INVITED PAPER

Translational Knowledge: From Collecting Data to Making Decisions in a Smart Grid

Converting data to knowledge is the main concern of this paper, which provides a new data processing solution and illustrates it through several applications.

By Mladen Kezunovic, *Fellow IEEE*

ABSTRACT | This paper focuses on the most critical task in monitoring and control of power systems: converting data to knowledge that will facilitate control action. With expansion of smart grids, which assume wide dissemination of intelligent electronic devices (IEDs) in substations supported by extensive communications, data associated with power system events becomes abundant. The fact that new data of better quality than before exists in smart grids does not assure that better decision making will be possible. In order to convert data to actionable knowledge, certain processing needs to be performed. This processing is referred to as Translational Knowledge, i.e., the knowledge that allows transition from data collection to action. This concept is illustrated through several applications such as fault locating, alarm processing, and protective relaying.

KEYWORDS | Alarm processing; fault location; power system control; power system monitoring; protective relaying

I. INTRODUCTION

The enhancement of the monitoring, control, and protection of power systems through smart grid solutions primarily means availability of more data of better quality than before and availability of new applications that will utilize the data to produce better decision making. To allow such benefits, the process of data collection, integration and usage needs to be improved. A unique approach to providing more reliable and more comprehensive data not available before was introduced a few years ago [1]. It resulted in development of a number of applications that took advantage of new information to gain better knowledge about power system events, particularly faults. At that time the smart grid terminology was not extensively used so the new concept was termed Data Integration and Information Exchange [2]. While the concept was the topic of research initially, outcomes are reduced to practice and demonstrated at a number of utility sites now [3].

What made the new solutions possible was the concept of Translational Knowledge, a set of procedures that allowed field data to be merged with models to produce better decision making. While there is extensive literature about Translational Science [4] going into details of assumptions and meanings of the science contribution to knowledge formulation, and its contribution to practical engineering decisions, this paper stays at the level of illustrating such concepts using several new solutions made possible through smart grid developments. Of particular attention in this paper are fault conditions in power systems where a number of decisions are made either automatically or by operators to detect faults, isolate faulted components, repair faulted elements and restore power system operation.

The first application discussed, which utilizes new data to create improved decision making after occurrence of a fault, is an optimal fault location scheme. While many solutions for fault location under particular system conditions have been developed so far [5], [6], there is still no approach that will select the best result for any network configuration and operating condition. An improved approach to locating faults with an optimal outcome was developed utilizing additional data about network conditions, as well as additional measurements [7]. Further details how the knowledge about algorithm constraints and performance leads to better decision

Manuscript received October 5, 2010; revised December 29, 2010, and January 19, 2011; accepted January 31, 2011. Date of current version May 17, 2011.
The author is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843-3128 USA (e-mail: kezunov@ece.tamu.edu).

Digital Object Identifier: 10.1109/JPROC.2011.2112631

making is illustrated through the use of network models and fault data matching capabilities [8], [9]. The next application introduces a new alarm processing concept where cause/effect reasoning is improved through knowledge about device operation embedded in a control model of protective relay operations. While intelligent alarm processing has been a subject of investigation for some time [10]–[12], a breakthrough in the ability to make more informed decisions came as a result of merging data from substation intelligent electronic devices (IEDs) with supervisory control and data acquisition (SCADA) data [13]. The last application that allowed a new approach based on availability of improved power system models is a neural network (NN)-based protective relaying scheme, which has many variations [14]–[18]. As a result, an improved approach to protective relaying has been developed with a distinct feature that creates benefits over existing solutions: no settings are used, hence no need for setting calculation and coordination [19]. This makes the scheme inherently adaptive, which is an ultimate need in smart grids where penetration of renewable generation and overloading of existing systems creates unique fault and operating conditions making setting coordination a challenging task. To create improved decision making, the knowledge about power system conditions has been utilized to train the NN solution and allowed transition of such knowledge into the decision making of a new relay design based on new input data choice.

The paper first addresses the spatial and temporal aspects of data collection. Each of the new applications is discussed in subsequent sections and conclusions; acknowledgements and references are given at the end.

II. BACKGROUND

This section gives a brief review of spatial and temporal considerations of data collection and discusses the impact on data integration and information extraction [20].

A. Temporal Considerations

1) Relative and Absolute Time as a Reference for Correlating Power System Events: Monitoring, control, and protection applications require knowledge of the instance of time when a given event has occurred. The *relative time* may be used to understand the time sequence between the various control actions. As an example, knowing the inception time of a fault, the time that takes relays and breakers to operate can be calculated relative to the incident time. Besides, an *absolute* time plays a role when various data related to a given disturbance is collected at multiple locations, and such data is used to improve knowledge about the event. As an example, operation of multiple relays and tripping of multiple breakers may be sensed by IEDs located in multiple substations, so absolute time needs to be known to be able to differentiate actions corresponding to the same event from actions caused by other but time adjacent events.

2) Sampling Clock Time as a Reference for Synchronous Signal Sampling Versus Scanning: Various measurements in the power system are performed by IEDs which 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 converter converts samples into a computer word, known as data. The clock signal used for initiating the S/H circuit operation 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 signals in the three phases were scanned. The mode of sampling clock control that results in synchronous sampling versus scanning is widely different in modern IEDs versus legacy remote terminal units of SCADA system.

3) Time as a Reference for Waveform Representation in Time and Frequency Domain: Many of the applications 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 (during faults) while the phasor representation is sufficient for steady state conditions (during normal operating 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 at a given time instance of interest.

A typical example of time synchronization between waveform samples is two-ended protection or fault location on a transmission line which is implemented using time-domain solutions. A typical example of synchronization of phasor samples is in the two-ended fault location where the measurements from two ends are phasors but may be used either as synchronized or unsynchronized. While the phasors extracted from sample may be used for many applications, taking phasor samples synchronously across all the IEDs is more involved since it requires understanding of how the sample calculation is performed.

4) Implementation of the Time Reference: To illustrate how various options of time synchronization mentioned above may be implemented, Fig. 1 shows various designs of the sample and hold and A/D conversion circuits used in legacy and new IED solutions. For ease of implementation of the smart grid solutions in general and in particular the

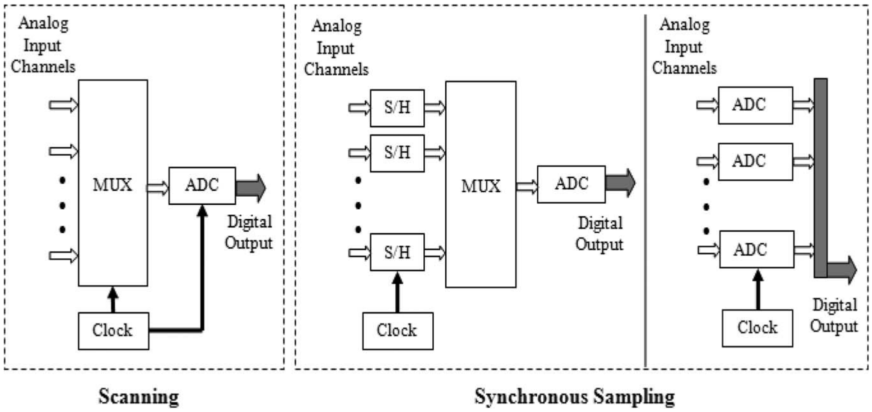


Fig. 1. *Synchronous sampling versus scanning.*

ones discussed in this paper, the sampling synchronization should be controlled by a common reference such as a GPS time synchronization signal [21]. GPS receivers typically provide both the sampling clock signal and absolute time reference as shown in Fig. 2, which may be then combined with the design shown in Fig. 1 to provide very precise control of all the temporal issues mentioned in this section.

B. Spatial Considerations

1) Location as a Reference for Data Processing and Information Extraction: For various events in the power system such as faults, only specific power system components are involved in the event and only local data from IEDs are used in the process of information extraction. For example, a transmission line protection relay uses only local data to analyze faults on the line. On the other hand, system integrity protection scheme (SIPS) monitors large area and requires wide-area data. Due to the huge diversity in IEDs, their technologies, and communication infrastructure, sometimes it is challenging to achieve good spatial considerations. There are some non-traditional ways that could be used to improve spatial considerations, like precise satellite images similar to Google Earth [22], National Lightning Detection Network (NLDN) [23], and geographic information system (GIS) technology [24] that can be used for enhanced presentation of spatial data obtained by substation IEDs.

2) Location as a Reference for Model Representation: In the Translational Knowledge approach, to create new knowledge, the extracted information needs to be

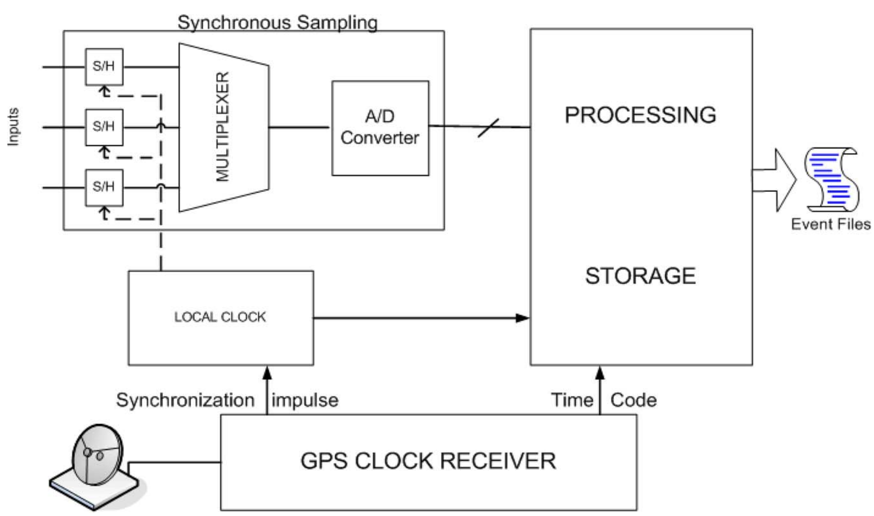


Fig. 2. *GPS synchronization.*

Table 1 Temporal, Spatial, and Model Requirements

Application	Temporal	Spatial	Model
Optimal Fault Location	Synchronized or unsynchronized Phasor or Sample Vector	Local and system-wide	Power System Network
Intelligent Alarm Processing	Synchronized or unsychronized Phasors	Substation and system wide	Petri-Net Logic
NN Based Fault Detection and Classification	Synchronized Sample Vector	Local	Result Pattern Space

supplemented by a predefined model. Choices of models vary depending on which type of information is utilized to create them. As an example, the model can represent a power system network, like in the case with the optimal fault location applications, or it can represent cause/effect relationships in control equipment operations like in the case of Petri-Net logic used in Intelligent Alarm Processors, or it can represent pattern space for resulting vectors in the neural network-based Fault Detection. Such cases are discussed later in the paper.

To illustrate how the different background considerations relate to the three applications discussed further in this paper, Table 1 shows temporal, spatial and model requirements for those applications. More details can be found in the following sections.

III. OPTIMAL FAULT LOCATION

This section gives a brief overview of the typical fault location approaches and implementation requirements and then focuses on a solution that illustrates additional benefits from the Translational knowledge approach.

A. Overview of Fault Location Algorithms

Transmission lines occasionally suffer from faults which are generally caused by random and unpredictable events. Transmission lines are normally equipped with protective devices (relays) at both/all ends. Relays sense the signal waveforms caused by such faults and isolate the faulted line as soon as possible by operating corresponding circuit breakers. Distance relays typically give some idea about the fault location, but the results may over-reach or under-reach due to several unknown parameters, such as prefault loading, fault resistance, remote in feed etc. In order to restore service quickly, an accurate location of the fault is highly desirable in all circumstances to help the maintenance crew find and repair the faulted line as soon as possible.

Transmission line fault location approaches can be broadly classified into three categories: [5], [6]:

- phasor based;
- time-domain based;
- traveling wave based.

Phasor based methods use a fundamental frequency component of the signal and lumped parameter model of the line while time-domain-based methods use transient components of the signal and a lumped or distributed parameter model of the line. Both of these methods can be subdivided into another two broad classes depending upon the availability of recorded data: single-end methods where data from only one terminal of the transmission line is available and double-end methods where data from both (or multiple) ends of the transmission line can be used [25]–[27]. Double ended methods [28]–[34] can use synchronized or unsynchronized phasor measurements, as well as synchronized or unsynchronized samples. Traveling-wave-based fault locations [35]–[37] use the reflection waves generated by the fault. They are based on the correlation between the forward and backward traveling waves along a line or direct detection of the arrival time of waves at line terminals. Each of the techniques requires very specific measurements from one or both ends of the line to produce results with desired accuracy. In some applications, the measurements from both ends of the transmission line are not available and single-ended algorithms cannot perform well due to the unique configuration of the transmission lines (multiterminal lines, taps, etc.). In such cases, some unconventional techniques based on wide area measurements may have to be used [8], [9].

B. An Optimal Solution That Matches Data With Models

To illustrate how Translational Knowledge is generated based on the field sampled data and power system model, a fault location scheme that matches the field data with a power system model to produce an optimal solution is discussed in detail next.

A typical smart grid scenario that will illustrate this unique approach is the situation when tapped lines are used to supply large industrial customers. In this case, measurements at all ends of the tapped line may not be available due to the lack of measurement sensors at the taps. Since the smart grid implementation assumes that measurements at other locations in the vicinity are available through communications, a scheme that takes advantage of the sparsely measured data is developed.

C. Data Processing

Traditionally, in a substation, remote terminal units (RTUs) acquire analog and digital measurements (bus voltages, branch flows, frequency, breaker status, transformer tap position, etc.), collectively called SCADA measurements, which are sent to the energy management

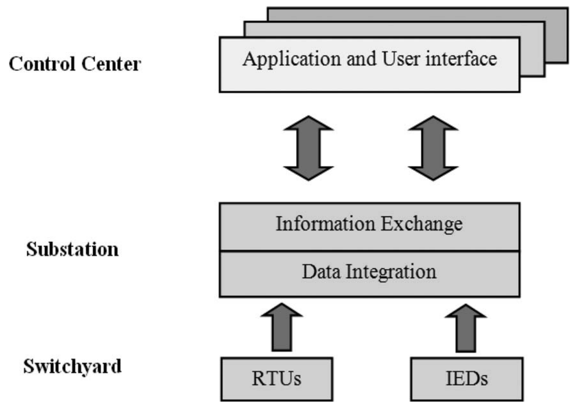


Fig. 3. ***Functional diagram for substation data flow.***

systems (EMS) every two to ten seconds. With the rapid advancement of technology, IEDs come into the picture. These computer-based devices can record and store a huge amount of data (both operational and non-operational) with a periodicity depending upon the intended purpose of the device. Their sampling rates are much higher than what is used in RTUs, and the data is taken synchronously versus being scanned in RTUs.

In a modern integrated substation automation solution, various types of IEDs are employed for monitoring, control and protection purposes. All of the newer models have data recording and exporting capabilities today. The IEDs are triggered by various events such as faults, and at that time they record data that is typically not captured by SCADA RTUs such as auto reclosing sequences, transients, and current dc offset. This additional data may be used to supplement SCADA data to create better results in fault location. The basic idea of integration of data is to collect all the IED data in a substation database and use it for extracting information automatically and then utilizing the extracted information for several power system applications. The functional diagram for proposed substation data flow is shown in Fig. 3.

To import the IED data into the central repository requires means of data format conversion and communication among different IEDs, all features that are widely required in the smart grid solutions. In addition to the automatically retrieved IED and RTU data, the database should contain several other data, such as the following:

- static system data containing description of the system components and their connections (i.e. topology and model parameters);
- SCADA EMS PI Historian data, which may be used to tune the static system model with real-time data;
- substation interpretation data that allows one to correlate the naming convention of recording devices and of the static system model with PI Historian data.

Although, integrating a huge amount of data provides improved information by exploiting redundancy, the quality of data is also a major concern. Distortion in magnitude and phase angle of current and voltage signal is introduced in each stage of the measurement chain. Ideally, the output waveform should be an exact replica of the input signal, but the error introduced in several data processing stages makes the output distorted. Thus the quality of data depends largely on the performance of devices used in the measurement chain. The performance of these devices and the information extraction schemes are discussed elaborately in [38]. Several commercial software packages are used to implement the fault location scheme. The static power system model is implemented using PSS/ETM 30 [39]. To tune the power grid with pre-fault data, SCADA PI-Historian data is used [40]. The required data for this approach can be broadly classified into:

1) System Level Data: This includes power system model data (in saved case format *.sav) and data reflecting real time changes in power system (PI Historian data). The power flow input data (in *.raw format) contains power flow system specification data for the establishment of a static system model used by PSS/E to run the power flow analysis. Sequence data (*.seq) contains the negative and zero sequence impedance data needed for short circuit study. SCADA PI Historian data contains the latest load, branch and generator data to tune the static system data with the actual pre and post fault conditions.

2) Field Data: These include event data recorded by different IEDs after occurrence of any abnormality. The field recorded data (digital protective relay or digital fault recorder data) should follow the COMTRADE format. Using this format, the measured data (*.dat), and configuration (*.cfg) are described. The digital protective relay (DPR) or digital fault recorder (DFR) data contains analog and digital sample values from signal input channels for a specific substation. The configuration data contains information for interpreting the allocation of measured data to the equipment in a substation. The DFR recorded data supplied in native DFR format, used in our example, are converted to COMTRADE file using DFR Assistant software [41] which can generate an analysis report (containing the type of fault and a possible faulted line) in addition to generating the COMTRADE files.

In addition, we need substation interpretation data for each substation correlating the nomenclature used in DFR files and those used in PSS/E files. The interpretation files should be modified as frequently as needed to reflect the DFR configuration or system model changes.

D. Information Extraction

The event data obtained from DFRs should be preprocessed before use in the applications. This section illustrates how the data captured by DFRs are converted into information and integrated with system level data to be used in the fault location application.

1) Extraction of Phasors: Once the disturbance events are obtained from the IEDs, two processing steps are taken to obtain phasors from the samples of recorded analog signals as follows:

- removal of high-frequency noise by low-pass filtering;
- use of an improved Fourier algorithm to effectively remove decaying dc-offset component and obtain the prefault and during-fault phasors of voltages and currents [42].

The prefault phasor can be calculated using the first cycle of the recorded waveform. The during-fault phasor can be calculated using any fault cycle following the fault inception and prior to fault clearance. The fault inception moment is determined from waveforms recorded by DFR. It is possible to select different fault cycles to calculate the during-fault phasors from the different DFR recordings. This may introduce fault location error, especially for the arcing faults during which the fault resistance is changing. Under this situation, selecting different fault cycles means experiencing different impacts of fault resistance. Another option is to use time synchronized phasors which are calculated based on the same data window determined through one for the data synchronization methods mentioned in the background section. An alternative is that the user checks the waveforms manually and specifies the matching fault cycles across all recordings.

2) Synchronization of Phasors: The PSS/E load flow study based on the modified system model (the real-time power system model obtained from SCADA PI Historian) is carried out to obtain the prefault phasors and during fault phasors. For a typical fault case, several DFRs may be triggered and the phasors calculated from the recorded waveforms may lack synchronism which will introduce phase angle difference among phasors. Thus time synchronization of the phasors obtained from different DFRs is necessary. The phasors calculated from each DFR recording may be synchronized by rotating them in reference to the phasors obtained by the load flow study, assuming the angle difference between the pre- and during-fault phasor for the corresponding recorded current or voltage is fixed. This way, all recorded pre- and postfault phasors are synchronized using the same reference. Another option is to use GPS-synchronized phasors.

3) Tuning the Power System Model With Real-Time Power System Conditions: The saved case model from PSS/E may not reflect prevailing operating conditions of the system when a fault occurs. To obtain simulated phasors corresponding to the time when a fault occurred, the static system model should be tuned with real-time power system conditions. This tuning procedure may consist of updating power grid topology (switching status) and then updating generation and load data near the substations of interest. This may be achieved by utilizing information from both DFR recordings and SCADA PI Historian data. The updated model is saved in a new saved case data (*.sav) which is used for further simulation.

Updating power grid topology: Topology describes connectivity of various components in a power system. In our case, the topology (i.e., connectivity between different buses), line status (whether the line is in service or not), line impedances and susceptances are extracted from the static system model provided in the simulation tool. The topology update is performed using information of the pre-fault breaker status and the pre-fault current magnitudes of the monitored branches derived from the DFR data. It is assumed that a zero magnitude (or smaller than 0.01 p.u.) of the current through a monitored branch indicates an out-of service status of the branch. If both the current and the breaker status of a branch are available, the current measurement will be used instead of the breaker status for topology update. This is based on the observation that the monitoring of currents is usually more reliable than the monitoring of the breaker status because the measurement contacts of the breaker may be unreliable or may not be monitored. In this way the service status (i.e., in or out of service status in the static model saved in the PSS/E file) of the branches will be updated.

Updating generation and load data: In a typical power system, the operator is able to track changes in real time using the SCADA system. Through the SCADA database, a low sampling rate recording typically used to capture short term and long term disturbances is available. Captured data is typically scanned every few seconds and it is usually phasor or RMS data, not sampled data. The PI Historian data provided by the utility is load, branch and generator data scan (typically at 10 s intervals) in a period before and after fault for each substation where DFRs are triggered. These data are used to update the system.

E. Translational Knowledge Through Matching System Wide Sparse Measurement and Power System Model

The system wide sparse measurement based fault location method [8], [9] uses phasor measurements from different substations located in the region where the fault has occurred. The measurements may be sparse, i.e., they may come from only some of so many transmission line ends (substations) in the region. The technique compares measured data with data generated by the short circuit simulation of possible fault locations. The measured and simulated data from the locations where measurements are taken is compared while the location of the fault is changed in the short circuit program. This process is repeated automatically until the measured and simulated values have minimal difference, which indicates that the fault location used in the short circuit program is the actual

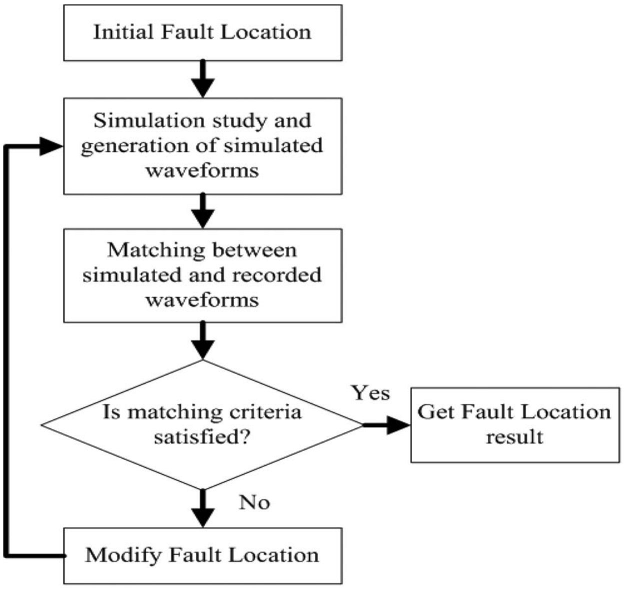


Fig. 4. *Flowchart of system wide sparse measurement method.*

one in the field. The overall process is described below using the following flowchart (see Fig. 4).

The waveforms can be matched using phasors or transients. In our present approach, field-recorded waveforms are used to calculate phasors, and they are in turn matched with the phasors obtained using short circuit study from short circuit model data (saved case data using PSS/ETM 30 was supplied by the utility).

The matching degree can be formulated as [33]

$$f_c(x, R_f) = \sum_{k=1}^{Nv} r_{kv}|V_{ks} - V_{kr}| + \sum_{k=1}^{Ni} r_{ki}|I_{ks} - I_{kr}| \qquad (1)$$

where

$f_c(x, R_f)$	the cost function used for phasors for matching;
x, R_f	the fault location and fault resistance;
r_{kv}, r_{ki}	weights for the errors of the voltages and currents, respectively;
V_{ks}, V_{kr}	simulated and recorded during-fault voltages, respectively;
I_{ks}, I_{kr}	simulated and recorded during-fault currents, respectively;
N_s, N_r	total number of voltage and current phasors to be matched, respectively;
k	the index of voltage or current phasors.

Theoretically the cost function should become zero when the simulated and recorded waveforms match completely. In a practical solution, the cost function is not zero and should be minimized. To obtain good waveform matching, the fault search range should be extensive. All possible faulted branches and fault resistance should be included in the search range which makes the search two-dimensional and exhaustive. For a large system, multiple searches should be run in parallel which can be achieved using population based optimization methods such as genetic algorithm (GA) [43]. Basically, in GA the objective is to maximize the fitness value. Therefore it can be applied to a maximization problem.

Equation (1) can be converted into a maximization problem by

$$f_f(x, R_f) = -f_c(x, R_f) \qquad (2)$$

where $f_f(x, R_f)$ is the fitness function to be used in GA.

F. Implementation

The architecture of the fault location scheme is shown in Fig. 5. The input data include the DFR data, the interpretation file for DFR data, the system model file in the PSS/E format and the PI Historian data matched to the model. The detailed description of this data requirement and handling of the data to extract information was discussed in earlier sections. The operation procedure of the software to utilize such information to obtain fault location is briefly described below.

The fault location solution using GA is performed in the following steps. First, the initial population is chosen randomly for this two dimensional (i.e., with two variables x, R_f) optimization problem. A fault location variable can be chosen from a range of zero to the length of the possible faulty line, and a fault resistance variable can be selected from typical possible fault resistance values. Second, short circuit studies are carried out using PSS/E, and the fitness is evaluated for each of the possible fault locations. Third, by using three GA operators (selection, crossover, and mutation), fault posing for the next iteration is obtained. By iteratively posing faults, running short circuit simulations, evaluating the fitness value, updating the fault location and resistance, the GA-based search engine guides the search process for a globally optimal solution.

G. Case Study

The software is implemented using the Java programming language. To interact between PSS/E activities and the Java programming language, IPLAN [44] language (which is a part of the PSS/E package) is used. The IPLAN language is able to modify the system topology, control the load flow and short circuit studies, and control the reporting of the results of the PSS/E activities. Like other programming languages, IPLAN language can be used to write programs, by which one can automatically control the PSS/E activities, as well as read and save the results in a disk file.

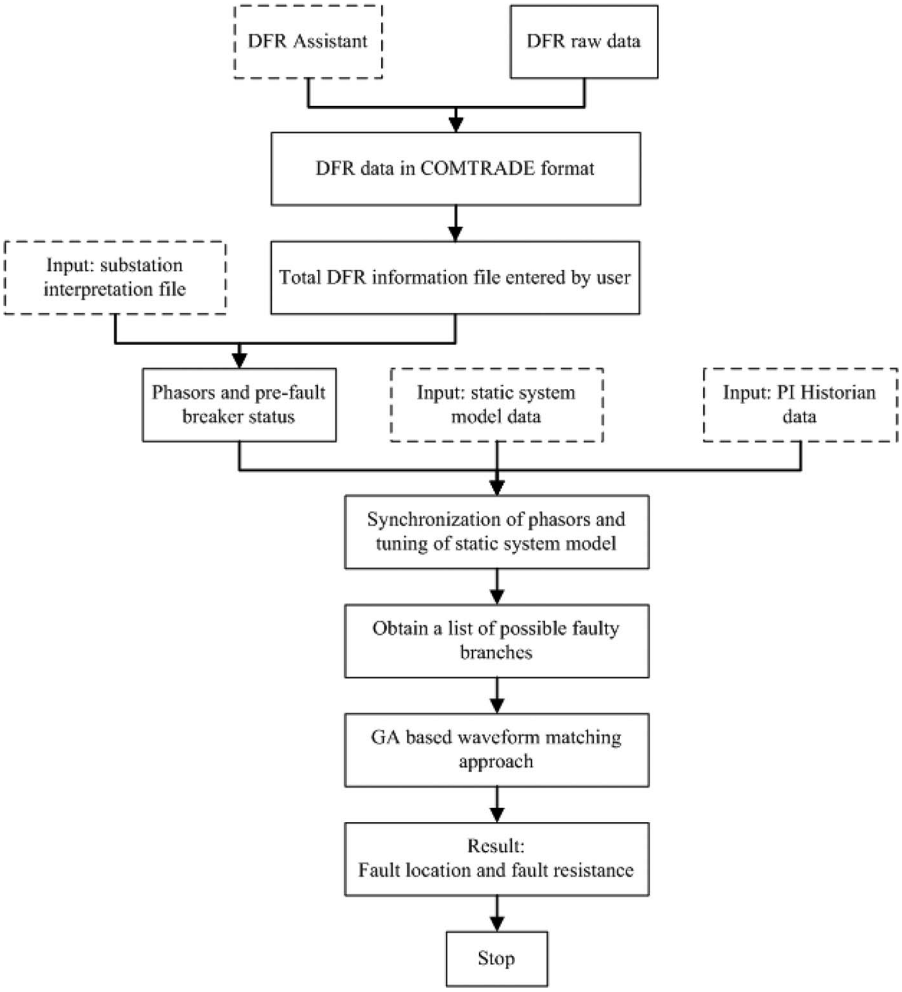


Fig. 5. ***Architecture of the fault location scheme.***

The parameters used for GA are: population size = 30, crossover probability = 0.8, mutation probability = 0.05, coding binary string length for fault location = 8, and coding binary string length for fault resistance = 8. Fault resistance ranges from 0 to 0.4 p.u.

Implementation of OFLA is evaluated for the following issues: using varying number of DFR files, specifying the search region, using preprocessed fault location estimation, using different quantities for the match between measured and simulated data, evaluating differences in accuracy when different input data are available and different assumptions are satisfied etc. These different options may produce different results. Test activities are conducted on the data collected from a real life electric power system [45], [46].

We have also tested the fault location scheme [36] for one test case provided by the utility. Two subsequent (in 5 ms gap) phase to ground faults occurred in one circuit between substation A and substation B. PI Historian data (extracted in *.xls format) is provided for both of the substations for a 10-s time interval for a duration from prefault to postfault. DFR triggered for only substation A for both of the faults. For the first fault, a DFR-resident algorithm could estimate fault location and for the second fault, the DFR-resident algorithm could not. The fault location scheme estimated the location accurately for both of the cases.

IV. ALARM PROCESSING

Alarm processing is an important part of power system operation that has been a traditional feature of the power system energy management system (EMS) studied over the past decades. Despite a variety of proposed solutions, operators still have a strong need for a better way to monitor the system than what is provided by the existing alarm processing software [10]. An EPRI study [11] has listed issues that operators face with alarms during their day-to-day operation of a power system:

- alarms which are not descriptive enough;

- alarms which are too detailed;
- too many alarms during a system disturbance;
- false alarms;
- multiplicity of alarms for the same event;
- alarms changing too fast to be read on the display;
- alarms not in priority order.

Operators are expected to monitor the system condition and take actions immediately after the alarms occur. However, when all problems mentioned above mix up, operators are severely constrained to perform efficiently in a timely manner.

A. Overview of Alarm Processing Algorithms

In recent years, several different concepts for filtering and suppressing alarms based on intelligent techniques have been implemented in many practical systems [12]. The major intelligent techniques used so far include the following.

Expert System (ES) technique: This technique [47]–[50] is well suited for a diagnosis problem such as fault section estimation because it mimics the behavior of fault analysis experts which perform fact-rule comparisons and search of consequent options. The disadvantage is that an expert system has to be developed using formalized knowledge that captures the expertise, which may require an extensive expert interviewing effort.

Fuzzy Logic (FL) technique: This approach offers a convenient means for modeling inexactness and uncertainties; hence a powerful solution to handle the imprecise and incomplete data may be implemented [51], [52]. The disadvantage is the need to have empirical data that helps determine the membership function and properties of fuzzy variables.

Petri-nets (PN) technique: It possesses the characteristics of graphical discrete event representation and parallel information processing [53]–[59].While very fast, the dynamic nature of the temporal change of the alarms cannot be easily captured with the standard Petri-net approach unless further adjustments are made.

Fuzzy Reasoning Petri-nets (FRPN) technique: This technique gains the advantages of expert system and fuzzy logic, as well as parallel information processing [50]–[59]. Some of the disadvantages of previously mentioned individual techniques may be offset by the benefits coming from combining the techniques.

B. An Improved Solution That Matches Data With Models

An implicit disadvantage of the traditional knowledge-based systems is that they may be incapable of handling complex scenarios that are not encountered during knowledge acquisition, implementation, or validation. They may also suffer from the slowness in analysis due to involved knowledge representation and inference mechanisms. Solutions based on discrete event view of Petri-nets also have several limitations. For instance, the number of initial inputs is limited and it is difficult to model inexactness and uncertainties. Consequently, to accurately identify fault sections under complex circumstances, substantial heuristic rules and information are additionally required [60].

An advanced FRPN diagnosis model is adopted in [60]. This intelligent alarm processor (IAP) model is expected to achieve the following goals:

- suppress multiple alarms from one event;
- generate a single conclusion through a logical cause-effect relationship;
- automate the process to get answers quickly;
- make graphical and numerical information concise and easy to follow.This approach assumes that filed data and model of the relay actions are matched leading to a cause-effect analysis described below.

C. Data Processing

It has been proven that the logic operand data of digital protective relays can be used as additional input to enhance the alarm interpretation [60]. In a digital protective relay, the pickup and operation information of protection elements is usually in the form of logic operands [61]. The pickup and operation logic operands are more reliable than SCADA data because they are more redundant and have less measurement uncertainty than relay trip signals and circuit breaker status signals.

In such a solution, input data such as relay trip signals and circuit breaker status signals are acquired by RTUs of the SCADA system. Relay logic operand signals are defined in their data memories and retrieved from relays by the SCADA front-end computers in substations. The data are acquired from different substations and transmitted to the control center through selected communication links such as microwave or optical fiber. In the control center, the SCADA master computer puts the input data into a real-time data base and keeps updating them at each scan time.

The detailed description of field data needed for this application is listed in Table 2.

D. Information Extraction

The alarm diagnosis algorithm and model used for this case are illustrated in an earlier paper [60]. The protection system configuration for this case is shown in Fig. 6.

The system consists of 9 sections, including 3 buses, 2 generators, and 4 transmission lines.

The alarm processing application includes two stage analyses.

- *First Stage*: The system's topology is analyzed based on circuit breaker status data from the real-time data base. The analysis includes all sections isolated by the opening of circuit breakers into a rough candidate set. The set is rough because it may include sections, which are not faulted but are isolated due to backup relay operation.

Table 2 Input Data List

Data from RTU of SCADA (Main data)	
1	CB status change alarms (Opening and Closing)
2	Trip signal of Main Transmission Line Relays
3	Trip signal of Primary Backup Transmission Line Relays
4	Trip signal of Secondary Backup Transmission Line Relays
5	Trip signal of Bus Relays
Data from Digital Protective Relays (*Additional data)	
1	Pickup & Operation signals of Main Transmission Line Relays
2	Pickup & Operation signals of Primary Backup Transmission Line Relays
3	Pickup & Operation signals of Secondary Backup Transmission Line Relays
4	Pickup & Operation signals of Bus Relays

- *Second Stage*: The FRPN diagnosis model as well as data in the real-time data base corresponding to each section in the rough candidate set is used and FRPN matrix operation is implemented.

When one or more faults occur on a given sections of the power system, protection devices will reach a certain status accordingly. The observed circuit breaker status signals obtained from RTUs of SCADA systems are used as inputs for estimation of the faulted sections. The logic reasoning method uses the relay status obtained from the online-database to validate each candidate fault section. The strategy is to build one FRPN diagnosis model for each section of the power system. Each model establishes reasoning starting from a set of SCADA data leading to the conclusion about fault occurrence on its section with a certain truth degree value.

E. Translational Knowledge Through Matching SCADA and Relay Data With the Reasoning Diagnosis Model

We use a backward reasoning concept to structure the FRPN diagnosis models and generalize the design for transmission lines and buses. Fig. 7 illustrates the backward reasoning concept for structuring transmission line and bus diagnosis models respectively [60]. The "AND-OR" structure concisely represents all the possible combinations of main, primary backup and secondary backup protection operations for inferring a fault.

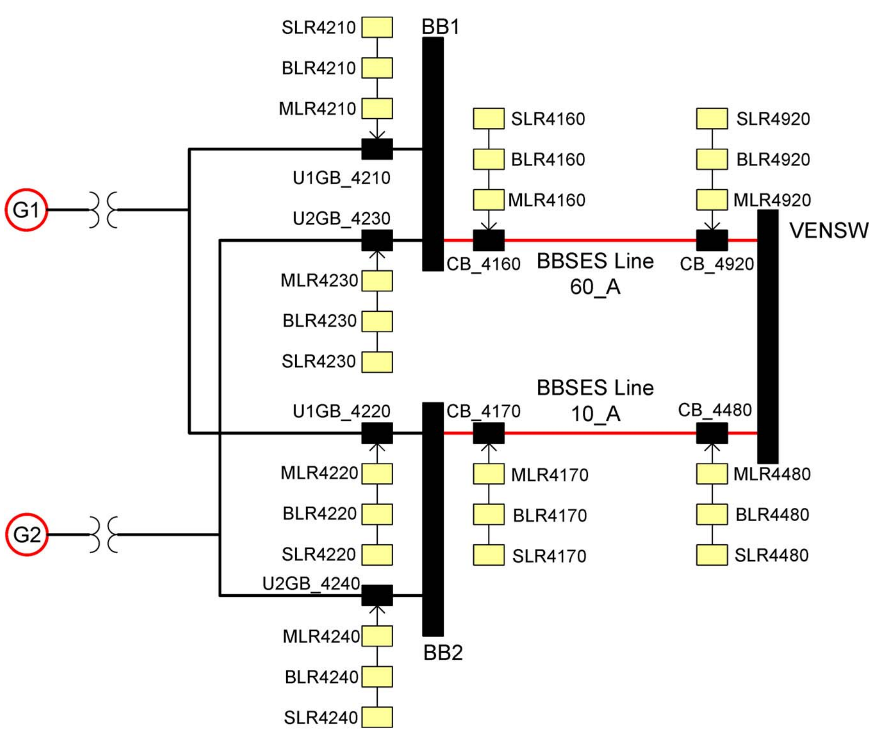


Fig. 6. *Protection system configuration diagram.*

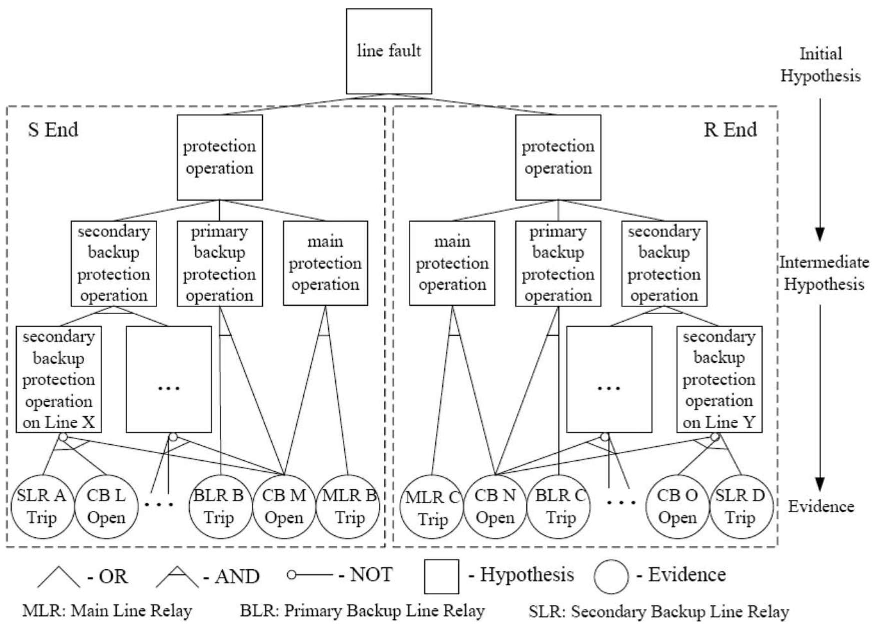


Fig. 7. ***Backward reasoning concept for structuring transmission line diagnosis models.***

F. Implementation

Based on the proposed structure introduced earlier [60], the FRPN diagnosis models are developed. As examples, Fig. 8 shows the FRPN models for the transmission line BBSES_60A and Unit 1, respectively.

Each proposition is given a "truth degree value" to illustrate the strength of confirmation. We use a "weighted average" operation when calculating the truth degree value of a consequent proposition from the truth degree values of its antecedent propositions. Fig. 9 illustrates the operation for r1 in Fig. 8.

The "weighted average" operation has two benefits [60]. First, the relative significance of antecedent propositions in implicating the consequent proposition is recognized by the weights of antecedent propositions. This is particularly meaningful when the cause-effect relation among antecedent propositions is considered. In our assumption, the circuit breaker opening is the effect of a relay trip. The "circuit breaker opens" proposition is generally given larger weight than that of the "relay trips" proposition because a circuit breaker opening indicates the completion of a protection operation more directly. For

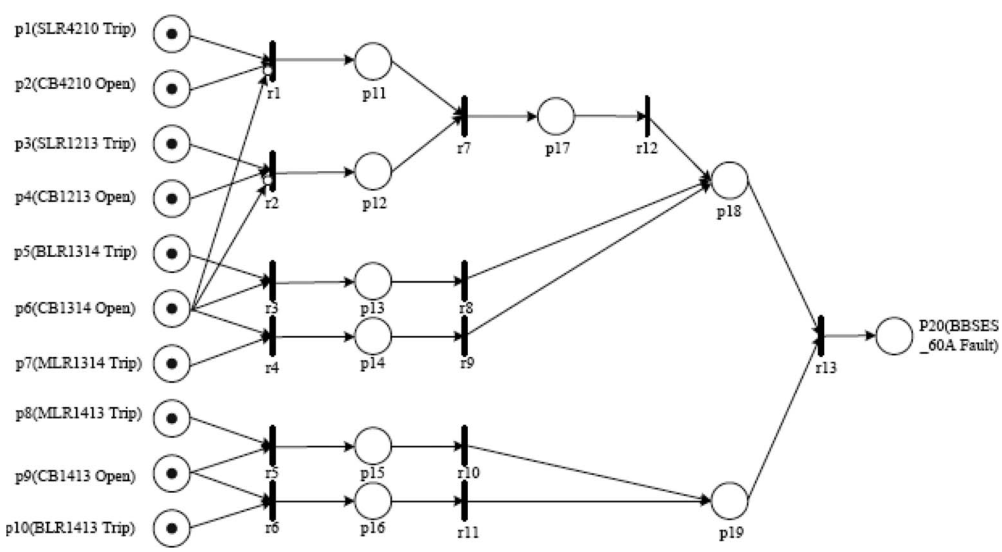
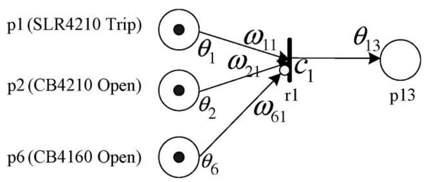


Fig. 8. ***A FRPN model for BBSES_60A fault.***



$$\theta_{13} = [\theta_1\omega_{11} + \theta_2\omega_{21} + (1-\theta_6)\omega_{61}]c_1$$

Fig. 9. ***An example of "weighted average" operation.***

example, regarding the rule r3 in Fig. 8, the proposition p5 "BLR4160 Trip" will be given a weight 0.4; the proposition p6 "CB4160 Open" will be given a weight 0.6.

Second, the false data problem is effectively handled by averaging the truth degree values of antecedent propositions. For example, when the relay MLR4160 trips and the circuit breaker CB4160 opens as a consequence of a fault on the line BBSES_60A, and "MLR4160 Trip" is not observed, p15, which stands for "main protection operates", it will still get a moderate truth degree value instead of 0, hence a moderate truth degree value for the final conclusion. It is apparent that the larger the number of input data, the impact of false data is more effectively countered.

G. Case Study

On September 5th 2007, a tornado in the area resulted in the tripping of two 345 kV lines and two generators in one of the substations. As a result of the unit trips, the bus split into two buses due to the configuration of the substation. Two cases taken from the SCADA database of the EMS responsible for monitoring the area are used. The first record was captured at approximately 07:49 AM, before the event occurred, and the second was captured at approximately 07:54 AM, just after the event occurred, as shown in Fig. 10.

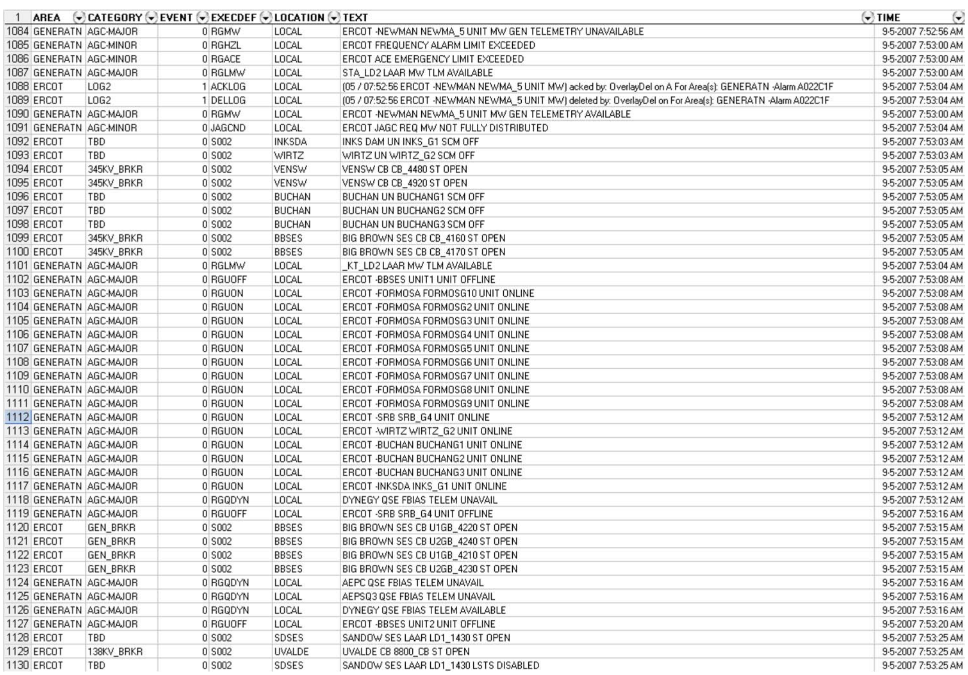
For the case in question, there were 2125 alarm messages that appeared within only 45 min. Obviously, this is beyond the capacity of any operator to handle. Thus, operators may not be able to respond to the unfolding events in time, and even worse, the interpretation by the operators may be either wrong or inconclusive. Operators also admitted that the list of large number of alarms provided little help for them.

Though the operator in this case was not able to provide relay data, our algorithm still works by having only SCADA data as inputs.

CASE 1: No protective relay signals. Circuit breaker CB4210, CB4220, CB4160, CB4920 status changes are detected as shown in Fig. 11.

Diagnosis result: Line BBSES_60A is faulted, and its truth value is 0.5130.

CASE 2: This report assumes the operation of the circuit breaker is tripped by the associated relays, thus allowing the relay status to be obtained to validate the fault. We assumed that we received the relay signals related to this case. All the devices worked correctly with no false signals. Operation of circuit breakers CB4210, CB4220, CB4160, CB4920 is detected as shown in Fig. 12.



1	AREA	CATEGORY	EVENT	EXECDEF	LOCATION	TEXT	TIME
1084	GENERATN	AGC-MAJOR	0	RGMW	LOCAL	ERCOT -NEWMAN NEWMA_5 UNIT MW GEN TELEMETRY UNAVAILABLE	9-5-2007 7:52:56 AM
1085	GENERATN	AGC-MINOR	0	RGHZL	LOCAL	ERCOT FREQUENCY ALARM LIMIT EXCEEDED	9-5-2007 7:53:00 AM
1086	GENERATN	AGC-MINOR	0	RGACE	LOCAL	ERCOT ACE EMERGENCY LIMIT EXCEEDED	9-5-2007 7:53:00 AM
1087	GENERATN	AGC-MAJOR	0	RGLMW	LOCAL	STA_LD2 LAAR MW TLM AVAILABLE	9-5-2007 7:53:00 AM
1088	ERCOT	LOG2	1	ACKLOG	LOCAL	(05 / 07:52:56 ERCOT -NEWMAN NEWMA_5 UNIT MW) acked by: OverlayDel on A For Area(s): GENERATN -Alarm A022C1F	9-5-2007 7:53:04 AM
1089	ERCOT	LOG2	1	DELLOG	LOCAL	(05 / 07:52:56 ERCOT -NEWMAN NEWMA_5 UNIT MW) deleted by: OverlayDel on For Area(s): GENERATN -Alarm A022C1F	9-5-2007 7:53:04 AM
1090	GENERATN	AGC-MAJOR	0	RGMW	LOCAL	ERCOT -NEWMAN NEWMA_5 UNIT MW GEN TELEMETRY AVAILABLE	9-5-2007 7:53:00 AM
1091	GENERATN	AGC-MINOR	0	JAGCND	LOCAL	ERCOT JAGC REQ MW NOT FULLY DISTRIBUTED	9-5-2007 7:53:04 AM
1092	ERCOT	TBD	0	S002	INKSDA	INKS DAM UN INKS_G1 SCM OFF	9-5-2007 7:53:03 AM
1093	ERCOT	TBD	0	S002	WIRTZ	WIRTZ UN WIRTZ_G2 SCM OFF	9-5-2007 7:53:03 AM
1094	ERCOT	345KV_BRKR	0	S002	VENSW	VENSW CB CB_4480 ST OPEN	9-5-2007 7:53:05 AM
1095	ERCOT	345KV_BRKR	0	S002	VENSW	VENSW CB CB_4920 ST OPEN	9-5-2007 7:53:05 AM
1096	ERCOT	TBD	0	S002	BUCHAN	BUCHAN UN BUCHANG1 SCM OFF	9-5-2007 7:53:05 AM
1097	ERCOT	TBD	0	S002	BUCHAN	BUCHAN UN BUCHANG2 SCM OFF	9-5-2007 7:53:05 AM
1098	ERCOT	TBD	0	S002	BUCHAN	BUCHAN UN BUCHANG3 SCM OFF	9-5-2007 7:53:05 AM
1099	ERCOT	345KV_BRKR	0	S002	BBSES	BIG BROWN SES CB CB_4160 ST OPEN	9-5-2007 7:53:05 AM
1100	ERCOT	345KV_BRKR	0	S002	BBSES	BIG BROWN SES CB CB_4170 ST OPEN	9-5-2007 7:53:05 AM
1101	GENERATN	AGC-MAJOR	0	RGLMW	LOCAL	_KT_LD2 LAAR MW TLM AVAILABLE	9-5-2007 7:53:04 AM
1102	GENERATN	AGC-MAJOR	0	RGUOFF	LOCAL	ERCOT -BBSES UNIT1 UNIT OFFLINE	9-5-2007 7:53:08 AM
1103	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -FORMOSA FORMOSG10 UNIT ONLINE	9-5-2007 7:53:08 AM
1104	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -FORMOSA FORMOSG2 UNIT ONLINE	9-5-2007 7:53:08 AM
1105	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -FORMOSA FORMOSG3 UNIT ONLINE	9-5-2007 7:53:08 AM
1106	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -FORMOSA FORMOSG4 UNIT ONLINE	9-5-2007 7:53:08 AM
1107	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -FORMOSA FORMOSG5 UNIT ONLINE	9-5-2007 7:53:08 AM
1108	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -FORMOSA FORMOSG6 UNIT ONLINE	9-5-2007 7:53:08 AM
1109	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -FORMOSA FORMOSG7 UNIT ONLINE	9-5-2007 7:53:08 AM
1110	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -FORMOSA FORMOSG8 UNIT ONLINE	9-5-2007 7:53:08 AM
1111	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -FORMOSA FORMOSG9 UNIT ONLINE	9-5-2007 7:53:08 AM
1112	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -SRB SRB_G4 UNIT ONLINE	9-5-2007 7:53:12 AM
1113	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -WIRTZ WIRTZ_G2 UNIT ONLINE	9-5-2007 7:53:12 AM
1114	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -BUCHAN BUCHANG1 UNIT ONLINE	9-5-2007 7:53:12 AM
1115	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -BUCHAN BUCHANG2 UNIT ONLINE	9-5-2007 7:53:12 AM
1116	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -BUCHAN BUCHANG3 UNIT ONLINE	9-5-2007 7:53:12 AM
1117	GENERATN	AGC-MAJOR	0	RGUON	LOCAL	ERCOT -INKSDA INKS_G1 UNIT ONLINE	9-5-2007 7:53:12 AM
1118	GENERATN	AGC-MAJOR	0	RGQDYN	LOCAL	DYNEGY QSE FBIAS TELEM UNAVAIL	9-5-2007 7:53:12 AM
1119	GENERATN	AGC-MAJOR	0	RGUOFF	LOCAL	ERCOT -SRB SRB_G4 UNIT OFFLINE	9-5-2007 7:53:16 AM
1120	ERCOT	GEN_BRKR	0	S002	BBSES	BIG BROWN SES CB U1GB_4220 ST OPEN	9-5-2007 7:53:15 AM
1121	ERCOT	GEN_BRKR	0	S002	BBSES	BIG BROWN SES CB U2GB_4240 ST OPEN	9-5-2007 7:53:15 AM
1122	ERCOT	GEN_BRKR	0	S002	BBSES	BIG BROWN SES CB U1GB_4210 ST OPEN	9-5-2007 7:53:15 AM
1123	ERCOT	GEN_BRKR	0	S002	BBSES	BIG BROWN SES CB U2GB_4230 ST OPEN	9-5-2007 7:53:15 AM
1124	GENERATN	AGC-MAJOR	0	RGQDYN	LOCAL	AEPC QSE FBIAS TELEM UNAVAIL	9-5-2007 7:53:16 AM
1125	GENERATN	AGC-MAJOR	0	RGQDYN	LOCAL	AEPSQ3 QSE FBIAS TELEM UNAVAIL	9-5-2007 7:53:16 AM
1126	GENERATN	AGC-MAJOR	0	RGQDYN	LOCAL	DYNEGY QSE FBIAS TELEM AVAILABLE	9-5-2007 7:53:16 AM
1127	GENERATN	AGC-MAJOR	0	RGUOFF	LOCAL	ERCOT -BBSES UNIT2 UNIT OFFLINE	9-5-2007 7:53:20 AM
1128	ERCOT	TBD	0	S002	SDSES	SANDOW SES LAAR LD1_1430 ST OPEN	9-5-2007 7:53:25 AM
1129	ERCOT	138KV_BRKR	0	S002	UVALDE	UVALDE CB 8800_CB ST OPEN	9-5-2007 7:53:25 AM
1130	ERCOT	TBD	0	S002	SDSES	SANDOW SES LAAR LD1_1430 LSTS DISABLED	9-5-2007 7:53:25 AM

Fig. 10. ***Alarm screen shot.***

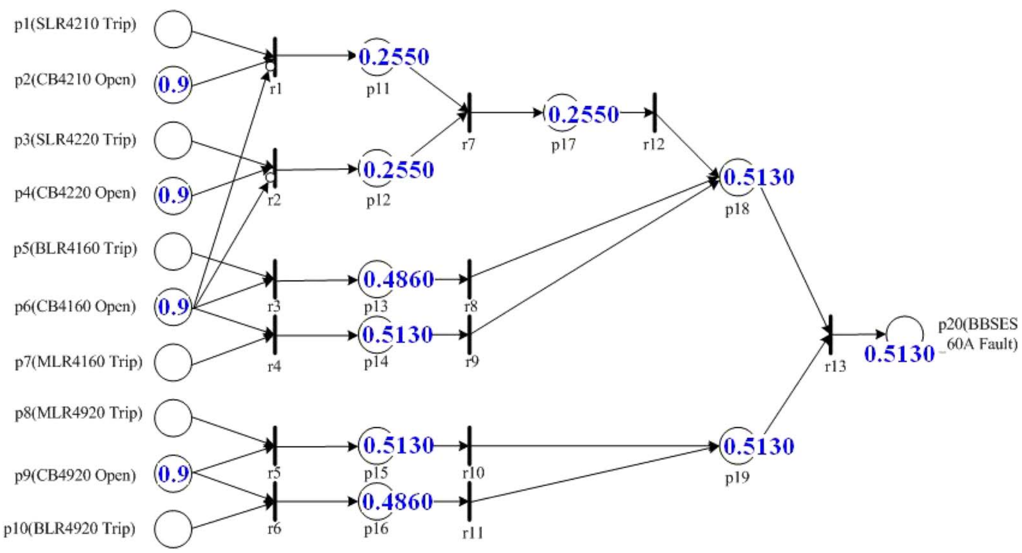


Fig. 11. *FRPN model analysis procedure for line BBSES_60A.*

Diagnosis result: Line BBSES_60A is faulted, and its truth value is 0.8550. With the input of the related relay signals, the fault certainness has been increased dramatically.

For this case, the extracted information is the cause-effect relationship between the fault event and alarms. The final conclusion will be that a fault occurs on the transmission line with a truth degree value. It greatly reduces the system operator's burden to deal with the overwhelming alarms during an event.

From the simulation test, we may draw the conclusion that by even using only SCADA data, our proposed IAP model works properly for the practical cases. Compared to current solutions, our model has the following advantages:

- The fault alarm analysis report can be generated automatically and immediately after the fault occurs.
- The FRPN models can be built in advance based on power system and protection system configurations and stored in files. In such a way, the FRPN models can be easily modified according to the changes of input data as well as power system and protection system configuration.

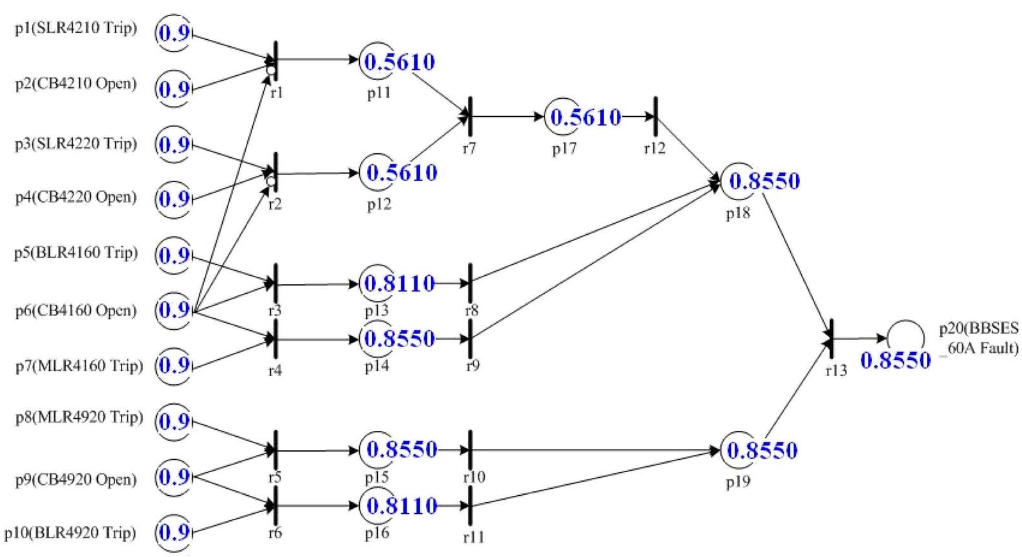


Fig. 12. *FRPN model analysis procedure for line BBSES_60A_with_assumed_relay_data.*

- This solution can use only SCADA data and does not need detailed data from IEDs or other measurement devices.

V. NEURAL NETWORK-BASED FAULT DETECTION AND CLASSIFICATION

The traditional transmission line protection schemes are mostly based on the calculation of three phase voltage and current phasor estimates and their comparison to predefined thresholds, which are a preset criteria. Calculating settings requires complex short circuit analysis to cover worst-case fault conditions and to coordinate the selectivity of each relay protection zone. The threshold-based algorithm needs extensive theoretical analysis and verification through elaborate field evaluations to make sure overload conditions and distributed generation injections are clearly differentiated from the faults. However, there are some inherent problems associated with the traditional distance relay principle. For example, the transient signal during a fault is a nonstationary signal containing fundamental frequency components, dc offset with damping, harmonics, etc. In some extreme situations, it will result in an inaccurate estimation of phasor representation of the faulted signal, which may cause the false judgment in relays [62]. This was actually the case in the 2003 Northeastern blackout in the United States and Canada when the relays could not differentiate between faults and overload conditions and cascading trips took place, bringing the entire interconnection into a blackout that cost the economy an estimated $50 billion [63].

A. Overview of Neural Network (NN) Algorithms

Neural network (NN) is an artificial intelligence technique, which has been studied in the power system area [14]–[16], and applied to solutions such as fault detection and classification [17], [64]–[70], fault location [71]–[73], fault direction discrimination [74], etc. By resembling the human brain, the neural network works as a parallel distributed processor made up of simple processing units, which has a natural capability for storing intelligence and making it available for generalization [75]. A neuron is the most fundamental information-processing unit of a neural network. The three basic elements of a neuron are synaptic weights, a summing junction, and the activation function. The input–output mapping of a neuron can be expressed by the following two equations [76]:

$$u_k = \sum_{j=1}^{n} \omega_{kj} x_j \tag{3}$$

$$y_k = \varphi(u_k + b_k). \tag{4}$$

Fig. 13 shows the structure of a single neuron.

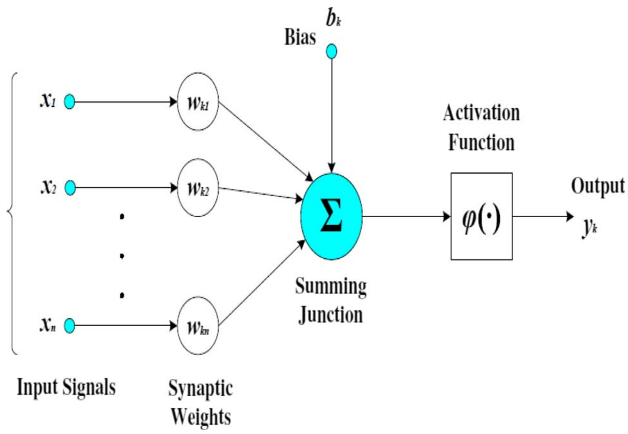


Fig. 13. *Basic neuron model of a neural network.*

The NN learning process shows the procedure of converting the power system field data to information, which then can be processed to form the knowledge, which is based on the interneuron connection strengths, or synaptic weights. In this section, we will discuss what data is required for neural network based fault detection and classification; what kind of model is used to convert the data to information; and how to the translation from information to knowledge leads to the final decision making step.

B. An Improved Protective Relaying Solution That Matches Data Patterns to Fault Types

The purpose of a NN-based protective relaying algorithm is to match a given set of input data patterns into several groups or clusters, so that each data pattern is assigned to a unique cluster. The procedure includes neural network training and testing.

Neural network training is the most important step when applying neural networks. The learning techniques for some neural networks can be classified into two broad categories: supervised learning and unsupervised learning [76]. In supervised learning, each input signal is associated with the labeled output. The task is the input–output mapping by adjusting the synaptic weights to minimize the overall error between the entire output set and their corresponding input data set. In unsupervised learning, the categories of the outputs are not known in advance. The network is self-organized by clustering techniques to identify the mutual similarity of the input patterns. The task is to adjust the network weights until the similar inputs can produce similar outputs. A neural network based on a combined unsupervised/supervised training scheme has proven to be more capable of handling large data sets of random fault scenarios than solely using supervised training schemes [19], [67]. The input data vectors of samples from voltage and/or current signals are mapped into clusters that contain information about fault existence and type.

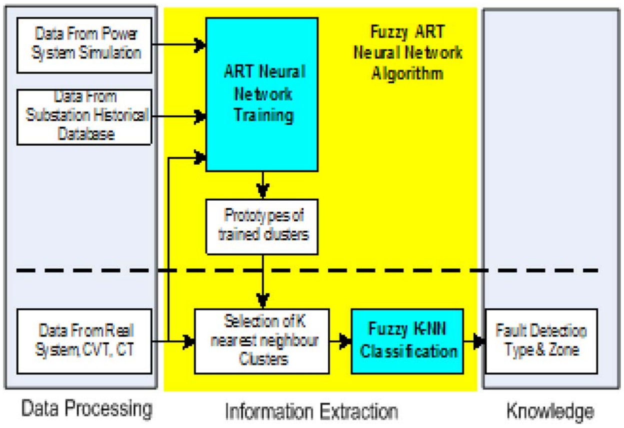


Fig. 14. *Fuzzy ART neural network algorithm for fault diagnosis.*

During the testing procedure, distances between each test pattern and established clusters are calculated. The outcomes of the testing are class labels assigned to test patterns according to the most common value among the K nearest prototypes.

The input into the neural network will be in the form of a moving data window containing samples of voltage and current signals from local measurements and simulations. Patterns are extracted from these data and placed together in one row to form feature vector components, which is shown in Fig. 17. Fig. 17 shows the patterns of the sampled data. The pattern is arranged using the postfault samples of three phase voltage and current signals. The zero sequence values of voltage $3v_0 = v_a + v_b + v_c$ and current $3i_0 = i_a + i_b + i_c$ are also included to precisely detect ground faults. In this case, all fault types can be differentiated very well [76]. Thousands of such patterns obtained from power system simulation or substation databases of field recordings are used to train the neural network offline, and then the pattern prototypes are used to analyze faults online by using the Fuzzy K-NN classifier.

C. Data Processing

Compared with traditional methods, neural network-based fault-diagnosis algorithms use the time-domain data: three-phase voltage and current samples as input instead of calculating phasors, which is shown in Fig. 14.

The neural network algorithm requires a large number of fault and non-fault cases to complete the process of training and testing for neural network tuning. Those training and testing cases are quite different for various transmission lines due to the selection of different simulation parameters and settings. To perform comprehensive tests, two categories of data will be acquired: field signal measurements and data from simulation cases. Since it is not possible to acquire enough fault cases from the field, generating the needed data files by simulation provide an alternative solution. Fig. 15 shows a block diagram for the fault and non-fault scenarios generation which is based on ATP/ATPDraw [77] and Matlab [78].

The power system of interest is first modeled in ATP/ATPDraw. A user can define the desired fault or nonfault cases by initializing the simulation setting parameters in Matlab. The measured three-phase voltage and current samples, which could also include time stamp, are extracted in the data format files defined by a user.

D. Information Extraction

The feature extraction is the procedure of information processing, which is shown in Fig. 14. The detailed unsupervised and supervised learning phases of neural network training are shown in Fig. 16 [67].

The neural network algorithm compares information from input voltage and current signals with prototypes instead of predetermined settings.

Typically, the data window length in each phase is one cycle or half cycle of the fundamental frequency signal. A longer window increases the number of samples for training and gives better information about the original signals. but it prolongs the training and testing procedures. The sampling frequency has a similar effect on forming the patterns as the data window does. Increased sampling frequency offers improved signal detection but also causes significant computational burden.

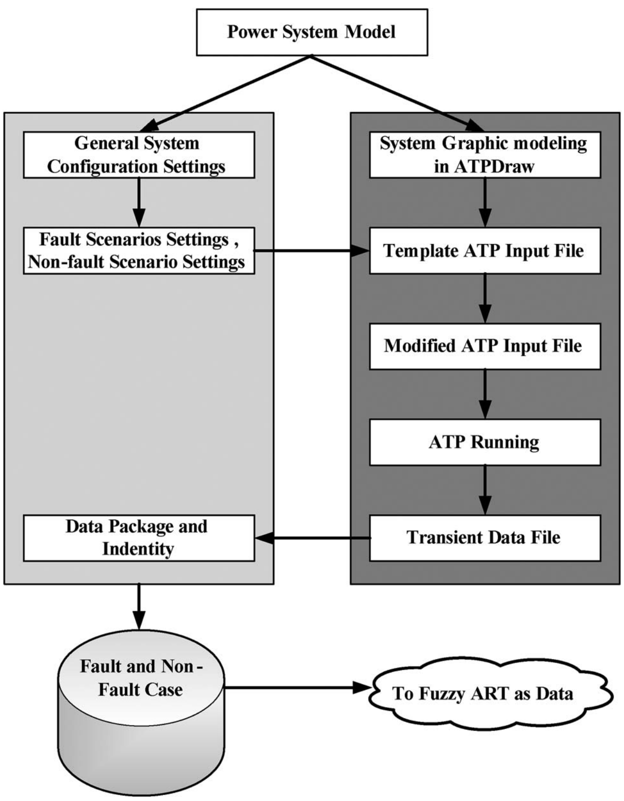


Fig. 15. *Simulations for fault and nonfault scenarios.*

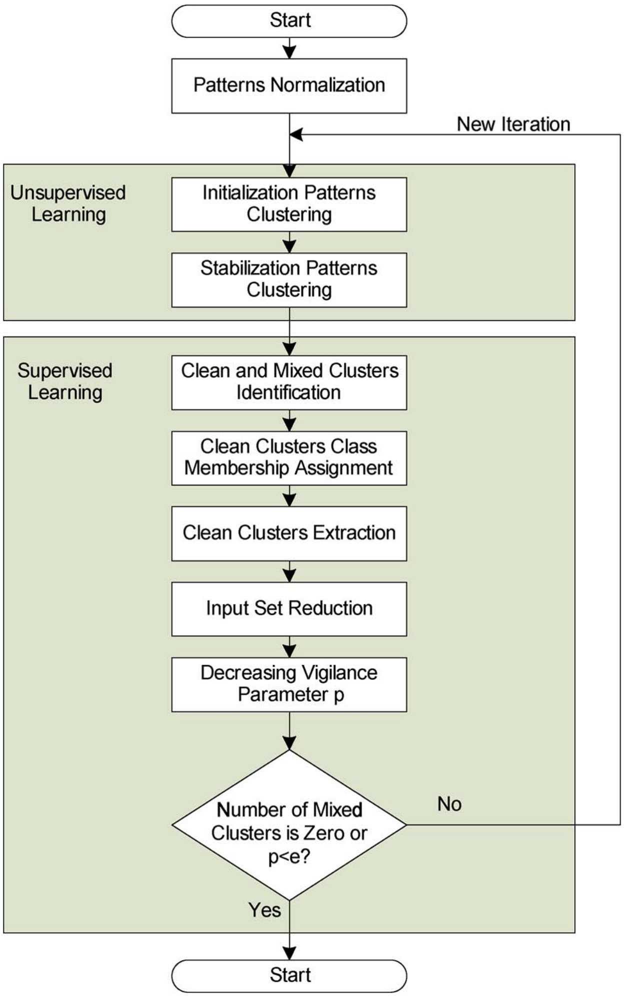


Fig. 16. ***Neural network clustering algorithm.***

E. Detecting and Classifying Faults by Mapping Data to Labeled Clusters

The purpose of the information processing is to form the knowledge, which could allocate the training patterns into homogeneous clusters by some grouping technique. Then the clusters are assigned to the classes, which are our expected fault events in the power system, such as fault type, etc. The number of clusters is increased and their positions are updated automatically during the learning, and there is no need to define them in advance. Fig. 18 shows the raw training patterns as information. Fig. 19 shows the clusters after the training processing, which is the knowledge of the neural network based fault detection and classification.

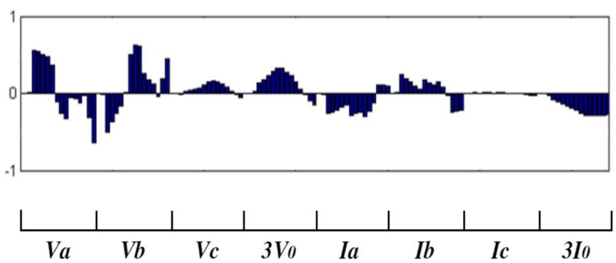


Fig. 17. ***Pattern arrangement for proposed neural network algorithm-data.***

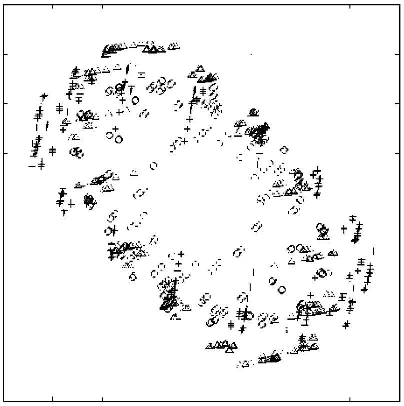


Fig. 18. ***The raw training patterns—Information.***

The typical types of classification are based on detecting the fault type and fault zone. The classification of mapping data to labeled clusters is performed by using the K-nearest neighbor rule (K-NN) [19].

Given a set of classified clusters, the standard K-nearest neighbors rule determines the classification of the input pattern x_i based only on the class labels of the K closest clusters in the cluster structure established during training:

$$\mu_j(x_i) = f\left[K, \mu_j(v_l)\right] \tag{5}$$

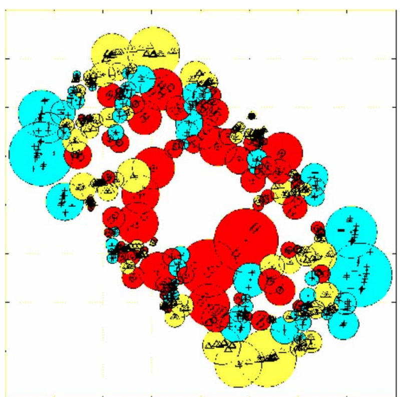


Fig. 19. ***The patterns are allocated to the clusters after a training processing—knowledge.***

where

$\mu_j(v_l)$ is membership value which determines the degree of belonging of cluster l to class j;

$\mu_j(x_i)$ is membership value of pattern i belonging to class j;

$i = 1, \ldots, P$ where P is number of patterns;

$j = 1, \ldots, C$ where C is number of classes;

$l = 1, \ldots, K$ where K is number of neighbors;

$v_1, v_2, \ldots, v_K$ denotes the centers of K nearest neighbors of pattern x_i.

$\mu_j(v_l)$ has only crisp values 0 or 1, depending on whether or not a cluster v_l belongs to class j:

$$\mu_j(v_l) = \begin{cases} 1, & \text{if cluster } l \text{ belongs to class } j \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

F. Implementation

K-NN rule based clusters have the equal importance, without taking into account their radii, and distances to the pattern that have to be classified. Thus when implementing the detecting and classifying faults by mapping data to labeled clusters, the advanced K-nearest neighbor technique, Fuzzy K-NN, is discussed and used to deal with transmission line fault diagnosis issues in a line protection scheme [67], [68]. The theoretical background can be found in [79], [80].

New patterns are classified based on the weighted distances (d_l) to K nearest clusters, as well as on relative size (r_l) and class labels (c_l) of these clusters. The Fuzzy K-NN calculates a vector of membership values $(\mu_1(x)\mu_2(x), \ldots, \mu_c(x))$ of input pattern x_i in the existing classes. The class membership values are calculated based on the following equation:

$$\mu_j(x_i) = f\left[K, \mu_j(v_l), d_l(x_i)\right] \tag{7}$$

where now may $\mu_j(v_l)$ take any value between 0 and 1, representing the relative size of the actual cluster l. Each cluster belongs to one of the existing classes, with its membership value defined by the following adopted relation:

$$\mu_j(v_l) = \begin{cases} \frac{r_l}{r_{\max}}, & \text{if cluster } l \text{ belongs to class } j \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

The membership degree of cluster v_l belonging to class j is equal to the ratio between radius (r_l) of actual cluster l and radius $(r_{\max})$ of the largest cluster in the cluster structure. The outcome is that the larger clusters have more influences than the smaller ones, and the clusters with longest radius have $\mu_j(v_l) = 1$.

By taking into account distances between pattern x_i and K nearest clusters, the distance $d_l(x_i)$ is selected to be a weighted Euclidean distance between pattern x_i and cluster l

$$d_l(x_i) = \|x_i - v_l\|^m. \tag{9}$$

The new fuzzier classification algorithm will help classify better a variety of input patterns.

G. Case Study

In this case, a simulation system for the 500 kV transmission line is built based on the ATP/ATPDraw and Matlab [81]. A large number of fault and nonfault cases has been generated, which includes different fault types, fault locations, fault resistance, and fault angles. The fault is initiated at 0.02 s into the simulation process, and cleared at 0.45 s.

The proposed neural network is trained by using the simulated fault and nonfault cases. There are 209 clusters altogether determined with labels of different fault types. Then 5000 cases are tested for the trained neural network. Two classification algorithms are used when performing the test procedures: the nearest neighbor algorithm and the fuzzy K-nearest neighbor algorithm. Fig. 20 shows the errors for the fault classification for the basic nearest neighbor algorithm and the fuzzy 4-nearest neighbor algorithm. From Fig. 19, we can see that the error for fuzzy 4-NN is stable at about 1.5%.

Using the prototypes of trained clusters, the Fuzzy K-NN classifier takes into account both the effect of weighted distances and the size of neighboring clusters for distinguishing new patterns. It is proven that it has better performance than a common K-nearest neighborhood classifier. The advantage of neural network based fault detection algorithms is that the neural network can form its knowledge translating it from information describing the clusters.

VI. OTHER APPLICATIONS

The examples of the translational knowledge approaches described in previous sections relate to three distinct applications related to fault analysis: Fault detection and classification, fault location, and alarm processing. If there is a single fault in the system each of the applications will perform as discussed in the related sections. A more complex translation knowledge processing occurs in a case of multiple faults. To further discuss such cases it is worth differentiating different types of multiple faults and their consequence on the approaches described earlier. The multiple faults may be classified as follows.

- *Evolving faults*: the fault starts as one type and ends up in a different type (as an example, the fault

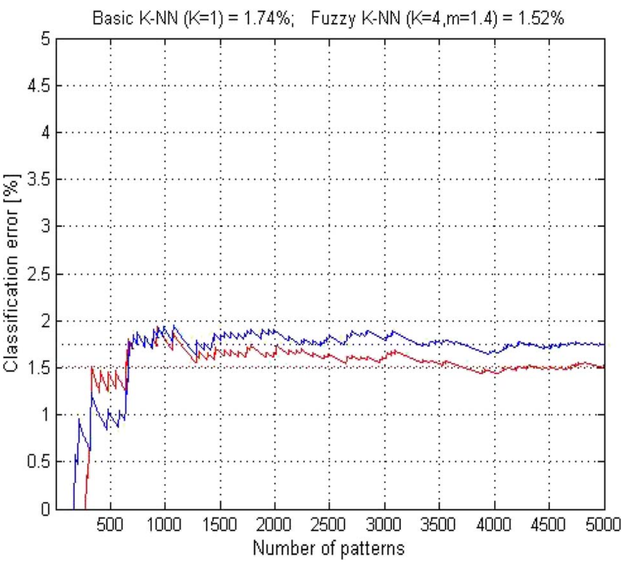


Fig. 20. *Error results of neural network fault classification tools.*

evolves from single phase-to-ground to two-phases-to-ground).

- *Cross-country faults*: such faults involve one transmission line and then extend to another (as an example, this may happen when two different transmission lines are on the same transmission tower).The lines may or may not begin/end in the same substations.
- *Multiple time-correlated faults*: Unrelated faults occurring in different parts of the system but in very short time proximity to each other (as an example, a storm comes through a region and lightning causes many faults where one may be initiated before the other clears). This may happen on the lines that may or may not begin/end in the same substation.

In such situation, while quite rare, the proposed applications are expected to perform as follows:

A. Fault Detection and Classification

The proposed technique is a pattern recognition based technique that eventually uses simulated or actual recorded cases for the NN training. If the above described cases can be replicated through simulation and/or captured record of actual waveform, the proposed technique can learn how to deal with such cases

B. Fault Location

The proposed technique is based on matching between phasors generated from a short circuit study and phasors extracted from measured waveforms captured in the field during a fault. Prior to the use of the technique the topology of the system and model of the system need to be determined and "calibrated" using SCADA PI Historian data. For multiple faults as described above the challenge becomes to model them in the short circuit program, capture the waveforms, and determine the model and topology using PI Historian data. If from the recorded waveforms it may be determined which type of the multiple fault has occurred, which requires further expertise beyond what was discussed in this paper, then the corresponding system model may be used for short circuit study, and the technique may be applied as proposed. If the representation of multiple faults in the model gets constrained due to the short circuit modeling tool limitations, a time domain modeling tool such as electromagnetic transient program (EMTP) may be used for more accurate model representation. In that case the technique will need to be extended to allow match between the simulated and recorded waveforms rather than phasors as originally proposed.

C. Alarm Processing

In this step one is really not dealing with fault characterization but rather with interpretation of alarms that come from relays, disturbance recorders, sequence of event recorders, etc. Hence it becomes important to understand how such devices react to multiple faults, and what types of alarms are generated in that case. Again, such study requires expertize beyond what is presented in this paper. If such expertize is developed and clear correlation between the various types of multiple faults and alarms is established, then the proposed technique may be applied. The models for cause–effect relationship between the alarms and types of faults would need to be developed, but once developed the same approach for drawing conclusions as originally proposed in this paper would be applicable.

The above cases of multiple faults have not been studied in this paper for a simple reason: the use of translation knowledge assumes that the overall process is eventually automated and hence provides a quick way of analyzing simple (and most common) fault cases. A common attempt by many in the past was to develop complex translational knowledge techniques that can help automate the analysis process for complex events such as multiple faults. The experience of this author has been that if one attempts to automate cases that are overly complex one runs into a "diminishing returns" situation rather quickly: the time spent on developing the solutions, and consequently maintaining them and updating them through the life-cycle of the solution, becomes so costly that the automation is not justified, and the whole approach defeats its basic purpose of being easy and cost effective to use and apply. This leads to practical recommendation: as one develops automated means for the translation knowledge techniques one should make a judgment what level of complexity is worth implementing using automated solutions and when one has to start relying on sound judgment of an expert that uses multiple means, many of them manual, to come to a conclusion. The case in point, multiple faults, may be better analyzed through an experienced person gathering different sources of information manually or automatically and making some judgment through quite often heuristic approach. The final massage of this experience is notable: complexity of the translational knowledge approach need to be carefully managed when practical automated solutions are developed and implemented.

VII. CONCLUSION

It is widely understood that smart grid solutions create large amounts of field data that has potential value to improve power system monitoring, control and protection. This paper has demonstrated how data may be converted to information, and then matched to different types of models for different applications to create knowledge for control actions. The benefits of this process, termed translational knowledge, are multiple:

- in the fault location case, an optimal solution was achieved by matching field data with model data in the cases when no other algorithms could reach the result;
- in the alarm processing case, by matching data to the model of relaying logic, a compact cause/effect analysis is achieved that is not possible with legacy solutions;
- in the protective relaying case, matching data vectors from the field signals with cluster of patterns designating fault types created an inherently adaptive protection that does not rely on accuracy of settings as in traditional relays. ■

Acknowledgment

The author gratefully acknowledge the contributions of his former Ph.D. students Y. Liao, S. Vasilic, and X. Luo whose dissertation work has been incorporated into this paper. His current Ph.D. students B. Matic-Cuka, P. Dutta, Y. Guan, and P. Chengzong are acknowledged for the extension of the prior work and case studies reported in the paper.

REFERENCES

[1] M. Kezunović, A. Abur, A. Edris, and D. Šobajić, "Data integration/exchange Part II: Future technical and business opportunities," *IEEE Power Energy Mag.*, pp. 24–29, May/Jun. 2004.

[2] M. Kezunović, A. Abur, A. Edris, and D. Šobajić, "Data integration/exchange Part I: Existing technical and business opportunities," *IEEE Power Energy Mag.*, pp. 14–19, Jan./Feb. 2004.

[3] P. Myrda, M. Kezunovic, S. Sternfeld, D. R. Sevcik, and T. Popovic, "Converting field recorded data to information: New requirements and concepts for the 21st century automated monitoring solutions," in *Proc. CIGRE General Session*, Paris, France, Aug. 2010.

[4] E. R. Dougherty, "Translational science: Epistemology and the investigative process," *Current Genom.*, vol. 10, pp. 102–109, 2009.

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

[6] *IEEE Guide for Determining Fault Location on AC Transmission and Distribution Lines*, IEEE Inc., IEEE Std. C37.114-2004, Jun. 2005.

[7] Y. Liao, "Automated analysis of power quality and transmission line fault location," Ph.D. dissertation, Texas A&M University, College Station, TX, May 2007.

[8] M. Kezunovic and Y. Liao, "Fault location estimation based on matching the simulated and recorded waveforms using genetic algorithms," in *Proc. Develop. Power Syst. Protection*, Amsterdam, The Netherlands, Apr. 2001.

[9] M. Kezunovic and P. Dutta, "Fault location using sparse wide area measurement," in *Proc. CIGRE Study Committee B5 Annu. Meeting Colloquium*, Jeju, Oct. 2009.

[10] B. F. Wollenberg, "Feasibility study for an energy management system intelligent alarm processor," *IEEE Trans. Power Syst.*, vol. PWRS-1, no. 2, pp. 241–246, May 1986.

[11] EPRI EL, *Human Factors Review of Electric Power Dispatch Control Centers*, vol. 1–6, Electric Power Research Institute, Palo Alto, CA.

[12] M. Pfau-Wagenbauer and W. Nejdl, "Integrating model-based and heuristic features in a real-time expert system," *IEEE Expert*, pp. 12–18, Aug. 1993.

[13] X. Luo, "Power system fault analysis based on intelligent techniques and IED data," Ph.D. dissertation, Texas A&M University, College Station, TX, May 2007.

[14] R. Aggarwal and Y. Song, "Artificial neural networks in power systems: Part I—General introduction into neural computing," *Power Eng. J.*, vol. 11, no. 3, pp. 129–134, Jun. 1997.

[15] R. Aggarwal and Y. Song, "Artificial neural networks in power systems: Part II—Types of artificial neural networks," *Power Eng. J.*, vol. 12, no. 1, pp. 41–47, Feb. 1998.

[16] R. Aggarwal and Y. Song, "Artificial neural networks in power systems: Part III—Examples of applications in power systems," *Power Eng. J.*, vol. 12, no. 6, pp. 279–287, Dec. 1998.

[17] M. Kezunovic, "A survey of neural net application to protective relaying and fault analysis," *Eng. Intell. Syst.*, vol. 5, no. 4, pp. 185–192, Dec. 1997.

[18] Y. H. Song, Q. X. Xuan, and A. T. Johns, "Comparison studies of five neural network based fault classifiers for complex transmission lines," *Electr. Power Syst. Res.*, vol. 43, no. 2, pp. 125–132, Nov. 1997.

[19] S. Vasilic, "Fuzzy neural network pattern recognition algorithm for classification of the events in power system networks," Ph.D. dissertation, Texas A&M University, College Station, Texas, USA, 2004.

[20] M. Kezunovic and A. Abur, "Merging the temporal and spatial aspects of data and information for improved power system monitoring applications," *Proc. IEEE*, vol. 93, no. 11, pp. 1909–1919, Nov. 2005.

[21] W. Lewandowski, J. Azoubib, and W. J. Klepczinski, "GPS: Primary tool for time transfer," *Proc. IEEE*, vol. 87, pp. 163–172, Jan. 1999.

[22] Google Earth. [Online]. Available: http://www.google.com/earth/index.html

[23] K. L. Cummins, J. A. Cramer, C. J. Biagi, E. P. Krider, J. Jerauld, M. A. Uman, and V. A. Rakov, "The U.S. National Lightning Detection Network: Post-upgrade status," in *Proc. Second Conf. Meteorolog. Applicat. Lightning Data*, Jan. 2006.

[24] D. J. Maguire, M. Batty, and M. F. Goodchild, *GIS, Spatial Analysis, and Modeling*, 1st ed. Redlands, CA: ESRI Press, 2005.

[25] T. Takagi, Y. Yamakoshi, M. Yamaura, R. Kondow, and T. Matsushima, "Development of a new type fault locator using the one-terminal voltage and current data," *IEEE Trans. Power Appl. Syst.*, vol. 101, no. 8, pp. 2892–2898, Aug. 1982.

[26] E. O. Schweitzer, III, "Evaluation and development of transmission line fault locating techniques which use sinusoidal steady-state information," in *Proc. 9th Annu. Western Protective Relay Conf.*, Spokane, WA, Oct. 26–28, 1982.

[27] L. Ericsson, M. Saha, and G. D. Rockefeller, "An accurate fault locator with compensation for apparent reactance in the fault resistance resulting from remote-end in feed," *IEEE Trans. Power Appl. Syst.*, vol. PAS-104, no. 2, pp. 422–436, Feb. 1985.

[28] M. Kezunovic, B. Perunicic, and J. Mrkic, "An accurate fault location algorithm using synchronized sampling," *Electr. Power Syst. Res. J.*, vol. 29, no. 3, pp. 161–169, May 1994.

[29] A. A. Gigris, D. G. Hart, and W. L. Peterson, "A new fault location technique for two- and three-terminal lines," *IEEE Trans. Power Delivery*, vol. 7, no. 1, pp. 98–107, Jan. 1992.

[30] M. S. Sachdev and R. Agarwal, "A technique for estimating line fault locations from digital impedance relay measurements," *IEEE Trans. Power Delivery*, vol. PWRD-3, no. 1, pp. 121–129, Jan. 1988.

[31] D. Novosel, D. G. Hart, E. Udren, and J. Garitty, "Unsynchronized two-terminal

fault location estimation," *IEEE Trans. Power Delivery*, vol. 11, no. 1, pp. 130–138, Jan. 1996.

[32] Y. Liao and S. Elangovan, "Improved symmetrical component-based fault distance estimation for digital distance protection," *IEE Proc.-Generation, Transmission and Distribution*, vol. 145, no. 6, pp. 739–746, Nov. 1998.

[33] D. A. Tziouvaras, J. Roberts, and G. Benmouyal, "New multi-ended fault location design for two- or three-terminal lines," in *Proc. Develop Power System Protection*, Amsterdam, The Netherlands, Apr. 9–12, 2001.

[34] A. Johns and S. Jamali, "Accurate fault location technique for power transmission lines," *IEE Proc.*, vol. 137, no. 6, pt. C, pp. 395–402, Nov. 1990.

[35] L. Lewis, "Travelling wave relations applicable to power system fault locators," *AIEE Trans.*, pp. 1671–1680, 1951.

[36] F. Gale *et al.*, "Fault location based on travelling waves," in *Proc. 5th Int. Conf. Develop. Power Syst. Protection*, 1993, pp. 54–59.

[37] F. Gale, J. Stokoe, and P. Crossley, "Practical experience with travelling wave fault locators on Scottish Power's 275 & 400 kV transmission system," in *Proc. 6th Int. Conf. Develop. Power System Protection*, Mar. 25–27, 1997, pp. 192–196.

[38] "Integration of substation IED information Into EMS functionality," PSerc Project T-32 Final Report (08-08). [Online]. Available: http://www.pserc.org

[39] PSSE Load Flow Siemens Power Transmission & Distribution, Inc., PTI, Schenectady, NY.

[40] [Online]. Available: http://techsupport.osisoft.com/

[41] Test Laboratories International, Inc., DFR assistant—Software for automated analysis and archival of DFR records with integrated fault location calculation. [Online]. Available: http://www.tliinc.com

[42] Y. Guo, M. Kezunovic, and D. Chen, "Simplified algorithms for removal of the effect of exponentially decaying dc-offset on the Fourier algorithm," *IEEE Trans. Power Delivery*, vol. 18, no. 3, pp. 711–717, Jul. 2003.

[43] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison Wesley, 1989.

[44] *IPLAN Program Manual*, PSS E 31.0, Release 17.

[45] M. Knezev, "Optimal fault location," M.S. thesis, Dept. of Electrical Engineering, Texas A&M University, College Station, TX, 2007.

[46] M. Kezunovic, "Optimized fault location," Apr. 2008, Final report for Concurrent Technologies Corporation, PSerc publication 08-07.

[47] A. A. Girgis and M. B. Johns, "A hybrid expert system for faulted section identification, fault type classification and selection of fault location algorithms," *IEEE Trans. Power Delivery*, vol. 4, no. 2, pp. 978–985, Apr. 1989.

[48] Y. L. Zhu, Y. H. Yang, B. W. Hogg, W. Q. Zhang, and S. Gao, "An expert system for power systems fault analysis," *IEEE Trans. Power Syst.*, vol. 9, no. 1, pp. 503–509, Feb. 1994.

[49] E. M. Vazquez, O. L. M. Chacon, and H. J. F. Altuve, "An on-line expert system for fault section diagnosis in power systems," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 439–444, May 2002.

[50] Y. C. Huang, "Fault section estimation in power systems using a novel decision support system," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 439–444, May 2002.

[51] H. J. Cho and J. K. Park, "An expert system for fault section diagnosis of power systems using fuzzy relations," *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 342–348, Feb. 1997.

[52] W. H. Chen, C. W. Liu, and M. S. Tsai, "On-line fault diagnosis of distribution substations using hybrid cause-effect network and fuzzy rule-based method," *IEEE Trans. Power Delivery*, vol. 15, no. 2, pp. 710–717, Apr. 2000.

[53] K. L. Lo, H. S. Ng, and J. Trecat, "Power systems fault diagnosis using Petri nets," *IEE Proc. Generation, Transmission and Distribution*, vol. 144, no. 3, pp. 231–236, May 1997.

[54] K. L. Lo, H. S. Ng, D. M. Grant, and J. Trecat, "Extended Petri-net models for fault diagnosis for substation automation," *IEE Proc. Generation, Transmission and Distribution*, vol. 146, no. 3, pp. 229–234, May 1999.

[55] A. A. E. Fergany, M. T. Yousef, and A. A. E. Alaily, "Fault diagnosis of power systems using binary information of breakers and relays through DPNs," in *Proc. Int. Conf. Power Syst. Technol.—PowerCon*, Kunming, China, Oct. 2002, vol. 2, pp. 1122–1126.

[56] H. Ren, Z. Q. Mi, H. S. Zhao, and Q. X. Yang, "Fault diagnosis for substation automation based on Petri nets and coding theory," in *Proc. IEEE Power Eng. Soc. General Meeting*, Denver, CO, Jun. 2004, vol. 1, pp. 1038–1042.

[57] S. M. Chen, J. S. Ke, and J. F. Chang, "Knowledge representation using fuzzy Petri nets," *IEEE Trans. Knowledge Data Eng.*, vol. 2, no. 3, pp. 311–319, Sep. 1990.

[58] T. V. Manoj, J. Leena, and R. B. Soney, "Knowledge representation using fuzzy petri nets c revisited," *IEEE Trans. Knowledge Data Eng.*, vol. 4, no. 10, pp. 666–667, Jul. 1998.

[59] M. M. Gao, M. C. Zhou, X. G. Huang, and Z. M. Wu, "Fuzzy reasoning Petri nets," *IEEE Trans. Syst., Man Cybern.*, vol. 33, no. 3, pt. A, pp. 314–324, May 2003.

[60] L. Xu and M. Kezunovic, "Implementing fuzzy reasoning Petri-nets for fault section estimation," *IEEE Trans. Power Delivery*, vol. 1, no. 1, Nov. 2006.

[61] S. M. Chen, J. S. Ke, and J. F. Chang, "Knowledge representation using fuzzy Petri nets," *IEEE Trans. KnowledgeData Eng.*, vol. 2, no. 3, pp. 311–319, Sep. 1990.

[62] J. S. Thorp, A. G. Phadke, S. H. Horowitz, and J. E. Beehler, "Limits to impedance relaying," *IEEE Trans. Power Apparatus Syst.*, vol. 98, no. 1, pp. 246–260, Jan./Feb. 1979.

[63] U.S.-Canada Power System Outage Task Force. (2004, Apr.). Final report on the August 14, 2003 blackout in the United States and Canada: Causes and recommendations. [Online]. Available: http://www.nerc.com

[64] T. Dalstein and B. Kulicke, "Neural network approach to fault type classification for high speed protective relaying," *IEEE Trans. Power Delivery*, vol. 10, no. 2, pp. 1002–1011, Apr. 1995.

[65] M. Kezunovic and I. Rikalo, "Detect and classify faults using neural nets," *IEEE Comput. Applicat. Power*, vol. 9, no. 4, pp. 42–47, Oct. 1996.

[66] Y. H. Song, Q. X. Xuan, and A. T. Johns, "Comparison studies of five neural network based fault classifiers for complex transmission lines," *Electric Power Systems Research*, vol. 43, no. 2, pp. 125–132, Nov. 1997.

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

[68] M. Kezunovic, I. Rikalo, and D. Sobajic, "Real-time and off-line transmission line fault classification using neural networks," *Eng. Intell. Syst.*, vol. 4, no. 1, pp. 57–63, Mar. 1996.

[69] D. V. Coury and D. C. Jorge, "Artificial neural network approach to distance protection of transmission lines," *IEEE Trans. Power Delivery*, vol. 13, no. 1, pp. 102–108, Jan. 1998.

[70] F. Zahra, B. Jeyasurya, and J. E. Quaicoe, "High-speed transmission line relaying using artificial neural networks," *Electr. Power Syst. Res.*, vol. 53, no. 3, pp. 173–179, Mar. 2000.

[71] H.-T. Yang, W.-Y. Chang, and C.-L. Huang, "Power system distributed on-line fault section estimation using decision tree based neural nets approach," *IEEE Trans. Power Delivery*, vol. 10, no. 1, pp. 540–546, Jan. 1995.

[72] A. J. Mazon, I. Zamora, J. F. Minambres, M. A. Zorrozua, J. J. Barandiaran, and K. Sagastabeitia, "A new approach to fault location in two-terminal transmission lines using artificial neural networks," *Electr. Power Syst. Res.*, vol. 56, no. 3, pp. 261–266, Dec. 2000.

[73] Z. Chen and J.-C. Maun, "Artificial neural network approach to single-ended fault locator for transmission lines," *IEEE Trans. Power Syst.*, vol. 15, no. 1, pp. 370–375, Feb. 2000.

[74] T. S. Sidhu, H. Singh, and M. S. Sachdev, "Design, implementation and testing of an artificial neural network based fault direction discriminator for protecting transmission lines," *IEEE Trans. Power Delivery*, vol. 10, no. 2, pp. 697–706, Apr. 1995.

[75] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Upper Saddle River, NJ: Prentice Hall Inc., 1999.

[76] N. Zhang, "Advanced fault diagnosis techniques and their role in preventing cascading blackouts," Ph.D. dissertation, Texas A&M University, College Station, TX, 2006.

[77] L. Priker and H. K. Hoidalen. ATPDraw Version 4.0 for Windows9X/NT/2000/XP. SINTEF Energy Research AS, Norway. [Online]. Available: http://www.eeug.org/files/secret/atpdraw

[78] The MathWorks, Matlab 7 getting started guide. [Online]. Available: http://www.mathworks.com/access/helpdesk/help/pdf_doc/matlab/getstart.pdf

[79] G. A. Carpenter and S. Grossberg, "ART2: Self-organization of stable category recognition codes for analog input patterns," *Appl. Opt.*, vol. 26, no. 23, pp. 4919–4930, Dec. 1987.

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

[81] M. Venkatasubramanian, M. Kezunovic, and V. Vittal. (Sep. 2008). Detection, Prevention and Mitigation of Cascading Events—Prototype Implementations. Power Systems Engineering Research Center (PSerc). Tempe, AZ. [Online]. Available: http://www.pserc.org

ABOUT THE AUTHOR

Mladen Kezunovic (Fellow, IEEE) received the Dipl. Ing. degree in electrical engineering from University of Sarajevo, in 1974 and the M.S. and Ph.D. degrees in electrical engineering from the University of Kansas, in 1977 and 1980, respectively.

Currently, he is the Eugene E. Webb Professor and Site Director of Power Engineering Research Center (PSerc), an NSF I/UCRC at Texas A&M University, College Station. He worked for Westinghouse Electric Corporation, Pittsburgh, PA, from 1979 to 1980, and the Energoinvest Company, in Europe from 1980 to 1986, and spent a sabbatical at EdF in Clamart from 1999 to 2000. He was also a Visiting Professor at Washington State University, Pullman, from 1986 to 1987 and The University of Hong Kong, during fall 2007. 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 and Registered Professional Engineer in Texas.