaptive reliability control system," *IEEE Trans. Power App. Syst.,* vol. PAS-86, no. 5, pp. 517–531, May 1967.

[3] ——, "Control of power systems via the multi-level concept," Ph.D Dissertation, Case Western Reserve University, Cleveland, OH, 1968.

[4] J. Zaborszky, A. K. Subramanian, T. J. Tarn, and K. M. Lu, "A new state space for emergency control in the interconnected power system," *IEEE Trans. Autom. Control,* vol. AC-22, no. 4, pp. 505–517, Aug. 1977.

[5] L. H. Fink and K. Carlsen, "Operating under stress and strain," *IEEE Spectr.,* vol. 15, no. 3, pp. 48–53, Mar. 1978.

[6] M. Kezunovic and A. Abur, "Data integration and information exchange: Impact on future substation and EMS applications and related implementation requirements," EPRI Final Rep. 1 010 898, May 2004.

[7] M. Kezunović, C. C. Liu, J. McDonald, and L. E. Smith, "Automated fault analysis (IEEE tutorial)," IEEE Power Engineering Soc., 2000.

[8] M. Kezunović and B. Peruničić, "Fault location," in *Wiley Encyclopedia of Electrical and Electronics Terminology*. New York: Wiley, 1999, vol. 7, pp. 276–285.

[9] S. Vasilić and M. Kezunović, "Fuzzy ART neural network algorithm for classifying the power system faults," *IEEE Trans. Power Del.,* vol. 20, no. 2, pp. 1306–1314, Apr. 2005.

[10] J. Keller, M. R. Gary, and J. A. Givens, "A fuzzy K-nearest neighbor algorithm," *IEEE Trans. Syst., Man, Cybern.,* vol. SMC-15, no. 4, pp. 580–585, July/Aug. 1985.

[11] A. Gopalakrishnan, M. Kezunović, S. M. McKenna, and D. M. Hamai, "Fault location using distributed parameter transmission line model," *IEEE Trans. Power Del.,* vol. 15, no. 4, pp. 1169–1174, Oct. 2000.

[12] K. A. Clements, O. J. Denison, and R. J. Ringlee, "A multi-area approach to state estimation in power system networks," presented at the IEEE PES Summer Meeting, San Francisco, CA, 1973, Paper C72 465-3.

[13] H. Kobayshi, S. Narita, and M. S. A. A. Hamman, "Model coordination method applied to power system control and estimation problems," in *Proc. IFAC/IFIP 4th Int. Conf. Digital Computer Application to Process Control,* 1974, pp. 114–128.

[14] T. Van Cutsem, J. L. Horward, and M. Ribbens-Pavella, "A two-level static state estimator for electric power systems," *IEEE Trans. Power App. Syst.,* vol. PAS-100, no. 8, pp. 3722–3732, Aug. 1981.

[15] Y. Wallach and E. Handschin, "An efficient parallel processing method for power system state estimation," *IEEE Trans. Power App. Syst.,* vol. PAS-100, no. 11, pp. 4402–4406, Nov. 1981.

[16] L. Zhao and A. Abur, "A two-level state estimator for multi-ISO operation," presented at the Proc. North American Power Symp., Rolla, MO, 2003.

[17] A. Abur, H. Kim, and M. K. Celik, "Identifying the unknown circuit breaker statuses in power networks," *IEEE Trans. Power Syst.,* vol. 10, no. 4, pp. 2029–2037, Nov. 1995.

[18] A. Monticelli, "The impact of modeling short circuit branches in state estimation," *IEEE Trans. Power Syst.,* vol. 8, no. 1, pp. 364–370, Feb. 1993.

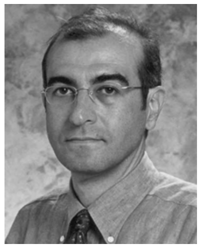
[19] K. A. Clements and A. S. Costa, "Topology error identification using normalized lagrange multipliers," *IEEE Trans. Power Syst.,* vol. 13, no. 2, pp. 347–353, May 1998.

[20] A. G. Exposito and A. de la Villa Jaen, “Reduced substation models for generalized state estimation,” *IEEE Trans. Power Syst.,* vol. 16, no. 4, pp. 839–848, Nov. 2001.
[21] S. Jakovljević and M. Kezunović, “Software for enhanced monitoring in integrated substations,” presented at the Power Tech 2003 Conf., Bologna, Italy.
[22] S. Zhong and A. Abur, “Implementation of a modified state estimator for topology error identification,” presented at the IEEE Power Engineering Society Annu. Meeting, Toronto, ON, Canada, 2003.
[23] M. R. Irving and M. J. Sterling, “Substation data validation,” *IEE Proc.,* vol. 129, no. 3, pp. 119–122, May 1982.

Mladen Kezunovic received the Dipl. Ing. degree in electrical engineering from the University of Sarajevo, former Yugoslavia, in 1974 and the M.S. and Ph.D. degrees in electrical engineering from the University of Kansas, Lawrence, in 1977 and 1980, respectively.

He is the Eugene E. Webb Professor and Director of Electric Power and Power Electronics Institute at Texas A&M University, College Station, where he has been since 1987. His main research interests are digital simulators and simulation methods for relay testing as well as application of intelligent methods to power system monitoring, control, and protection.

Dr. Kezunovic is a Member of CIGRE—Paris.



Ali Abur (Fellow, IEEE) received the B.Sc. degree from Orta Dogu Teknik Universitesi, Turkey, in 1979 and the M.Sc. and Ph.D. degrees from Ohio State University, Columbus, in 1981 and 1985, respectively.

He is currently a Professor in the Department of Electrical Engineering at Texas A&M University, College Station.

He is an editor for the IEEE TRANSACTIONS ON POWER SYSTEMS and *IEEE Power Engineering Society Letters*.