IEEE TUTORIAL ON AUTOMATED FAULT ANALYSIS

COURSE COORDINATOR

Mladen Kezunovic

Texas A&M University

Department of Electrical Engineering

College Station, TX 77843-3128-, USA

COURSE INSTRUCTORS

Mladen Kezunovic, Texas A&M University

Chen-Ching Liu, University of Washington

J.R. McDonald, University of Strathclyde, Scotland

Larry Smith, Alabama Power Company

July 2001

Contributors to the Tutorial Text

Section I: Requirements –- Utility Perspective
Larry Smith, Alabama Power, USA

Section II: Equipment Characteristics
Mladen Kezunovic, Texas A&M University, USA

Section III: Scope of the Analysis
Mladen Kezunovic, Texas A&M University, USA

Section IV: Implementation Using DFR Data
Mladen Kezunovic, Texas A&M University, USA

Section V: Implementation Using Status Data and Sequence of Events
Juhwan Jung and Chen-Ching Liu, University of Washington, Massimo Gallanti, ENEL, Italy

Section VI: Implementation Using SCADA Systems
J.R. McDonald and S.D.J. McArthur, University of Strathclyde, Scottland, U.K.

PREFACE

This tutorial has been organized to educate practicing engineers about the latest trends in automated analysis of faults. The history on this subject dates to the late eighties when first expert systems for automated analysis were introduced. These solutions were built using early expert system implemented with programming languages such as LISP, PROLOG, and OPS. The data acquired by the Supervisory Control and Data Acquisition (SCADA) Systems was utilized for the analysis. In the last fifteen years the field has advanced with new developments being pursued in two general directions. One direction was introduction of a variety of intelligent techniques, besides expert systems, such as neural nets, fuzzy logic, genetic algorithms, object oriented programming, etc. The other direction was the use of data from other substation recording equipment, besides SCADA Remote Terminal Units (RTUs), such as Digital Fault Recorder (DFRs), Sequence of Event Recorders (SOEs), Digital Relays (DR), etc. As a result, several practical solutions were developed and implemented at different utilities around the world. This tutorial is primarily aimed at discussing the requirements for the developments and characteristics of various systems being developed and implemented at this time.

In interest of the efficiency needed in presenting a one day tutorial, the authors opted not to include discussion of the fundamental issues related to the intelligent system techniques. This material can be found in some excellent IEEE Tutorials recently presented on the subjects of Expert Systems, Neural Nets and Fuzzy Logic. The attendees of this tutorial are strongly encouraged to use the published material from the mentioned tutorials if they have a need to get acquainted with the subject of the intelligent techniques. For the completeness of the subject matter, this tutorial will also venture into discussing some of the issues of the intelligent techniques and the implementation approaches. However, the main focus of this tutorial will remain a presentation of the different options and approaches in developing and implementing automated systems for fault analysis.

The first section of the report relates to the utility perspective on the subject. Mr. Larry Smith from Alabama Power has made a comprehensive summary of the requirements discussing the required data, identifying the potential users of the results, reviewing possible implementation approaches, and outlining the expected benefits coming from automating the analysis.

The next two sections are written by Mladen Kezunovic from Texas A&M University. These sections make a general introduction to the subjects of the equipment used for recording the data and the goals of the development. It is pointed out that the recording equipment characteristics may be quite different leading to different capability regarding the analysis implementation. It was also indicated that the final analysis goals may also be quite different depending on the end-users and hence the system implementations may have to be tailored to the specific type of the users.

The last three sections are discussing three different approaches taken in developing some recently implemented systems for automated analysis. The different approaches are based on the use of different recording equipment and the type of the users that were to utilize the systems were diverse.

Section IV describes several systems implemented using Digital Fault Recorders as the main source of data. Professor Kezunovic describes systems developed for Reliant Energy HL&P, TXU Electric, and WAPA. The basic idea incorporated in all the systems was developed by Texas A&M University, and later on was further enhanced by Texas A&M University and Test Laboratories International, Inc for the needs of the mentioned utilities.

The next section describes a development that was undertaken by University of Washington under guidance of Chen-Ching Liu. The original developments were based on the use of SCADA data base, but later on were enhanced to include data from the Sequence of Events Recorders as well. Professor Liu and co-authors discuss features of a system developed jointly by EPRI and ENEL and implemented in the utility in Italy.

The last sections introduces a comprehensive approach to fault analysis that utilizes data from a number of data recording devices located at various substations in the power system. The system has been developed by a group of researchers from the University of Strathclyde in Scotland, U.K., headed by Jim McDonald. In this section Professor McDonald and co-author outline the goals of the development and the implementation features of the new system.

At the end of each section, a detailed list of references on the subject matter treated in the section is enclosed to facilitate further understanding of the issues discussed.

Mladen Kezunovic — College Station
Tutorial Editor — July, 2001

TABLE OF CONTENTS

Section I

REQUIREMENTS: UTILITY PERSPECTIVE

Larry Smith
Alabama Power Company

Abstract. This section discusses a utility prospective of the fault analysis requirements. The needs of operations, maintenance and protection staff are presented. Data requirements and uses as well as the expected benefits and advantages are outlined at the end.

1. DATA REQUIREMENTS

In the past the utility engineer often struggled with the fact that little data was available when attempting to analyze problems within power systems. They also did not have enough information to predict or ascertain the level of maintenance needed or when it would be required for the major equipment located within their substations. As new and higher levels of technologies have made their way into the utility environment, these same engineers are now suffering from data overload. They have more data than can be processed and assimilated in the time available. Thus important knowledge concerning the status of substation equipment is just lying stagnant and not being used to the betterment of either the personnel or the equipment.

This "data overload" not only has impact on each piece of equipment, or substation, but also at the system level. The data might be coming from sensors in breakers or transformers or some could even be available in other monitors already located in the substation (i.e. fault recorders, event recorders, RTUs and microprocessor based relays).

Where and when this data can and should be automatically converted to information is also a question. Then once the information is available how is it turned into knowledge. For knowledge is power, and can only be used for the betterment of the system once it (knowledge) is available. In the past this process was performed within the brain of someone. Today the challenge is to automatically convert data to knowledge, which frees the manpower to implement corrective or preventive action.

The massive quantities of data, the diverse points of origin, and the vast array of implementations make this a very complex area of discussion. How do you correctly interpret the data? How is the right decision reached? To whom should the decision be supplied? In what time frame does the critical path lie? How many different people need to be supplied with what different pieces of information?

These many different questions need to be discussed and decisions reached so each utility can provide its' own personnel with the information and knowledge needed to adequately operate and maintain their electric systems.

Sometimes the data available and the data required may be two very different things. The following sections discuss what information is required by the different areas of responsibility within a utility. Then the final section discusses the data needed to supply the required information. A basic understanding of who will be using the data and how they will be using it, is required in order to understand the type of information required and the necessary accuracy.

2. USERS OF INFORMATION

In the utility environment, the users of the information extracted from different recording devices and systems are:

- Operations
- Maintenance
- Protection

Operations. The information requirements for operating personnel from an automated analysis package are somewhat limited. Operating personnel are charged with returning to service as much of the electric system as practical within the shortest time possible. Therefore, the items of interest to them are two-fold.

First, where or what is the problem? Did the line re-close and stay in? These are the first questions that should be answered. The information necessary should be presented as quickly as possible.

The second, item of interest is: What equipment operated? Did everything work correctly? If so can it be returned to service? If not, what has to be isolated?

If automated fault analysis is to be of any significant assistance to operating personnel the completed analysis should be available for use within 5 (five) minutes from the conclusion of the event. (i.e. The results of the analysis should be presented to the operators.)

As previously discussed, fault location and operational correctness will allow the operating personnel to make their first decision concerning return to service. If a line operates one time and re-closes successfully, then all that is required is to forward the fault location to line personnel for later inspection. If the re-closing was not successful then the operating personnel should use the fault location as an indication of how to isolate the problem.

If the fault location indicates line problems, the operator should then use that location to return to service as much of the line as possible. Thereby isolating the problem area. If the fault location indicates that some piece of major equipment is involved (i.e. circuit breakers, transformers, switches, etc.) then the operator proceeds to isolate that equipment and return to service other outages equipment.

Maintenance. The information requirements from a maintenance standpoint are entirely different. Maintenance personnel are charged with repairing and returning outages equipment to service. Therefore, they require information concerning what is damaged or operating outside of normal parameters. Then maintenance personnel can then be dispatched to correct any problems (i.e. downed conductors, failed lightning arrestors, breakers, or transformers. Normal time requirements for maintenance personnel depends on the way a utility operates and the type of equipment, and its' criticality in serving load. This usually is in the region of two hours or less, to allow notification of the correct personnel to solve the problem.

Protection. The most difficult job falls to the protection engineer, who is charged with the final assessment of the correctness of any response to a given fault condition. Except in extremely rare cases of catastrophic failure, normally the protection engineer is given adequate time to collect all data necessary for a complete evaluation of any event (i.e. fault clearing or equipment failure, etc).

Among the many questions that must be answered are: "Did the right thing, respond the right way?" "Did the wrong thing respond the wrong way?" "Did the right thing, respond the wrong way?" These are all very simple questions with complex answers. The "thing" in the above questions might be a relay, relaying system, circuit breaker, or switch. The conjecture lies in the hands of the protection engineer to make the assumptions necessary to correct the system problem that caused an erroneous system response.

This area of analysis can not be fully automated, because all system mis-operations can not be defined. The raw data is very important here and as much of it as possible allows for a timelier resolution of the given condition.

3. ANALYSIS IMPLEMENTATIONS

Data availability is the prime mover in the type of automated analysis that is possible. Implementations will be discussed involving data from Sequence of Events Recorders, SCADA Systems, and Digital Fault Recorders. Each provides data that can solve problems of differing types.

The data that is required to provide as complete an analysis as possible differs with the answers required. The following section will attempt to address the data only and thus allow the different implementations to acquire and use the data necessary to provide the answers expected.

Basically, only fault clearing analysis will be addressed here. If conditions such as Power Swings, Undervoltage, Fault Location Information, Underfrequency, Dynamic Conditions on Generators, or Data for Relay Testing are to be automatically analyzed other data may be required.

The commonality of time also plays a role in the ability to analyze complex faults using data from different pieces of equipment. The basic assumption here is that all data for the Sequence of Events solution and the Digital Fault Recorder solution is synchronized to within one millisecond. For the SCADA solution there presently exists two different types of time data. Older SCADA systems only report time to the second, with some time tagging after receipt at the SCADA master. Some newer systems time tag to the millisecond at the remote terminal unit (RTU) and should be considered as a Sequence of Events implementation for that reason.

The operational data needed to analyze the performance of the protective relay systems, and interrupting mechanisms during faults on the electrical system should be available from the following systems:

- Protective relays
- Circuit breakers
- Other major substation equipment
- Automatic re-closing

To provide data necessary for automated analysis, the following quantities should be available:

A. **Bus phase voltages:** Depending upon the station configuration, bus and/or line voltages (phase-to-ground) should be monitored. A minimum of three phases per voltage class should be provided.

B. **Bus residual voltage (derived from phase quantities):** This is necessary for the analysis of potential polarized relaying schemes.

C. **Line phase voltages:** These are necessary for the analysis of re-closing schemes and relay operations.

D. **Line phase currents:** Currents from all three phases on each line monitored.

E. **Line residual current (derived from phase quantities):** This quantity is an absolute must to analyze any type of disturbance on a power system. Most faults can be analyzed by using this quantity and the three phase to ground voltages.

F. **Pilot channel data:** The data should include both transmitter and receiver status on all power line carrier, fiber optic, audio tone, pilot wire, and microwave circuits.

G. **Breaker, station tripping, and blocking status data:** Auxiliary contacts show when status points changed state. Breaker position, lockout relays position (differential and breaker failure) and pilot channel blocking all aid in the assessment of both correct and incorrect operations.

H. **Control contact performance:** Trip and close contact information is a must. Trip and close initiation helps to better define what happened and why (i.e. which relays initiated the trip or close).

Table 1-I. DATA AVAILABILITY BY DEVICE TYPE

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
DFR	x	x	x	x	x	x	x	x	x	-	x	x	x	x	x	x
SER	-	-	-	-	-	x	x	x	x	-	x	+	-	-	-	-
SCADA	x	x	x	x	x	-	+	+	x	+	+	-	-	-	-	-
RELAYS	x	x	x	x	x	+	+	x	-	x	+	x	x	x	x	x

X **Data available**
- **Data not available**
+ **Partial Data may be available in some form**

I. **Alarm contacts:** If possible an alarm contacts such as relay failure, breaker air and/or gas pressure and transformer temperature and cooling should be included whenever possible, in order to allow a more complete analysis.

J. **Relay target data:** Presently, outputs with this information are not available from all relaying systems, however, any system that is to perform automated power system analysis should include target data. Microprocessor relays and relaying systems provide ASCII and contact output of target data. It would be extremely beneficial if these targets can be included in the analysis.

K. **Time code information:** Time should be accurate to a minimum of one millisecond and it should be possible to determine the time for each data sample. This accuracy is necessary when comparing fault data from different stations. If data from different sources are to be merged to make calculations such as phase angle between sources it is necessary that time information have an accuracy of 50 microseconds (1 degree of 60 Hz) or better.

L. Fault duration: Used in analyzing relay and breaker performance.

L. **Clearing time (all phases):** Used in determining pole openings of breakers and fault clearings.

N. **Magnitude of the fault current:** Magnitude use for fault locations and determining which relays should have operated.

O. **Type of fault (single phase, multiphase, evolving):** The analysis of faults depends on knowing how faults evolved.

P. **Phases involved in the fault:** Useful in fault locating and relay system analysis

SCADA analog quantities will not be instantaneous quantities and will not provide data during the fault.

If possible a record of pre-disturbance data would be of benefit to establish pre-fault levels and relay memory quantities.

4. BENEFITS TO THE INDUSTRY

As stated earlier presently the mountains of data that is created with any event of the electric power system is tremendous. If this data can be automatically analyzed at least or ascertain the correctness of simple operations then personnel will have more time to spend on the more complex problems. These systems might also be of benefit in bringing problem areas to the attention of the proper personnel. An automated analysis system reaps benefits in the following areas:

- More consistent analysis
- Reduced time for analysis
- Reduced cost of serving different groups

More consistent analysis. Once an analysis system is in place it should insure a more consistent and stable classification of events. With several different people within an organization analyzing events it is very likely that different standards and suppositions are being used. Thereby creating very different solution and false solutions.

Reduced time for analysis. If an automatic analysis can be performed as quickly as possible during or following an event, then operating and maintenance personnel can respond very quickly to problems. The major advantages are:

1. Reduce outage time for customers.
2. Quicker response to problems that might develop further into catastrophic events.
3. Increased reliability by reducing problem areas prior to erroneous operations.

Reduced cost of serving different groups. Presently, analysis personnel may be required to produce reports required in different areas of the organization. With a completely automatic system these basic reports could be delivered to the requesting areas via electronic transmission and the analysis personnel only having to report on the more complex problems. So the reduced cost is also a productive improvement.

5. CONCLUSION

The information presented here has been a combination of ideas and tools developed after twenty-five years of using this data. It also represents input provided by various departments within a utility. Most importantly it is an example of what needs to happen to give the correct personnel the information needed in a timely fashion. This process may be shown as the following evolution.

DATA	should provide
INFORMATION	that yeilds
KNOWLEDGE	that invokes
ACTION	which in turn yeilds more data

As the work continues the one thing that has to remain in the forefront is the fact that we are attempting to determine the correct action that will benefit the customer in the best way. Timely return to service, increased reliability, cleaner power, etc.

Section II

EQUIPMENT CHARACTERISTICS

Mladen Kezunovic
Texas A&M University
College Station, TX 77843-3128
USA

Abstract. This section provides an overview of the basic issues associated with analysis of recordings in substations: recording equipment, data obtained, and uses of data. As a conclusion, a discussion related to the hierarchical levels of data processing and possible implementation architectures will also be given.

1. TYPICAL RECORDING EQUIPMENT

In the substations, the recording equipment can be quite versatile [1]. This may depend on many factors including the history of the substation construction and upgrades, utility operating practices, strategic importance of the substation, etc. In any case, the following are different recording equipment types that are typically used in modern substations:

- Digital protective relays (DPRs)
- Digital fault recorders (DFRs)
- Sequence of event recorders (SERs)
- Remote terminal units (RTUs) of a SCADA system
- Intelligent Electronic Devices (IEDs) used for variety of monitoring and control applications
- Fault locators (FLs) developed for stand alone high accuracy fault locating

Digital protective relays (DPRs). The modern DPRs today represent a complex recording and measurement instrument equipped with a decision making control logic and multitude of monitoring functions [2]. It is important to note that modern DPRs also have a variety of settings (both user selectable and internal) as well as a number of internally computed measurement and logic signals that may be accessed by the user. The main obstacle in deriving a generalized model of such a relay (which may be essential for the analysis) is the fact that the relay designs from various vendors may be quite different and levels of data access provided for the outside use may also be significantly different. The analysis is generally based on some type of a hypothesis about the relay operation, and the hypothesis is directly tied to the availability of a model as well as the measured data to verify the expected vs. the actual relay behavior. Last, but not least, due to a relatively low sampling rate of some of the earlier DPRs their waveform recording function provides only a limited frequency representation of the waveforms. This may impair the ability to perform a detailed analysis based on the recordings.

Digital fault recorders (DFRs). The modern DFRs are highly accurate recording instruments providing sampled waveform and contact data using relatively high sampling rate (typically above 5KHz) [3]. Their use in the analysis is quite appropriate since they provide recordings of the waveforms that were also "seen" by the DPRs. However, various DFRs provide different triggering mechanisms, and the performance and sensitivity of the triggers may affect the ability to capture relevant waveforms. Some of the newer DFR designs allow the user to program a customized triggering mechanism, which in turn can assist the analysis process. It should be noted that DFRs are generally pretty expensive when a unit cost per channel is considered. Due to a large number of input channels typically being required in a substation, an attempt may be made by the field personnel to connect only the crucial monitoring signals for the recording. In that case not all the channels of interest for the analysis may be available (recorded). Further potential obstacles are associated with the DFR data formats that may be proprietary not allowing implementation of an "open" data recording system which can easily be interfaced with the analysis function.

Sequence of Event Recorders (SERs). The modern SERs are complex recording instruments implemented today most likely using programmable logic controllers (PLCs) and analogue waveform data acquisition subsystems [3]. The SERs are capable of monitoring changes in the switching equipment status with high precision due to a high data sampling rate. Combined with measurements of analog signals, the SERs can record the status change for variety of controllers including the ones that are based on analog set points. Most of the SERs can also be set to provide control function through a number of control outputs. For the analysis purposes, existence of SERs in a substation is very important. Unfortunately, most of the utilities will have SERs possibly only in larger substations due to an excessive cost. In any case, the SERs also present a potential obstacle of not being designed as an "open" system, which in turn may reduce an ability to interface the data recorded by those systems with the data from other sources used for the analysis. If available and accessible through standard communication interfaces, SERs may be very important recording infrastructure to be used for the analysis.

Remote terminal Units (RTUs) of a SCADA system. The modern RTU can be a very sophisticated recording instrument that may have a recording performance of a DFR, and at the same time may be producing a variety of pre-calculated quantities [4]. In addition, some advanced RTUs will provide an extensive SER and some limited DPR functions. Due to the fact that RTUs are a part of the Supervisory Control and Data Acquisition (SCADA) system, the data is readily available for the analysis at the centralized location through a SCADA database. However, the "open" system design provision remains an issue with the RTUs as well since they are primarily designed to interface to the EMS SCADA data base using mostly customized communication protocols and database formats. Another potential problem is related to a limited opportunity for the user to access the recorded data locally before it is sent to a centralized location. In the analysis hierarchy it may be desirable that some locally recorded data is available in a substation for a local analysis, and this may be difficult to implement due to a lack of local substation user interfaces in most of the RTU designs.

Intelligent electronic devices (IEDs) used for variety of monitoring and control applications. The modern IEDs are available today for variety of applications ranging from simple stand alone controllers and dedicated data recording systems to pretty complex integrated devices for monitoring, control and protection of the entire substation bay [4]. The main issue with IEDs, when used for the analysis, is the "open" communication architecture and data recording performance. Since the IEDs are not standardized even regarding the functions they perform, it may be very hard to find detailed enough description that will allow generic models to be developed and used for the analysis. In any case, not withstanding the limitations, the IEDs are indeed a good addition to the data recording infrastructure needed for a comprehensive analysis to be performed related to substation equipment operation.

Fault locators (FLs) developed for stand alone high accuracy fault locating. The modern stand alone FLs are designed to provide very accurate fault location, but for that purpose they may have to have a fairly advanced built-in data acquisition system [5]. Because of this, the cost of the stand alone FLs is almost prohibitively high, and only a few companies are using them extensively. More typically, these instruments will be used occasionally and only on the most critical transmission lines. Further discouragement for the use of stand alone FLs comes from the fact that the fault location function is provided today almost free of charge on most DFRs, DPRs and even RTUs. Hence there may be very little motivation for the additional investment for possibly not too significant increase in the accuracy. In any case, should these instruments be available, the analysis can be significantly enhanced as long as the data from the FLs is readily available for the analysis use.

2. DIFFERENCES IN THE RECORDING EQUIPMENT PERFORMANCE

Possibly the least understood issue related to the analysis is the recording equipment performance. As much as the basic recording function of all the equipment may be the same in that the same signals may be recorded, the type of the data recorded and performance of the recording function may be quite different. The following are some typical recording approaches that are found in different equipment, which in turn make the difference in the performance of the data recording function:

- Synchronized data sampling vs. scanning
- Continuous recording (after triggering) vs. reporting by exception
- Local synchronizing of sampling vs. system-wide synchronizing
- Low precision (10 bit) vs. high precision (16 bit) A/D conversion
- Low sampling rate (16 s/c) vs. high sampling rate (64 s/c)
- Recording of pre-calculated values vs. recording of samples
- Pre-filtering of data (beyond just the antialiasing filtering needs) vs. only the antialiasing filtering

Synchronized data sampling vs. scanning. It is well known that some recording instruments perform synchronized data sampling on all the channels connected to the instrument. In this case a sample and hold (S/H) circuit is provided on each input channel, and all of them are strobe at the same time via a common sampling clock signal. The basic techniques for deriving and synchronizing the sampling clock may be implemented by using either an independent external clock or utilizing a phase lock loop tuned to the frequency of the analog signal being sampled. In any case, the signal samples obtained this way can be used to recover not only the basic properties of the signal, but also to establish a phase difference between the signals presented at different inputs. In the analysis, knowing the phase difference may be critical. This is the case when the fault detection and fault type classification need to be performed in a three-phase system. On the other hand, some instruments use the data scanning techniques where the samples are taken from each channel with one and the same S/H circuit. This is typically implemented using a multiplexer that switches the sample and hold circuit among various channels. The time of switching is calculated to accommodate the time required for A/D conversion so that each next channel sample is taken only after the previous channel sample has been converted. Using this technique the samples are taken at different time points on each of the channels hence prohibiting an easy way of establishing the actual phase difference among signals connected to various channels. DPRs, DFRs, FLs, some SERs and some IEDs typically perform synchronized data sampling, while RTUs, and some other types of SERs and IEDs would typically use data scanning techniques.

Continuous recording (after triggering) vs. reporting by exception. Some of the recording instruments are set up to perform continuous recording where a certain length of a data buffer is continuously upgraded with new samples while the old ones are discarded. This technique is called "circular buffer" recording. After the instrument is triggered by an event or disturbance, the recording is continued beyond the circular buffer length, while the circular buffer data is preserved. The length of the recording may be as long as there is the memory available, and the recording appears to be "continuous" after tripping, beyond the "circular" buffer length. On the other hand, the recording may also be initiated by a command or a trigger asserted by an operator. In this case there may not be a pre-history available, and the history starts with the trigger initiation of the recording. Another similar approach is taken by setting a threshold for a measured value (RMS for example), and each time the threshold is exceeded, the value is recorded and reported. The analysis may be affected by the history available for the recorded data, and different instruments will provide different performance in this respect due to the above mentioned recording techniques utilized. Typical example of the use of the pre-event history is the calculation of the load flow data prior to a given fault. This data may be critical in understanding the line loading and substation switching status, which in turn may sometimes be crucial in trying to locate the fault by fitting the recorded data to the modeling and simulation data. DFRs typically maintain several cycles of pre-fault data, while DPRs do the same but the length of the pre-fault data may vary depending on the fault detection technique implemented. Again, the ability to perform the analysis may be dependent on the length and type of the data history recorded.

Local synchronizing of sampling vs. system-wide synchronizing. The data sampling techniques require a synchronizing clock signal to be available and its synchronization to be maintained over a period of time. The two typical implementation approaches are: local synchronization and system-wide synchronization. The local approach assumes that there is a local clock which is fairly stable, and hopefully re-synchronized from a more accurate centralized clock. In some instances the re-synchronization is done manually, and the drifts in the local clock may be significant. On the other hand, the synchronization may be done using "continuous" synchronizing to a system-wide accurate clock. Such an example is the case of the use of the Global Positioning System (GPS) of satellites. In this case each of the recording devices needs to be equipped with a GPS receiver, and synchronization accuracy achieved is within several microseconds. The type of sampling synchronization and its importance in the analysis applications is directly tied to an attempt to establish either a phasor or data sample correlation at two adjacent substations or within a given substation between different recording instruments. It should be understood that the phasors or data samples may also be synchronized using software techniques, but the process may be more complex and may require additional processing time. The analysis may require that certain phasor or sample values from different recording devices are used (computation of accurate fault location, for example), and in that case the synchronization and associated accuracy need to be maintained and understood. Very few instruments in the use today provide GPS receiver for synchronization, but in the future this feature may become standard.

Low precision (10 bit) vs. high precision (16 bit) A/D conversion. Obviously, this discussion is related to the recording of analog waveforms. The "vertical" A/D converter resolution vs. "horizontal" resolution is the issue which affects accuracy of the signal measurement vs. signal representation respectively. The horizontal resolution is determined by the sampling rate and will be discussed in the next paragraph. The vertical resolution is associated with the dynamic range of the signal. The larger the dynamic range, the higher the need for a more precise A/D converter. For example, to capture the dynamic behavior of the current signal before and after the fault, the data conversion using a 16 bit A/D converter may very well be needed. On the other hand, the dynamic change in the voltage signal may be measured rather accurately using a 12 bit A/D converter due to a smaller dynamic range. If the analysis requires very accurate representation of a wide dynamic range of the signals, a high precision A/D converter may have to be used. It is worth mentioning that only the latest products for substation recording may be equipped with a 16 bit A/D (which is the desirable accuracy in most application cases), while most of the older products have either a 10 bit or 12 bit A/D converters. Another point to be understood is the signal scaling technique used in some recording devices to adjust signal dynamic range before the A/D conversion is applied. For the analysis purposes the scaling technique may have to be known in advance to be able to interpret the recorded waveforms and perform the analysis correctly.

Low sampling rate (16s/c) vs. high sampling rate (64 s/c). The issue of the sampling rate used in a recording instrument may be quite important for at least two reasons: the antialiasing filter selection and accuracy of signal representation. The antialiasing filter selection is associated with the requirements that the sampling rate be at least twice the highest frequency to be represented in the sampled signal. This requirement comes from the well-known sampling theorem, and is selected based on the application at hand. Since some of the recording instruments have other applications implemented on the same device as well (DPRs are a good example), it becomes very important to understand the constraints placed by the main application when an auxiliary application is being defined based on the data coming from the same instrument. This may particularly be the case when the data from a DPR is used for the analysis purposes. On the other hand, the accuracy of signal representation is also dependent on the selection of the sampling rate. In general, the higher the sampling rate, the better the representation. As well known, the higher sampling rate may not contribute to a better measurement accuracy unless the selection of the A/D converter meets the dynamic range requirement. In addition, the antialiasing requirement may be satisfied with a lower sampling rate, but a better signal representation may be achieved by increasing the sampling rate. If one is observing given measurements that represent pre-calculated quantities, one should note that some of the calculated quantities may be optimized if a given sampling rate is selected (some techniques for phasor measurements, for example). In some other measurement approaches there is no clear guidance for the optimal selection of the sampling rate (measurements based on the differential equation solution). The selection of the sampling rate is a multifaceted issues and needs clear understanding when the analysis applications are considered.

Recording of pre-calculated values vs. recording of samples. The analysis may require that some values already pre-calculated by the recording device be reconstructed for the analysis purposes using samples. This poses an interesting question as to what is available from a given device: samples, pre-calculated values, or both. The instruments that are designed to perform only the recording function (DFRs for example), would provide the samples. Dedicated instruments such as PQ meters or RTUs may provide only pre-calculated values. The DPRs may be able to provide both. In any case, it may be important to understand the exact signal processing that has taken place in the recording device when the data available for the analysis was acquired. This information may not necessarily be available to the end user, and the analysis may have to be simplified if such information is not available. On the other hand, over simplification of the analysis may not produce the desirable end result of a reasonable accurate and trustworthy analysis.

Pre-filtering of data (beyond just the antialiasing filtering needs) vs. only the antialising filtering. This issue is leading one further into the details of the signal processing undertaken by any recording device. If the details are known, the original input signals may be reconstructed more faithfully, and this in turn may lead to a better analysis. The pre-filtering process in some of the recording devices (DPRs in particular) may be quite involved and data difficult to reconstruct unless full design information for the data acquisition subsystem is provided. Typical example is the use of the different techniques for filtering of the DC offset and " undesirable" harmonics. Some of the DPRs have the filtering part separated from the part that reconstructs a measured quantity, while some DPRs have both function implemented in one and the same digital filter. Being able to distinguish the differences may be an important part of the analysis, but again the level of detail in the analysis may be the driving factor for wanting (or not wanting) to know the design details of the recording equipment. Of course, the recording equipment that uses only the antialiasing filters of a simple design may be doing the least of pre-processing and hence reconstructing the signal may be the simplest.

3. DIFFERENCES IN THE PRACTICE OF USING THE EQUIPMENT

The existence of a given recording infrastructure may not uniquely define the analysis logic, implementation approach, and the end user interfacing. One of the reason for this is the different practice of using the equipment found at different utilities. To further clarify this statement, the following aspects of the practice are discussed in more details:

- The analysis logic
- Implementation infrastructure
- The end user interfacing

The analysis logic. To say that an analysis of the recordings in substations is to be performed does not give enough details to understand what would be the data analyzed, with what purpose, and how the reports will be organized. One approach regarding the data selection is to identify if the analog waveforms, contacts, or both will be used. Immediately after that, a clarification related to the exact type of signals and their relationship to the equipment and/or power system behavior needs to be established. Subsequently, a detailed description of the analysis hypothesis that is to be evaluated needs to be given. Finally, the signal pre-processing, calculated quantities and comparison thresholds need to be specified or derived. The simplest approach is to identify the entire analysis logic based on a given recording instrument infrastructure. This in turn may simplify the definition of all the mentioned components of the logic since some personnel will have extensive experience with a manual analysis of the data coming from this type of the recording system. The process of defining the logic for an automated analysis will be somewhat simplified since one would need to concentrate primarily on formalizing the procedures already performed by the operators. However, if the analysis is to be done based on some new data, or a new recording infrastructure consisting of several traditional instruments, the process of defining the analysis logic may be much more difficult and even controversial. The purpose of the analysis may become a complex issue if the analysis is aimed at some new practices yet to be established, or at the use of some additional equipment and/or power system performance criteria. A discussion of the purpose would have to be carried out early enough in the design stage, and it should be made almost transparent to the details of how the logic may be implemented to avoid overlapping the purpose and implementation issues. This discussion leads to the final question of how the analysis reports should be organized. Depending on the experience of the design team and the support received from the prospective users, the best approach may be to try to define the final outcome (format of the reports) first, and then to proceed backward towards definition of the purpose and eventually to the data selection and the recording system implementation details.

Implementation infrastructure. Once the basic concept of the analysis logic is resolved, and detailed functional specifications are developed accordingly, the implementation approach needs to be defined and related hardware and software infrastructure put in place. As with any good design, the implementation approach needs to reflect the ability for the user to make changes, additions and future upgrades. In the modern software architectures available as a part of the standard operating systems and programming languages, it is essential to recognize the software tools that can allow the above mention design criteria to be met. One component of the consideration is the definition of the hierarchical processing and data storing structure that will correspond to the physical power system hierarchy. For the analysis, and corresponding implementation, it is crucial to decide if the analysis is to be performed in the substation, close to the source of data, or if it is going to be performed at a centralized place. Of course, a combination of the above is a possible option as well. The location of the analysis introduces very important implementation issues such as the data base organization, data communication and processing requirements. Once this decision is made, one may consider using standard software tools and architectures that may be utilized to configure the application. Typical examples of this approach is the use of the Internet and web server/browser configuration tools that can allow for application "plug ins" to be constructed to cover the specific applications. In addition, extensive database tools for storage, upgrades, protection and management of the data may be utilized to assure the required data access. Finally, a concept of the client/server architecture may be applied to a great extent to define the analysis processing applications, and the user report viewing applications. The use of the Microsoft NT operating system tools combined with a powerful MFC programming tools of C++ may be utilized to implement the client server architecture that is quite flexible and modular.

The end user interfacing. As much as the issue of the end user may have been identified and discussed as the part of the analysis logic definition, it still remains to be decided as how flexible the end user distribution of the reports should be. Today's advances in the computer communication area provide a variety of options for the end user interfacing. For example, all the interested users connected to the company LAN, Web, cell phone, fax, and pager systems may be notified about different aspects of the analysis. In some cases the users may only be prompted that a report exists, while in some other instances the users may be provided with a condensed and/ or expanded version of the report. Of course, all of these features need to be coordinated with the basic company operating polices and practices. The mentioned technologies allow the user to be reached almost at any time and at any place, and notified about the results of the analysis. A decision needs to be made about a reasonable level of detail and complexity of the report prepared for company wide distribution. Least to say, the cost of providing the interfacing needs to be accessed as well.

4. CONCLUSION

As a result of the discussion in this section one can note a variety of issues that need to be addressed when the analysis of recordings in substations is to be performed and automated. One obviously needs to start with the recording equipment available in substations. The decision needs to be made regarding the choice and type of the equipment to be used. In particular, if some new equipment is to be added in the future, this needs to be known in advance as well. The next step is to reach a detailed understanding of the equipment performance and recording capabilities. For the new equipment, the users need to become intimately familiar with the equipment performance. This issue becomes very critical if the analysis is to be performed using data from different recording systems. Mixing and matching the data is not a trivial issue, and a consistent way of handling, processing and storing the data from different sources has to be defined. After that, a clear understanding of the practice of the equipment use needs to be established. If the equipment has been used in the past, the operators of the equipment need to be involved in the discussion of the functional specifications for the analysis due to their extensive experience regarding manual analysis performed in the past.

5. ACKNOWLEDGEMENT

The author wishes to acknowledge the financial support from Electricite de France (EdF) that was provided during the author's sabbatical stay at EdF and made the work on this section possible. EdF is also acknowledged for permitting the author to use parts of the text otherwise developed for some internal EdF reports.

6. REFERENCES

[1] A. Newbold, P. Fauquembergue, M. Kezunovic, A. Sanz, M.V. Gonzales Sabato, " Use of Intelligent Systems within Substations," *Electra*, No. 181, pp. 93-111, December 1998.

[2] M. S. Sachdev (Co-ordinator), "Advancements in Microprocessor-Based Protections and Communications," Tutorial Text, *IEEE Publication No. 97-120-0*, Operations Center, Piscataway, NJ, 1997.

[3] IEEE Special Publication, "Application of Fault and Disturbance Recording Devices for Protective System Analysis," *IEEE Publication No. 87TH-0195-8 PWR*, Operations Center, Piscataway, NJ, 1987.

[4] Proceedings of the Symposium, "Integrated Control and Communication Systems," *CIGRE*, Helsinki, Finland, 1995.

[5] Proceedings of the CIGRE Colloquium, Study Committee 34, Preferential Subject 2: Fault Location and System Restoration, Florence, Italy, October 1999

Section III

SCOPE OF THE ANALYSIS

Mladen Kezunovic
Texas A&M University
College Station, TX 77843-3128
USA

Abstract. This section sets the stage for definition of the scope of the analysis considered in this tutorial. The analysis is related to the operation of protective relays and related equipment such as circuit breakers, communication channels, etc. To further define the scope, the following subjects related to the analysis will be discussed: goals, approach, levels of complexity, and implementation techniques. The conclusion will outline the application focus within the scope.

1. ANALYSIS GOALS

The analysis goals discussed are driven by the latest trends of making the monitoring of power systems more cost effective and focused. This in turn should make the service of delivering power to the customers more reliable and power system operation performance more competitive. In order to meet the mentioned objectives, the analysis goals should be as follows:

- Efficient utilization of existing data
- Minimum interruption time of power delivery
- Increased automation
- Maximum equipment performance

Efficient utilization of existing data. One of the major goals in the analysis is to try to utilize the existing data to a maximum. This requirement may translate into two important criteria: no additional equipment should be installed unless absolutely necessary; data from one type of the equipment should be utilized for multiple purpose as much as possible. If no additional equipment is to be installed, then a natural question arises: what is the equipment that should be selected as the most appropriate, and further more, should data from several different types of equipment be combined and used. In making a decision, one should be fully aware of the performance limitations of different types of recording equipment as discussed in Section 2 of this tutorial. One should also observe the potential difficulties associated with combining data from different recording systems as also discussed in Section 2. The preference of some of the utilities seem to be the use of the digital protective relays (DPRs) as the recording system. This preference may be associated with the fact that DPRs are becoming the most sophisticated recording systems in substations, and yet will be wide spread in the near future. The cost of adding other types of recording systems may not be justifiable if it can be proven that the use of DPRs as the source of data can meet the goals. As per the second criteria related to the most efficient utilization of data, it also translates into a need for the analysis to provide the results that may be used by a number of different staff including operations, maintenance and protection engineers.

Minimum interruption time of power delivery. The increased competition in the utility industry has been quite often interpreted as the ability to provide an improved service to the customers without a significant increase in the cost of the energy being delivered. One very simple criteria for the improved service is the ability to maintain the service with minimum interruption of the power delivery. One of the major goals of the analysis is to provide enough information to all the utility staff to be able to understand the reasons for the interruption better, and provide as quick and as focused as possible an action to restore the power delivery in the cases the analysis shows that the service was interrupted. If the service was not interrupted, and it should have been (as is the case when a relay fails to operate when a permanent fault occurs), the analysis should indicate that a quick manual action is needed to isolate the faulted part as soon as possible. Finally, the analysis should also provide enough understanding of the status of the equipment so that a preventive set of measures can be put in place to reduce the likelihood of the damage and service interruption to occur due to deterioration of the equipment “health”. The analysis should be very much focused towards addressing all of the mentioned goals and should not be developed in a vacuum where the goals are not clearly specified and quantified. This is the best way of making sure that the final tools and methodologies are going to be understood and accepted by the staff, which is not a trivial goal in itself to reach.

Increased automation. Almost all of the modern recording systems provide some level of automation in collecting and storing the recorded data. However, the analysis process is almost as a rule left to the operator to perform through manual operation associated with selecting and viewing the files. The analysis as discussed throughout the most of this tutorial is to be performed automatically with a minimum interaction from the utility staff. This requirement comes from at least two criteria. One is associated with the system operators and aims at the analysis performed as quickly as possible so that an action of restoring the system can be taken immediately if needed. The other one is associated with protection engineers and maintenance staff and aims at minimizing the time spent on the routine manual tasks so that the staff can spend additional time on analysis of complex cases, if needed. As a result of the mentioned requirement for increased automation, the recording system infrastructure has to be enhanced to provide for automated collection, storage, and processing of the recorded data as well as for automated generation, storage and distribution of the analysis reports. The reason for mentioning the need for the enhancement is the fact that almost none of the existing recording systems have the ability incorporated at this time. Some of the enhancements needed are quite simple and may come as a by-product of the software technology and tools used to implement the analysis solution. As a side comment, one should pay particular attention to the software environment used for the implementation of the analysis application since the selection of the commercial software tools can tremendously affect the efficiency and flexibility of the automation.

Maximum equipment performance. The equipment operating performance is of concern to the utilities since poor performance may lead to undesirable operating conditions and even failures. An example of the poor performance is a slow operation of circuit breakers. If too slow, the operation may affect the system stability in the cases where the operation is initiated by the protective relays with a goal of clearing the fault as soon as possible. An example of the failed equipment is a stuck breaker that can not perform switching operation when called upon to do it. To maintain desirable level of the equipment operating performance, the equipment operation needs to be monitored, analyzed and reported on. The analysis may be done for at least two purposes: to enhance the maintenance procedures, and to evaluate the operating practices. In both cases, the analysis details and levels are very much dependent on the type and number of the signals recorded. In some instances (circuit breaker monitoring for example), if only the 52a and 52b contact status is recorded, this may lead to the ability to monitor and evaluate the switching action only [1]. In this case, for a more elaborate circuit breaker performance monitoring and evaluation, the current waveforms together with the contact status as well as current and voltage waveforms from the DC control circuit may have to be recorded.

2. OBJECT ORIENTED ANALYSIS APPROACH

In performing the analysis, it becomes pretty important to decide on the focus by selecting appropriate objects to be analyzed. An object may be either a part of the power system, the whole power system, a particular type of event, peace of the equipment, etc. The point is that an object needs to be well defined from the stand point of the analysis to be performed. As such, an object may be virtual, that is it may correspond to a combination of the physical equipment and analysis tasks. The following are some important properties of the analysis objects:

- Inputs and outputs
- Purpose of the analysis
- Hypothesis assumptions
- Time response
- Intended uses and users

Inputs and outputs. It is obvious that each analysis object has to have some inputs and outputs defined. The inputs have to be clearly defined including the sources of data, type of data, means of collecting and storing data, formats for data presentation, etc. In defining the input data one has to make sure that all the data recording peculiarities discussed in Section 2 of this tutorial are taken into account if the data is coming to a given object from a recording instrument. The outputs are typically related to some raw data and various reports.

Purpose of the analysis. Again, it is pretty clear that the purpose of the analysis needs to be specified in some simple but yet comprehensive way. For example, if a peace of equipment is analyzed, it has to be clearly stated if this analysis is related to the correct operation, maintenance schedules, performance assessment, catastrophic failures, etc. Further more, one needs to define if the purpose includes any possible action to be undertaken as a consequence, or is this a troubleshooting action for the reasons of better understanding a given situation, etc. Last but not least, the purpose needs to be justified from the standpoint of the investments and possible gains obtained should this analysis be implemented and put to service.

Hypothesis assumptions. In any analysis, one needs to define a hypothesis that the analysis is trying to match. The whole concept of the analysis is related to developing certain assumptions about what is a known or a correct event, and then performing the analysis to confirm or deny the assumption. In the process of doing so a set of input data is used and then the data is processed in many different ways to see if the results can resemble a known hypothesis about the event.

Time response. In many analysis applications the time it takes to produce an output after the input was made available is critical. The analysis objects that are time critical need to be defined by explaining where the time requirements are coming from. The time requirement may be associated with the ability/inability to provide inputs in a timely fashion The requirement may also be associated with the processing time needed to complete the analysis, or with the time required to provide outputs, in particular the reports, in a given format.

Intended uses and users. The analysis object needs to be put in the context of particular uses of the results as well as the type of users. First, it has to be decided if the immediate uses are intended for the other analysis objects or personnel. Hence the definition indicating if the users are humans or other software modules. Whenever the humans are involved, special care has to be exercised to make sure that the analysis object is not competing, or is not in conflict with activities of the personnel. Sometimes it may appear as obvious that the analysis object will help the operators to do the job better, but the personnel in question may think differently due to whatever reason.

3. LEVELS OF THE ANALYSIS COMPLEXITY

It is pretty important to note that the power system and its associated monitoring, control and protection equipment may have complex behavior, and the related analysis can get really complicated. In this sense, it is important to recognize the following levels of the analysis complexity:

- Individual events
- Individual devices
- Events and devices operating in a cause and effect mode
- System-wide interactions including multiple events and devices

Individual events. Typical example of this level of the analysis complexity is determination if a given deviation from the normal power system waveforms (for example 50Hz positive sequence phasors) is a fault, normal change in system loading, or a disturbance caused by a power quality related event. This may be concluded observing the change in the phasor frequency, amplitude or phase, as well as certain other signal features calculated using signal samples. In any case, if the event is a fault, further complexity of the analysis is associated with trying to determine the type, location, inception angle, and resistance of the fault. Of course, there are many other examples of the individual events that may need to be analyzed to better understand a related power system operation.

Individual devices. Typical example of this level of the analysis complexity is determination if a relay has operated correctly, and if not, what are the causes. This can be done using a hypothesis associated with intended operation of the relay. In any case, there are at least two possible situations regarding the data available for the analysis one is that relay input and output data are recorded by an independent recording system; the other is that the analysis data is taken from the relay itself. In the case the relay input data is recorded by a separate recording system, it may be possible that the analog waveforms recorded are not taken from exactly the same CTs and CCVTs as the ones taken by the relay, but the recorded waveforms resemble closely the relay input data. The analysis approaches are dependent on the data types.

Events and devices operating in a cause and effect mode. Typical example of this level of the analysis complexity is the case of a relay being connected to the power system and the circuit breaker. It may be operating in an interactive mode using a communication scheme to communicate with the relay at the other end, and going through a re-closing sequence due to the nature of the fault. In this case the analysis is not only related to the two previously described levels of the analysis (event and device), but also to a new level associated with analysis of the interactions between the relay and communication channels as well as between the relay and circuit breaker. The overall analysis follows a hypothesis of a cause and effect mode of operation, namely an event causing the relay to operate, and then the relay causing the breaker to operate. Due to the mentioned interactions, the analysis needs to be more detailed representing various action/reaction steps.

System-wide interactions including multiple events and devices. Typical example of this level of analysis is the case when the operation of the entire protective relaying system in a given power system is analyzed for the contingencies including cascading trips leading to a black out or partial blackout (brown out). In this situation the analysis may consist of a number of the individual cases described above (events, devices, cause-effect operation). Further more, a pretty complex set of additional interactions between the power system and protective relaying system may be included by making appropriate correlation among the previously described cases.

4. IMPLEMENTATION TECHNIQUES

To provide an implementation framework for the analysis discussed in this tutorial, some of the most common techniques for implementing the automated analysis systems are discussed. Obviously, the manual analysis is also an option, but it is not further discussed since it is well known and understood. The choice of the techniques used for the automated analysis is made based on the prevailing usage reported in the technical literature, and does not in any sense represent the ultimate choice. The following techniques used in particularly suitable for a given type of the analysis: automated analysis are discussed emphasizing some inherent properties of each of the technique making them

- Signal Processing
- Expert Systems
- Neural Nets
- Fuzzy Logic
- Hybrid solutions

Signal Processing. It is obvious that some form of the signal processing will take place in almost any analysis implementation that involves analogue waveforms, but some signal processing may also be used for processing of contact status information as well [2]. In the past, the most common signal processing techniques used were the ones based on orthogonal transforms such as the Fourier transform and its derivatives: Fast Fourier Transform (FFT), and Discrete Fourier Transform (DFT). The reason was pretty simple: most of the analysis approaches were based on extraction of phasors, and those techniques are pretty powerful in that respect. However, as the analysis approaches have improved and expanded recently, some new signal processing techniques such as wavelets and variety of other digital filters were introduced.

Expert Systems. The use of expert systems in the implementation of the analysis applications is probably the oldest approach taken and the early implementations date back to the late seventies [3]. The reasons for using this group of techniques is obvious: the analysis is a decision making process aimed at a number of comparisons and consequent searching steps. The expert system techniques are very well suited for that purposes. Actually some of the earliest expert system solutions were associated with medical applications where the diagnosis of the illnesses was the aim. The power system fits this concept since one is trying to diagnose the power system and related equipment behavior based on a number of measurements and hypothesis similar to what is done in the medical, or any other diagnosis applications.

Neural Nets. It is well known that neural nets can be a powerful approach to parallel processing of input signals where rather simple and computationally efficient implementation of otherwise complex non-linear relationships can be achieved. The analysis may however place some unique requirements on the neural net selection and application where a simple nets with as few as possible of the processing layers may have to be selected in order to achieve the time performance requirements. Even though the neural nets have been shown to act as very powerful pattern recognizers, some draw backs to their use in the analysis are quite serious [4]. As the neural net may have to be extensively trained to become an acceptable classifier, the issue of the selection of the training sets and methodology needs special attention since different analysis tasks may require different approaches to this issue.

Fuzzy Logic. This set of techniques is often used when dealing with imprecise and/or incomplete data [5]. The analysis function may have quite a few cases where data is of the described nature either due to the lack of an accurate data recording system, or due to the lack of certain types of data altogether. The theory of defining the Fuzzy sets, variables and logic operations is well know and straight forward. However, applying the theory to the analysis tasks may not be as straight forward since a considerable knowledge about the event/ device being analyzed is needed to be able to make a selection of the variables and their typical values. Most of the fuzzy logic applications to fault analysis proposed so far were primarily related to the classification of the power system and equipment states, but the information provided by the authors regarding the guidance and justification of some of the above mentioned selections was minimal. For future use of these technique, better understanding of the benefits and constraints needs to be achieved.

Hybrid Solutions. The nature of the analysis process may vary depending on the data processing and logic decisions made at different stages and complexity levels. Hence, in a given case, it may be very difficult to utilize any single techniques mentioned and achieve the required performance. This is due to the variety of needed properties that none of the techniques may have on its own. Quite often the best way to proceed would be to utilize several techniques in a hybrid solution [6].

5. CONCLUSION

The scope of the analysis in this tutorial focuses on the analysis goals, approach, complexity level and implementation techniques. It is considered crucial that the goals are clearly stated indicating the main reasons for the implementation as well as the expected benefits. It was suggested that the analysis complexity should be structured using hierarchical levels allowing for gradual progression from less to more complex analysis objects. The implementation techniques discussed in this section are quite versatile and have been used in the past for variety of the analysis applications. A very large volume of literature has been produced so far describing the solutions and the performance characteristic. However, very few solutions published so far gave a clear justification why a given technique was used for a given application. This of course is the core of a sound engineering solution, and the lack of this information does not help when practical applications need to be implemented and evaluated. For any new application, it is essential that the selection of the proposed implementation techniques as well as the approach for performance evaluation are well understood.

6. ACKNOWLEDGEMENT

The author wishes to acknowledge the financial support from Electricite de France (EdF) that was provided during author's sabbatical stay at EdF and made the work on this section possible. EdF is also acknowledged for permitting the author to use parts of the text otherwise developed for some internal EdF reports.

7. REFERENCES

[1] ANSI/IEEE C37.100-1992, *Standard Definitions of Power Switchgear*, IEEE Operations Center, Piscataway, NJ, USA.

[2] Kezunovic, I. Rikalo, "Use of Intelligent Systems and Advanced Signal Processing Techniques in Automated Analysis of Disturbances and Protective Relay Operations," *1998 Southern African Conference on Power System Protection*, Johannesburg, South Africa, November, 1998.

[3] Z.Z. Zhang, G.S. Hope, O.P. Malik, "Expert Systems in Electric Power Systems – A Bibliography Survey," *IEEE Transactions on Power Systems*, Vol. 4, No. 4, October, 1989.

[4] M. Kezunovic, "A Survey of Neural Net Applications to Protective Relaying and Fault Analysis," *Engineering Intelligent Systems*, Vol. 5, No. 4, pp. 185-192, December 1997.

[5] J.A. Momoh, X.W. Ma, K. Tomsovic, "Overview and Literature Survey of Fuzzy Set Theory in Power Systems," *IEEE Transactions on Power Systems*, Vol. 10, No. 3, pp. 1676 – 1690, August 1995.

[6] M. Kezunovic, I. Rikalo, C.W. Fromen, D.R. Sevcik, "An Advanced Hybrid Solution for Automated Substation Monitoring Using Neural Nets and Expert System Techniques," *Stockholm Power Technology Conference*, Stockholm, Sweden, June 1995.

Section IV

IMPLEMENTATION USING DFR DATA AND SYNCHRONIZED SAMPLING

M. Kezunovic
Texas A&M University
College Station, TX 77843-3128, USA

Abstract: This section describes a new concept of automated fault analysis where fault transients and changes in power system equipment contacts are processed on-line. This allows faster confirmation of correct equipment operation and detection of unexpected equipment operations, as well as increased accuracy of fault location and analysis. In addition, the section discusses three independent utility examples that illustrate automating some aspect of the fault analysis process. One approach is the substation level analysis, where local digital fault recorder (DFR) data is processed at the substation to obtain accurate fault location and analysis. Another approach is DFR data analysis at the master station location, where all DFR data files from remote locations are concentrated and processed. Finally, an example of a highly accurate fault location system for series compensated using global positioning system (GPS) synchronization is presented.

1. INTRODUCTION TO THE PROBLEM

Development of new technologies such as intelligent systems and synchronized sampling as well as increased utility deregulation and competition are leading to the introduction of new applications and solutions in the fault analysis area.

The early approaches to fault analysis using intelligent techniques were related to alarm processing in a Supervisory Control and Data Acquisition (SCADA) system [1]. At that time, expert system techniques were utilized to implement an automated analysis of alarms [2]. The SCADA based solutions did not have the capability to calculate fault location, and processing of analog waveforms was not done due to the lack of sampled waveform data. Further improvements of the overall solution were achieved using neural network (NN) implementations [3,4].

A study of the possible approaches to fault analysis using digital fault recorder (DFR) data revealed some advantages due to the ability to calculate fault location and correlate waveform samples with protective relay and circuit breaker contact operation. This has enabled a new approach to fault analysis to be implemented using expert systems and DFR data [5-7]. Further developments in this area indicated that a very accurate fault location approach can also be developed using DFRs enhanced with accurate data acquisition interfaced to global positioning system (GPS) receivers [8,9]. Use of neural nets for fault detection and classification was also investigated to enhance the overall fault analysis solution [10,11].

The general concept of automated fault analysis using substation data, expert systems and synchronized sampling is presented first. The new concept is related to analysis of the data coming from variety of substation equipment such as DFRs, sequence of event recorders (SERs), protective relays, and other intelligent electronic devices (IEDs).

This section also summarizes results from three different projects aimed at automated fault analysis. Two projects demonstrate possible approaches to automated fault analysis using DFR data and expert systems. The third project illustrates a highly accurate fault locator based on synchronized sampling using commercial data acquisition system with GPS receivers. A set of conclusions and related references are given at the end.

2. A NOVEL APPROACH TO AUTOMATED FAULT ANALYSIS

The ultimate fault analysis system should provide results of a detailed system-wide analysis of an event to the system dispatchers and protection engineers within seconds after the event occurred. This may not be feasible with the existing SCADA solutions. The main reason is the lack of detailed information about transient waveforms and contact changes that are not readily available through Remote Terminal Units (RTUs) of a SCADA system. On the other hand, such information is available through other Intelligent Electronics Devices (IEDs) including DFRs, Sequence of Event Recorders (SERs) and Digital Relays (DRs). A new concept for fast and accurate fault analysis can be developed using this equipment technology, high-speed data communication infrastructure and advanced software techniques.

Various types of users have different needs regarding the time response and/or extent of information provided by the fault analysis system. The system dispatchers are interested in getting the condensed fault analysis information as soon as possible after the valid fault occurs. Their main interest is determination of accurate fault location and switching equipment status that enables them to make decisions about the system restoration. The protection engineer, on the other hand, is most interested in getting detailed and specific information regarding the operation of the protection system and related equipment during the event. The time factor is not as strict as for the system dispatcher.

In this subsection we present the concept of an integrated fault analysis system that can be built with existing technology and can satisfy both types of users. The subsequent subsections give brief presentations of various research and development projects that are an illustration of possible steps towards the final system solution for automated fault analysis.

Figure 4.1 presents the block diagram of one possible implementation of an integrated system. Each substation is equipped with a PC (low end Pentium machine) that collects data from different devices (e.g., DFR, SER, DR), and analyzes that data locally. The results as well as raw data files are communicated to the central file server in a common COMTRADE format [12]. The substation analysis provides fault location and fault type based on the data recorded at this location. This data can be made available to the system dispatcher and protection engineer within a minute after the recording was made by appropriate device. The information is communicated in the form of a fax.

After this initial faxing, the substation PC establishes communication with the Central File Server and uploads event data to it. The System Wide-Analysis software monitors incoming event files and correlates files coming from different locations based on their accurate time stamps and samples that are synchronously taken at all substations using GPS receivers for synchronization. The system fault analysis is then executed using data from various locations to produce a summary report for protection engineers.

It is important to note that such an integrated solution is not yet fully implemented since the design provisions to implement the synchronized sampling for all substation data acquisition systems are not readily available [13]. In addition, utilities are still searching various options to provide standard communication architecture allowing high-speed substation-wide data acquisition and transfer to the centralized substation and system level location [14].

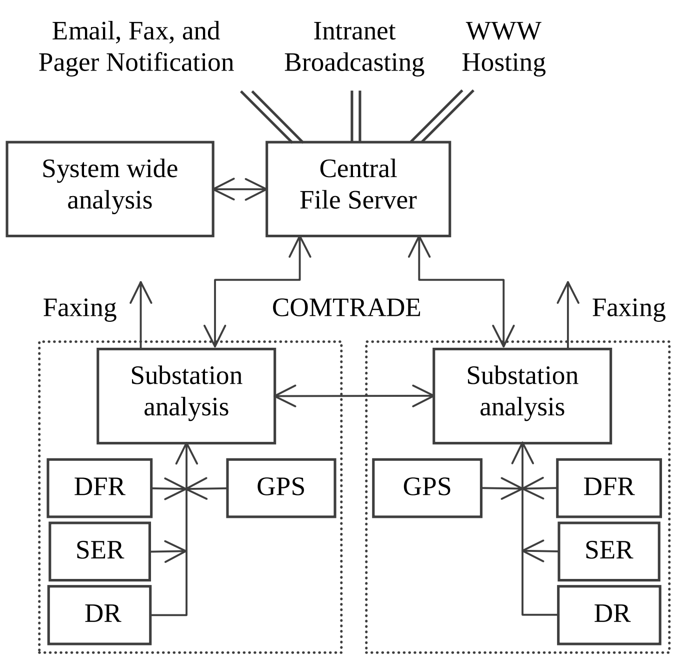


Figure 4.1. Conceptual diagram for the New Fault Analysis Approach

An assessment of the existing technology reveals that future developments in the IEDs for substation metering, control and protection are leading to the following expected improvements [13]:

- Increased accuracy of the A/D subsystem
- Synchronized sampling on all input channels
- Multiple communication interfaces for high-speed data transfer
- Availability of extensive signal processing for calculation of various measuring quantities

Combined with on-going developments in the utility communication architecture, it is expected that the substation IEDs will provide an impressive level of detail of the data needed to perform automated fault analysis.

In the meantime, some solutions that are less involved can be implemented using the technology that is readily available. In particular, DFR data analysis can be automated using expert system technology. In addition, fault location accuracy can be improved using GPS receivers to synchronize customized data acquisition units located at two ends of a transmission line. The remaining discussion in this section illustrates how several different approaches can be undertaken using the existing technology. It is expected that in the future, the substation equipment will have all the required design provisions so that an optimized use of the technology can be achieved maximizing the cost-performance benefit.

3. A CASE STUDY OF RELIANT ENERGY HL&P

Occurrence of a fault on a major transmission line may endanger the operation of a bulk power system and potentially lead to costly outages. If the fault analysis results are not available to system operators shortly after the fault occurred they might not be able to reach an optimal decision regarding the restoration of a line. In the era of increased competition between utilities due to the open access and retail wheeling, any unnecessary delay of energy supply restoration compromises a utility's competitive position.

The project described in this section is aimed at utilizing existing DFR data to provide a system dispatcher with accurate and timely information regarding the fault type and fault location, as well as an analysis of the operation of protection system and related equipment. The dispatcher can use this information to decide if a transmission line should be restored back to service or a maintenance crew dispatched.

3.1 System Architecture

The block diagram of the system that was developed is shown in Figure 4.2. The expert system communicates with DFR over a high-speed parallel link. It interrogates the DFR for new recordings on a continual basis. A new data file is copied from a recorder and immediately analyzed. The analysis report is created and faxed to the system dispatcher and to the protection engineer's office. The whole process takes less than a minute, so valuable information is available to the system dispatcher in a relatively short period of time after a fault was recorded by the DFR.

The analysis report typically contains the following information:

Table 4-I. Behavioral Patterns of the Basic Parameters

Event Type	**0 Sequence Current**	**Faulted Current**	**Unfaulted Current**	**0 Sequence Voltage**	**Faulted Voltage**	**Unfaulted Voltage**	**Line Voltage**
a - g	$I_0 > .2I_a$	$I_a > 1.4I_p$	$I_{b,c} < .331_a$	$V_0 > .04V_n$	$V_a < .9V_n$	$V_{b,c} > .96V_n$	$V_{ab} \approx V_{ca}$
a - b	$I_0 < .01I_a$	$I_a > 1.4I_p$ $I_b > 10I_p$	$I_c < 11_a$	$V_0 < .01V_n$	$V_a < .8V_n$ $V_b < .7V_n$	$V_c > .99V_n$	$V_{ab} < .8V_{luf}$
a - b - g	$I_0 > .1I_a$	$I_a > 1.4I_p$ $I_b > 10I_p$	$I_c < .11_a$	$V_0 > .05V_n$	$V_a < .8V_n$ $V_b < .8V_n$	$V_c > .98V_n$	$V_{ab} < .8V_{luf}$
a - b - c	$I_0 < .03I_b$	$I_f > 10I_p$		$V_0 < .01V_n$	$V_f < .8V_n$		$V_{lf} < .8V_{ln}$
a - b - c- g	$I_0 < .03I_b$	$I_f > 10I_p$		$V_0 < .01V_n$	$V_f < .8V_n$		$V_{lf} < .8V_{ln}$

- Event date/time stamp and DFR identification
- Fault type, fault location, and transmission line involved
- Relay tripping times, breaker opening times, and carrier signaling
- Snapshot of RMS values for selected analog channels

The fault analysis logic incorporated in the expert system's knowledge base relies on signal processing algorithms to extract aggregate parameters such as RMS values for phase currents and bus voltages from samples recorded by DFR. These parameters are then passed through the set of rules that represent relationships between system variables during different fault (or normal) operating conditions. The mathematical relationships between various parameters for certain fault types are shown in Table 4-I [5].

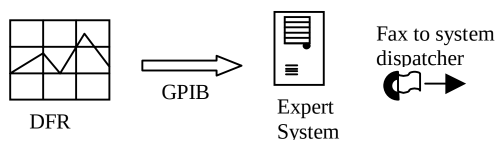


Figure 4.2. Diagram of the Substation Analysis System

3.2 Knowledge Acquisition and Rule Definition

The knowledge necessary for disturbance analysis was acquired by interviewing experts (protection relay engineers) and by using an empirical approach based on EMTP simulation studies. The reasoning process includes the following steps: fault detection, fault classification, event analysis, and protection-system and circuit-breaker operation analysis.

Fault detection and classification can be described by the following procedure, as outlined by the experts:

- Fault inception instant is detected by looking for the abrupt change in signal waveforms.
- Voltage waveforms are checked for a change in the fundamental harmonic amplitude. A voltage decrease indicates the possible faulted phases.
- Current channels of the phases that experienced a significant voltage decrease are checked next. The current that experienced the greatest amplitude increase indicates the probably faulted circuit.

The overall change in voltage and current waveforms indicates the type of fault (e.g., phase A to ground). It also points to other characteristics of the fault and the behavior of the protection system (fault clearing, reclosing).

Event and protection system operation analysis includes the following checks:

- Relay and breaker contacts' state is checked for a change. A status change is an indication that the protection system has detected a fault.
- If the protection system operation is detected and the presence of a fault is not identified, it is an indication of a protection system misoperation.
- If a fault is detected and there is no protection system operation, it is an indication of a possible protection system failure.

The reasoning required to perform classification and analysis of the event is implemented by using a set of rules. The reasoning process is separated into two stages. In the first stage, the system reasons on the basis of the analog-signal parameters, and in the second step, it reasons by using the protection-system parameters. Analog-signal and protection-system parameters are obtained by processing the recorded samples and extracting the relevant features of the signals recorded on the line that had experienced the largest disturbance.

A typical set of rules based on the analog parameters is shown in Figure 4.3. A sequence of checks is indicated by the circled numbers next to the rule definitions.

This set of rules represents the application's knowledge about the operation of a power system section in the form of "rules of thumb". The rule base is expandable and can be changed

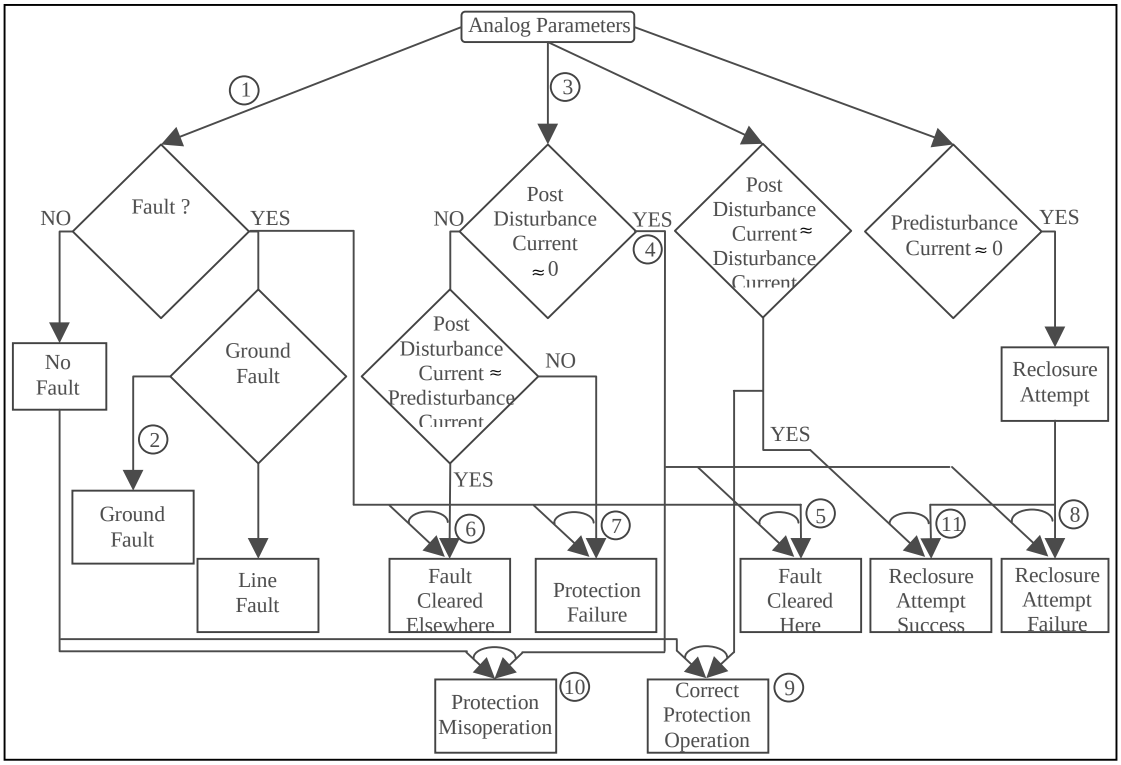


Figure 4.3. Rules for Fault Detection, Classification, and General Event Analysis Using Analog Parameters

over time, when a better understanding of particular operations of power system equipment becomes available.

To facilitate modularity and extensibility of the analysis logic, a "C Language Integrated Production System" (CLIPS) expert system tool was embedded in the application. This tool allows addition of new rules which specify a new set of actions to be performed for a given power system operating condition. Figure 4.4 shows an example of a CLIPS rule to determine if particular conditions for a phase-to-ground fault are met. The exact thresholds (multiplication coefficients) will change from substation to substation, and may need to be determined by trial and error as well as modeling and simulation.

```
(defrule  AG_fault
    (Ipre  ?Iap ?Ibp?Icp ?I0p)            (Iflt ?Ia ?Ib ?Ic ?I0)
    (vpre ?vap ?vbp ?vcp ?v0p)            (vflt ?va ?vb ?vc ?v0)
    (vlp ?vabp ?vbcp ?vcap)               (vl ?vab& ?vbc ?vca)
    (test (> ?I0 (* 0.20 ?Ia )))          (test (> ?Ia (*1.40 ?Iap)))
    (test (< ?Ib (* 0.33 ?Ia )))          (test (< ?Ic (* 0.33 ?Ia )))
    (test (> ?v0 (* 8.00 ?v0p)))          (test (< ?va (* 0.90 Ivap)))
    (test (> ?vb (* .96 ?vbp)))           (test (> ?vc (* 0.96 ?vcp)))
    (test (< (abs (-?vb ?vc))  (* .05 ?vbp)))
    (test (< (abs (- ?vab ?vca))  (* .2 ?vabp)))
    (test (< (abs (-?vbc ?vbcp)  (* .2 ?vbcp)))
=>
    (format t "AG_Fault fired%n")
    (assert (FaultType "phase A to ground fault")) )
```

Figure 4.4. Example of a CLIPS Rule

The expert system software is fully automated. Once configured, no operator interaction with the system is needed. The system reports is operating status on a daily basis by sending a fax message to the dispatcher's and protection engineer's office.

4. A CASE STUDY OF TXU ELECTRIC

The objective of this research project was to streamline DFR data files that are coming from many different locations and archive them on a corporate LAN using certain classification criteria. The basic data flow diagram is shown in Figure 4.5. The DFR Master Station PC's #1 through #3 are responsible for communicating with remote recording units via dial-up modem lines. The Master Station units can be configured to automatically poll remote recorders on periodic basis and retrieve new events, or substation DFRs can be setup to automatically call a Master Station when they have a new event to report. For this project, the second option was used.

It is worth mentioning, at this point, that this system is configured to classify files coming from DFRs made by two different vendors. The classification system has been generalized to allow easy incorporation of additional vendor's digital recording systems, as long as the particular DFR vendor provides DFR file format description. For utilities that may have DFRs from multiple vendors, this classification system feature provides a common platform for fault analysis and the distribution of results. In addition, the common platform eliminates the need to train employees to use multiple DFR manufacturer analysis packages.

4.1 The Classification Engine

To facilitate the classification process as well as distribute classification results across the corporate Intranet, a dedicated

File Server PC and a Classification Engine are secured. The Classification engine diagram is the "brain" of the system. It monitors assigned incoming file directories on a File Server and detects any new DFR data file that has been copied from Master Stations. These new files are processed using built-in logic to produce a classification report. Finally, the Classification Engine automatically converts the raw DFR data file into the COMTRADE format [12] and copies it with associated classification report to an assigned directory on Windows NT File Server.

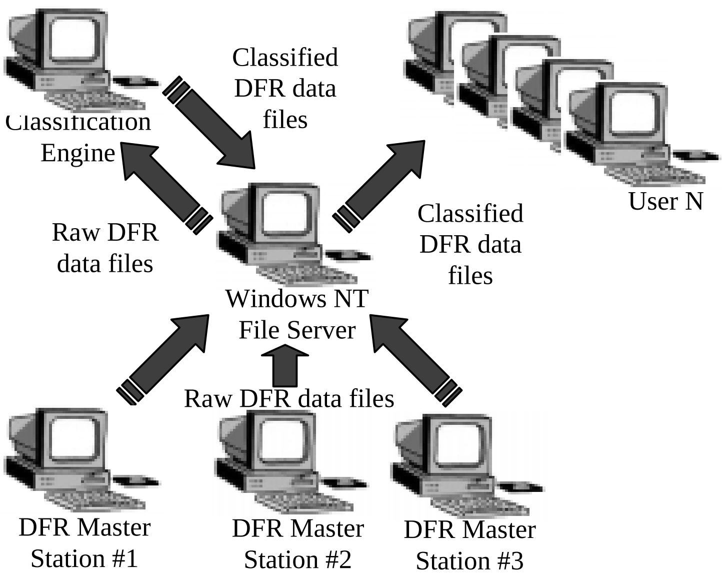


Figure 4.5. Classification System Block

The File Server is a repository of both the raw DFR data files in its native format and the processed DFR files in the COMTRADE format. The easy access to the processed DFR files is facilitated by archiving all data files into three categories (high, medium and low priority) based on certain criteria. A separate software package for viewing the processed DFR files is available. This package provides an integrated environment for displaying both the conclusions about the analysis of a selected DFR file, as well as graphics in the form of analog and digital traces.

One of the tasks of the Classification Engine is to reduce the time that system protection personnel spend on manual examination and archival of DFR records. This system automatically classifies and filters DFR records based on the following broad criteria:

- The fault condition exists and clearing time is satisfactory.
- The fault condition exists and clearing time is longer than expected.
- The fault condition exists and breaker restrike and/or ferro-resonance occurs during fault clearing.

Table 4-II. Input Signals Used by Classification Engine

Digital	Primary and backup relay trip	
	Breaker open position	
	Breaker close position	
	Breaker failure (BF) contact	
	Carrier Start and Carrier Received contacts	
Analog	At least two phase currents	(Ia, Ib,Ic)
	Residual current	(Ir)
	All three phase (bus side) voltages	(Va, Vb, Vc)
	Residual (neutral) voltage	(Vr)

Table 4-II shows the list of signals that are used for the DFR file classification logic. If the signal is not monitored in a particular DFR configuration, associated classification logic cannot be implemented. In the case where only the two phase currents plus residual current are monitored, the third phase current will be calculated automatically by the Classification Engine.

The following parameters are extracted and/or calculated from every DFR record:

- Relay trip times and relay pick-up time
- Breaker open/close times and breaker pick-up time
- Breaker failure start times
- Carrier start/receive times
- Calculated fault inception time

The classification logic is based on the analysis of the above parameters. The following are the events that can be recognized and flagged by the system:

- Slow relay clearing
- Breaker failure or slow breaker clearing
- Breaker restrike
- Carrier misoperation
- PT Ferro-resonance
- Reclosure failure, Line lockout

The Classification Engine keeps a detailed log of the system events during its operation. System events such as the names of the corrupted DFR files, or names of the incomplete DFR files are time stamped and recorded. The logging capabilities help the administrator troubleshoot the system operation on a daily basis.

The Classification Engine archives all incoming DFR files into three categories depending on the type of the event. These categories are High, Medium, and Low priority (see Figure 4.6). Events such as the normal fault clearing, or reclosure success will be categorized as high priority events and archived in the High priority folder on the central file server. Events such as the normal fault clearing, or reclosure success will be categorized as the medium priority. And finally, the events such as no operation will be stored in the Low priority folder.

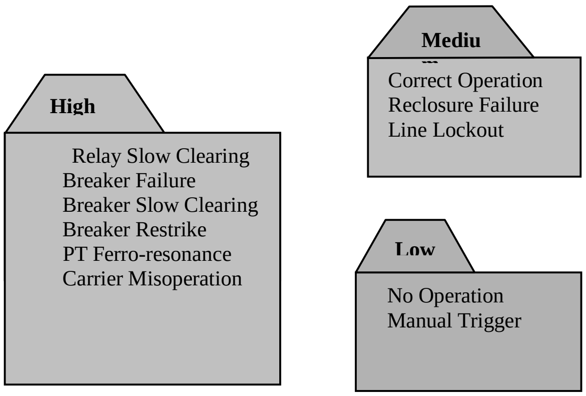


Figure 4.6. Classification Logic Categories

4.2 The Report Viewer

The Report Viewer is the Windows 95 client software used for accessing classification reports form the central file server. The module has an extensive graphical user interface (GUI) that allows users to access DFR reports and data files either locally (when directly connected to corporate LAN) or remotely (when connected to corporate LAN over a dial-up modem line). The Report Viewer application's main window consisting of three parts: network/local director display, waveform display, and classification report display.

The user may choose the event priority that he/she wants to access and display in the directory view. The default priority is High. The network/local directory display contains three columns: name of the DFR that recorded particular event, date/time stamp and short description of the event. For accessing the data over WAN (via dial-up connection), the application provides a caching function, similar to Internet browsers. This means that the data once down-loaded will be saved in the caching directories on the local drive, thus eliminating the need to retrieve the same event files over the WAN multiple times. In addition, the caching function enables the user to view the downloaded data files off-line.

The textual display of the Report Viewer presents the following information to the user (Fig. 4.7):

- Event Date/Time Stamp
- Event Type (e.g. breaker failure, etc.)
- Event Size (prefault, fault, postfault cycles)
- DFR Type and Recorder ID
- Breaker Operation Time
- Operation of Carrier Channels
- RMS values for associated Breaker ID's analog channels per cycle for every cycle in the record (the display will be color coded for prefault, fault and postfault intervals)
- Harmonic content of associate analog channels in a tabular form.

The waveform display of the DFR file presents graphs of analog and digital signals. This display has the following properties:

- Selectable DFR channels to display
- Tickmarks on the x and y axis
- Auto-scalable x and y axis
- Selectable time axis (milliseconds, cycles, or samples)
- Selectable waveform coloring
- Colored markers on the analog traces where the digital channel operation occurred
- Zooming capability
- Legend containing channel description and values of analog and digital signals at cursor position
- Measure of the time span between two points on the screen
- Waveform printing and print preview

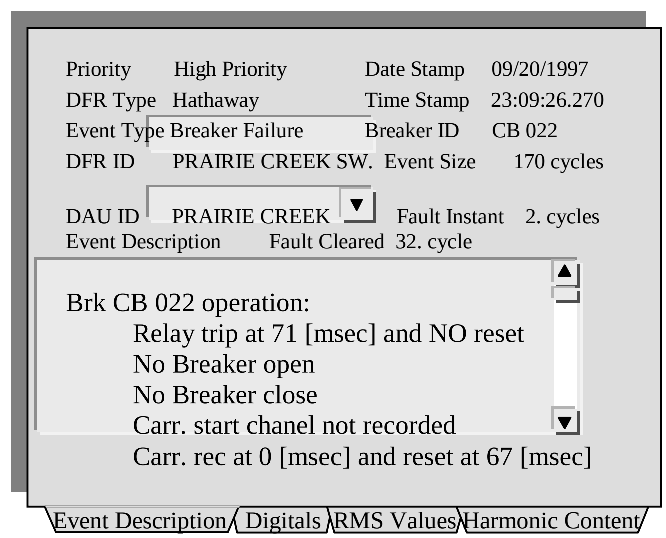


Figure 4.7. The Report Viewer Text Display

5. A CASE STUDY OF WESTERN AREA POWER ADMINISTRATION

As noted earlier, the accurate and timely information regarding fault location, after a transmission line fault has occurred, is most important to system operators. They need to confirm and isolate the faulted section before any system restoration is attempted. Then dispatch maintenance crews directly to the fault site.

Most of the existing fault location algorithms use data from one line end, due to the large cost of additional equipment involved in obtaining the data from the other end as well [15,16]. Recently, the cost of the necessary hardware is rapidly decreasing, which makes implementation of two ended fault location algorithms cost effective for critical transmission lines. The two ended fault location algorithms are inherently more accurate and robust than single ended ones [8,17].

5.1 Fault Location Using Synchronized Sampling

The fault analysis system presented in Figure 4.1 incorporates design features needed for implementation of an advanced fault location technique based on synchronized sampling. Figure 4.1 shows the case where two neighboring substations are equipped with GPS receivers. The GPS receivers are used for accurate synchronization of recording devices. Two substation PCs communicate with each other via dial-up modem lines and exchange fault waveform samples taken synchronously.

One of the most important requirements for this fault location algorithm is a fast, reliable and accurate data acquisition subsystem. This can be achieved either by using separate data acquisition with customized signal conditioning hardware, or making improvements in the existing data acquisition subsystem built in the customized DFR [18]. The first approach increases the cost and complexity of the hardware installed in the substation. The second is preferred if the existing DFRs can be upgraded.

As total cost of implementing this advanced fault location system decreases over time, we expect wider acceptance of the technology by utilities that want to gain comparative advantage by having accurate and up-to-date information regarding their transmission grid.

High sampling rate requirements are imposed on the data acquisition system due to the fact that the fault location method is based on discretization of Bergeron's traveling wave equations or lumped parameter line equations [19,20]. In order to derive these equations we can consider the unfaulted long transmission line shown in Figure 4.8. A transmission line longer than 150 miles 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

$$v_F(t) = \tfrac{z}{2}\left[i_S(t-\tau_x) - i_S(t+\tau_x)\right] + \tfrac{1}{2}\left[v_S(t-\tau_x) - v_S(t+\tau_x)\right] \tag{1}$$

$$i_F(t) = -\tfrac{1}{2}\left[i_S(t-\tau_x) + i_S(t+\tau_x)\right] - \tfrac{1}{2z}\left[v_S(t-\tau_x) - v_S(t+\tau_x)\right] \tag{2}$$

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 to point *F* from *S*. They are defined as

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

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

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

The distance to the fault does not appear explicitly in the equation. When the equation is discretized based on the sampling interval, the travel times to the point *F* from either end will not be exact any more. The right hand side of Equation 4 will have a finite non-zero value. Now, based on the sampling time step, the line can be divided into a number of discrete points, and Equation 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 fault point.

This method is strongly dependent on the sampling frequency. To reduce this requirement, the approximate point is used as a guideline. Once the minimum error point is obtained, the voltages and currents at the points adjacent to this point can be computed using the discretized versions of equations 1 and 2, the single end equations.

The line section between the adjacent points is now modeled as a short transmission line and the fault location is calculated more accurately. Further accuracy improvements can also be achieved for mutually coupled lines if the synchronized measurements are available from the terminals of the coupled lines [22].

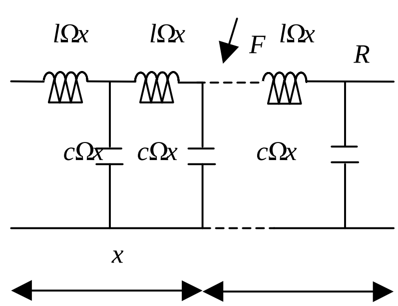


Figure 4.8. Unfaulted Long Transmission Line

5.2 Results From Testing The Algorithm

This section presents some results from testing the algorithm developed in the previous section. All simulations are done in EMTP [22].

Utility Transmission System

The testing of the algorithm was then carried out using data generated from the EMTP simulation of a real power system, belonging to Western Area Power Administration. The one line diagram of the system is shown in Figure 4.9. The line of interest is the 525 kV line from the Mead to Westwing sub-station. The line is 242.4 miles long and is mutually coupled to the 345 kV Mead-Liberty lines are equipped with series compensation capacitors at each end. The capacitors are protected by Meal Oxide Varistors (MOVs). For a detailed description of the system, please see reference [23].

The voltage and current data are acquired from the Mead 525 and 345 kV buses, the Westwing 525 kV bus and the Liberty 345 kV bus, to give a total of 4 three phase voltages and 4 three phase currents. No data is available from the Palo Verde-Westwing double line.

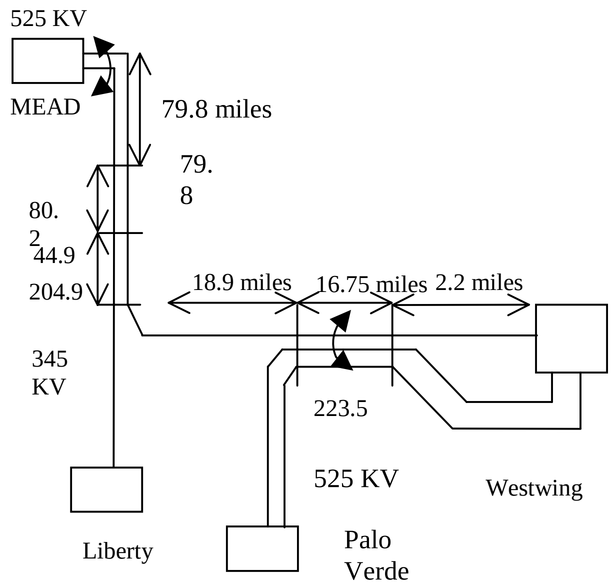


Fig. 4.9 One Line Diagram of the Sample Power System

On this system, extensive EMTP simulations were carried out, with the following variations in the simulation conditions:

- Fault Locations; Faults were introduced at 40.0, 79.8, 160.0, and 223.5 miles from Mead.
- Fault Types: Four types of faults were considered; Phase A to Ground, Phase B to C, Phase B to C to Ground and Three Phase to Ground.
- Fault Incidence Angle: 0º and 90º.
- Fault Impedance; 3Ω and 50 Ω.
- Series Capacitors: Capacitors In and Capacitors Bypassed.
- Instrument Transformers: Data collected from the primary or from the secondary of the instrument transformers. The instrument transformer models used in this study were developed earlier [24,25].

This gives a total of 4 x 4 x 2 x 2 x 2 x 2 = 256 test cases.

Test Results

The following tables show the results for Phase A to Ground faults, and Phase B to C faults. Each column in the tables correspond to the following four system conditions respectively:

- **PC**: Data from Primary, with Series Capacitors
- **PN**: Data from Primary, without Series Capacitors
- **SC**: Data from the Secondary, with Series Capaciators

The error % is shown in each cell of the table, and below it is the distance as computed by the fault location algorithm. The error % is the worst case error of four different fault scenarios which are obtained for two fault impedances and two incidence angles. The error is calculated by the formula given below:

Table 4-III. Maximum Fault Location Error (%) – A-G Fault

Actual Location	Error and Computed Location			
	PC	PN	SC	SN
40.0	0.355 39.15	0.708 38.28	1.498 43.63	1.604 43.88
79.8	0.367 80.68	0.437 80.86	0.246 80.40	1.405 83.21
160.0	1.076 162.61	1.128 162.73	1.084 157.37	1.106 162.9
223.5	2.727 230.11	2.041 228.45	1.570 227.30	1.339 226.75

$$Error\,(\%) = \frac{|Actual\ loc. - Computed\ loc.|}{Line\ Length} \times 100\% \qquad (5)$$

The sampling frequency is 20 kHz. All distances are measured in miles from Mead Substation. The single line to ground fault shows a better accuracy than the multi-phase fault for most of the testing scenarios. The accuracy of the fault location when the data is measured at the primary of the instrument transformers (scenarios PC and PN) is a little better than when the data is measured at the secondary (scenarios SC and SN), for a number of cases. The largest error occurs for the Phase A to Ground Fault at 223.5 miles form Mead, with the series capacitors included in the simulation and with the data measured at the primary. Other than this, there is no clear pattern in the error percentages. The testing also showed that the maximum errors were not associated with one particular fault incidence angle or fault impedance. Errors for the Phase B to C fault are higher in general, than the errors for the single line to ground fault. When data is measured form the secondary of the instrument transformers, the error percentages are lower than the cases when data is measured from the primary. The largest error is 3.791% for a fault at 79.8 miles from Mead, with the series capacitors included and data taken from the primary. This corresponds to a error in the fault location of 9.19 miles. The lowest errors are seen for faults at 40.0 miles. At 79.8 and 160.0 miles from Mead, the errors increase, and then drop again at 223.5 miles. As in the case of the Phase A to Ground fault, there was no association between the maximum error and any particular incidence angle or fault impedance.

Table 4-IV. Maximum Fault Location Error (%) – B-C Fault

Actual Location	Error and Computed Location			
	PC	PN	SC	SN
40.0	1.650 36.00	1.833 35.56	0.402 39.03	0.098 39.75
79.8	3.791 70.61	3.283 71.84	3.513 71.28	0.913 77.59
160.0	3.013 152.70	1.526 156.30	2.880 153.02	0.848 157.94
223.5	1.856 228.00	0.757 2258.34	0.828 221.49	1.058 220.93

Factors Affecting the Fault Location Accuracy

One of the main factors affecting the fault location is the sampling frequency of the data. The sampling interval determines the length of each discrete segment of the long transmission line. In our algorithm however, the fault location does not stop with determining the approximate point. It is further refined by applying the short line algorithm on a segment of the line around the approximate fault point. Therefore, the sampling frequency does not play as important a role as it does for the authors in [26]. However, the sampling frequency must be high enough to ensure that the refinement of the fault location is carried out on a segment that is reasonably short. In our case, this segment is around 16 miles, at a sampling frequency of 20 kHz.

The main factors affecting the accuracy of the fault location are:

- Transposition Points in the Power System: Looking at Figure 4.9, transposition points are seen at 79.8, 160.0, 204.9 and 223.5 miles in the line. These points are sources of reflected traveling waves, which affect the terminal voltage and current data. Our algorithm however, considers the line to be made up of three homogeneous sections: Mead to Fault Point, Fault Point to end of Mutual Coupling (204.9 miles) and from 204.9 Miles to Westwing. If the fault is outside the mutually coupled section, then the algorithm assumes that the line is made up of two homogeneous sections. The reflections from the transposition points affect the accuracy of the voltage and current reconstruction at the points adjacent to the approximate point, which in turn affects the accuracy of the fault location.
- Coupling with the Palo Verde Lines: There is one section 16.73 miles long, where the Mead-Westwing line is coupled with two 525kV line from Palo Verde. This coupling is not considered by the algorithm and is another source of error.

Not modeling components like series capacitors and surge arresters in the fault location algorithm did not affect the accuracy of the algorithm. As can be seen from the results, there is no clear trend to indicate if the presence or absence of the series capacitors in the simulations, affected the errors in any way, since all voltage and current measurements were taken from the line side of the circuit breaker.

6. CONCLUSION

This paper has introduced a general concept of automated fault analysis utilizing data collected by various substation data acquisition equipment, and synchronized using GPS receivers. Since the technology for a full-blown solution is not yet readily available, a variety of solutions can be implemented using existing advanced technology. The following are the projects and related benefits that are being implemented by Texas A&M University and its utility partners:

- High-speed automated substation based fault analysis using DFR data.
- Integrated system wide automated analysis of DFR data from different DFR systems.
- Accurate fault location utilizing synchronized samples from two ends of a transmission line using GPS receivers.

7. ACKNOWLEDGEMENTS

A number of Texas A&M University students and staff from participating utilities have contributed to the results presented in this section. The contribution of former TAMU students Predrag Spasojevic and Igor Rikalo is acknowledged. The following utility advisors also greatly helped in the developments: C.W. Fromen, D.R. Sevcik, and B. Lunsford from Reliant Energy HL&P, W.M. Carpenter and S.L. Goiffon from TXU Electric and S.M. McKenna and D. Hamai from Western Area Power Administration.

8. REFERENCES

[1] S.N. Talukdar, et al., “Artificial Intelligence Technologies for Power System Operations,” *EPRI Report EL-4323*, January 1986.

[2] Z.Z. Zhang, G.S. Hope, and O.P. Malik, “Expert Systems in Electric Power systems – A Bibliographical Survey,” *IEEE Transactions on Power Systems*, Vol. 4., No. 4, October 1989.

[3] H.T. Yang, W.Y. Chang and C.L. Huang, “A New Neural Networks Approach to On-Line Fault Section Estimation Using Information of Protective Relays and Circuit Breakers,” *IEEE Transactions on Power Delivery*, Vol. 9, pp. 220-230, 1994.

[4] C. Rodriguez, et al., “A Modular Neural Network Approach to Fault Diagnosis,” *IEEE Transactions on Neural Networks*, Vol. 7, pp. 326-340, 1996.

[5] M. Kezunovic, P. Spasojevic, C.W. Fromen and D.R. Sevcik, “An Expert System for Substation Event Analysis,” *IEEE Computer Applications in Power,* Vol. 7, No. 2, April 1994, pp. 15-19.
[6] M. Kezunovic, I. Rikalo, C.W. Fromen and D.R. Sevcik, “Expert System Reasoning Streamlines Disturbance Analysis,” *IEEE Computer Applications in Power*, Vol. 7, No. 2, April 1994, pp. 15-19.
[7] L. Cederblad, et al., “A Knowledge-Based System for Automatic Evaluation of Disturbance Recordings,” *1993 Texas A&M Relay Conference,* College Station, Texas, April 1993.
[8] M. Kezunovic and B. Perunicic, “Automated Transmission Line Fault Analysis Using Synchronized Sampling at Two Ends,” *IEEE Transactions on Power Systems*, Vol. 11, No. 1, February 1996.
[9] S.M. McKenna, D. Hamai, M. Kezunovic and A. Gopalakrishnan, “Transmission Line Modeling Requirements for Testing New Fault Location Algorithms Using Digital Simulators,” *Second International Conference on Digital Power System Simulators – ICDS ’97*, Montreal, Quebec, Canada, May 1997.
[10] M. Kezunovic and I. Rikalo, “Detect and Classify Faults Using Neural Nets,” *IEEE Computer Applications in Power,* Vol. 9, Iss: 4, pp. 42-47, October 1996.
[11] M. Kezunovic, I. Rikalo, C.W. Fromen and D.R. Sevcik, “An Advanced Hybrid Solution for Automated Substation Monitoring Using Neural Nets and Expert System Techniques,” *Stockholm Power Tech International Symposium on Electric Power Engineering,* Vol: 4, pp. 335-340, June 1995.
[12] IEEE Standard Common Format for Transient Data Exchange (COMTRADE) for Power Systems, *Power System Relaying Committee of the IEEE Power Engineering Society,* Revision Draft 5.3, December 1996.
[13] M. Kezunovic, “Future Requirements for digital Substation Control and Protection Systems,” *SEE Conference,* Paris, France, March 1997.
[14] M. Kezunovic, “Data Acquisition and Communication Requirements for Advanced Monitoring, Control and Protection,” *CIGRE Symposium on Integrated Control and Communication Systems,* Helsinki, Finland, August 1995.
[15] M. Takagi, et al, “Development of a New Fault Locator Using the One-Terminal Voltage and Current Data,” *IEEE Transactions on PAS*, Vol. PAS-101, No. 8, pp. 2892-2898, August 1982.
[16] L. Erickson et al., “An Accurate Fault Locator with Compensation for Apparent Reactance in the Fault Resistance Resulting from Remote-End In-Feed,” *IEEE Transactions on PAS-104,* No. 2, pp. 424-436, February 1985.
[17] D. Novosel, et al., “Unsynchronized Two-Terminal Fault Location Estimation,” *IEEE Transaction on Power Delivery*, Vol. 11, No. 1, pp. 130-138, January 1996.
[18] M. Kezunovic and I. Rikalo, “Automated System Solutions for Automated Fault Analysis Using DFR Files,” *Precise Measurements in Power Systems and Fault Disturbance Conference,* Arlington, November 1996.
[19] M. Kezunovic, B. Perunicic and J. Mrkic, “An Accurate Fault Location Using Synchronized Sampling,” E*lectric Power Research Journal*, Vol. 29, No. 3, May 1994.
[20] M. Kezunovic, B. Perunicic and J. Mrkic, “An Accurate Fault Location Using Synchronized Sampling at Two Ends of a Transmission Line,” *11th PSCC,* Avignon, France, September 1993.
[21] B. Perunicic, A.Y. Jakwani and M Kezunovic, “An Accurate Fault Location on Mutually Coupled Transmission Lines Using Synchronized Sampling,” *Stockholm Power Tech International Symposium on Electric Power Engineering*, Vol. 4, June 1995.
[22] EPRI, “Electromagnetic Transients Program (EMTP),” *Version 2, Revised Rule Book,* Vol. 1: Main Program, EPRI-EL-4541-CCMP, Palo Alto, California, March 1989.
[23] S.M. McKenna, D.M. Hamai, M. Kezunovic and A. Gopalakrishnan, “Transmission Line Modeling Requirements for Testing New Fault Location Algorithms Using Digital Simulators,” *Proceedings of the Second International Conference on Digital Power System Simulators,* Montreal, Quebec, May 1997.
[24] M. Kezunovic, Lj. Kojovic, V. Skendzic, C.W. Fromen, D.R. Sevcik, S.L. Nilsson, “Digital Models of Coupling Capacitor Voltage Transformers for Protective Relay Transient Studies,” *IEEE Transactions on Power Delivery*, Vol. 7, No. 4., October 1992, pp. 1927-1935.
[25] M. Kezunovic, Lj. Kojovic, A. Abur, C.W. Fromen, D.R. Sevcik, F. Phillips, “Experimental Evaluation of EMTP—Based Current Transformer Models for Protective Relay Transient Study,” *IEEE Transactions on Power Delivery*, Vol. 9, No. 1, January 1994, pp. 405-413.
[26] A.O. Ibe and B.J. Cory, “A Traveling Wave Fault Locator for Two- and Three-Terminal Networks,” *IEEE Transactions on Power Systems,* Vol. PWRD-1, No. 2, April 1986, pp. 283-288.

Section V

IMPLEMENTATION USING DATA FROM SEQUENCE OF EVENTS RECORDERS

Juhwan Jung Chen-Ching Liu
University of Washington

Massimo Gallanti
ENEL, Italy

Abstract. This section summarizes the problem of power system fault diagnosis and its solutions. Our emphasis in this section is on the use of detailed Sequence-of-Events Recorder (SER) messages for analysis of the fault location(s) and possible malfunctioning device(s).

1. INTRODUCTION TO THE PROBLEM

Power systems normally contain a large number of transmission lines, busbars, generators, transformers, and protective devices. These protective devices may be connected with communication devices in order to transfer their operations/status to a central (or regional) location where these devices are monitored. When one (or more) of these components fails, the failure may develop into cascading outages causing a wide spread failure. In case of cascading outages, the amount of information received at the control centers may become too large for dispatchers to analyze within a short period of time.

Fault diagnosis depends on knowledge of the power system state, i.e., the information available at the control centers. The system state has to be clarified before restorative actions can take place. Fault diagnosis involves identification of the fault location (or component) and type. However, it is possible that some protection (or communication) devices malfunction, making it more difficult to analyze a fault event. Thus, a fault diagnosis system should have the capability to identify the malfunctioning device(s). These considerations in fault diagnosis are discussed in section 3. Much of the Artificial Intelligence (AI) application to power systems is related to the system fault diagnosis. These developments can be grouped based on:

Table 5.-I. Developments of Fault Diagnosis Based on Sensors/Protective Devices

Type of devices	Information	Capability
Breaker	The contact status of breakers	Identification of the fault zone is possible. Insufficient for diagnosis of apparatus [1-4].
Current Transformer (CT)	Current measurement with Kirchhooff's law	Identification of the fault zone is possible. Insufficient for diagnosis of apparatus [10].
Digital Fault Recorders (DFRs)	Voltage and current waveforms	Identification of the exact fault location is possible. Useful for diagnosis of apparatus [11].
Sequence of Events Recorders (SERs)	Breaker and relay operation	Identification of the fault zone is possible. Useful for diagnosis of apparatus [8,12].

The level of details that can be analyzed by the fault diagnosis system: fault distance / fault zone based on relay protection zone, and diagnosis of apparatus, and the intelligent system technique used by the fault diagnosis system: Knowledge-Based System [1-3], Abductive Reasoning [4], Model-Based Reasoning [5-6], Neural Network [7], Logic-Based Reasoning [8-9].

More details about the methodologies are included in section 4.

The accuracy of fault diagnosis relies on the *completeness* of event messages received from the sensors or protective devices that may differ from a power system to another. The types of information commonly used for fault diagnosis are described in Table 1. Due to technological improvements in computers and electronics, the newest relays have communication capabilities that allow these devices to transfer their status to a central location. Recording devices such as Digital Fault Recorders (DFRs) and SERs provide detailed waveforms or events that are important for fault analysis. Another important consideration for development of a fault diagnosis system is *verification* and *validation* of the system. At the fault diagnosis system design level, it may be impossible to test all the scenarios that are likely to happen. The method for verification and validation is discussed in section 6 with the fault diagnosis system, Generalized Alarm Analysis Module (GAAM), developed at University of Washington, as an example.

2. PROTECTIVE DEVICES IN FAULT DIAGNOSIS

Fault diagnosis systems are designed to analyze the protection device operations based on their settings and coordination.

Thus, the designers of fault diagnosis systems have to be aware of what types of devices are installed in the target power system. In this section, the common protective devices and corresponding types of faults are discussed. The types and settings of protective devices may vary from a power system to another. The following are the devices of special interest:

- Distance relays
- Differential relays
- Overcurrent relays (OCRs)
- Breaker failure devices (BFDs)
- Bus coupler breakers (BCBs)
- Sequence of events recorders (SERs)

Distance relays. Distance relays are used to protect transmission lines. Generally, a pair of distance relays is needed to protect a transmission line. When a fault occurs on a transmission line, two distance relays at both ends of the line determine the fault impedance first. If the fault impedance exceeds the corresponding threshold setting, the distance relay sends a trip signal to the corresponding breaker immediately. Distance relays have multiple protection zones that provide back up in case of a failure of a primary protection device. The distance relay that detects the fault in the first zone is designed to trip first. The back up capability is achieved by zone-2 or zone-3 settings of the distance relays. The time settings for protection zones vary from a power system to another. To avoid errors for a long transmission line, a distance relay does not cover 100 % of the line protected by the relay. Usually, 80 % of the line is protected by the primary distance relays. When a fault occurs on the remaining 20 % of the line, the relay close to the fault location can isolate the fault in the first zone, while the relay at the other ends is designed to isolate the fault in zone-2. To account for the difference in operating times, two relays can communicate with each other so that they can both trip at the zone-1 time setting. The resulting protection scheme is called '*pilot relaying*'.

Differential relays. A differential relay can be installed for protection of a busbar. If the difference between incoming and outgoing currents at the busbar is not zero, then the differential relay recognizes a busbar fault and trips all breakers surrounding the busbar. A differential relay is also used to protect a transformer against internal faults.

2-3. Over Current Relays (OCRs)

Overcurent relays (OCRs). OCRs provide backup protection for a transformer when the differential relay fails to trip. The fault current is usually greater than the normal load current. When OCRs sense the fault current and the primary differential relay fails to trip, OCRs can trip the breakers on both sides of the transformer. Another application of the OCRs is the protection of transformers against external faults. These external faults occur on other power system components (lines or busbars), causing an over-current through the transformer.

Breaker failure devices (BFDs). For a transmission line fault, as introduced above, distance relays trip the corresponding breakers. However, the breaker(s) may fail to open due to a breaker internal failure such as a trip coil failure or loss of DC trip power supply. BFDs monitor breaker operations following a tripping command from a distance relay. If the breaker that received a tripping command does not open within a given period time, a BFD sends tripping signal to all breakers surrounding the busbar connected to the transmission line to isolate the fault.

Bus coupler breakers (BCBs). A normally-open breaker connecting a pair of busbars is called 'Bus Coupler Breaker'. A change of the BCB's contact status affects the power network configuration. For fault diagnosis systems, it is important to identify the configuration of power networks. BCBs are usually operated by dispatchers.

Sequence-of-Events Recorders (SERs). An SER is a recording device that reports detailed events of the circuit breakers and protective relays with an accuracy level within milliseconds. Examples of the SER messages are:

Date	Time	Sub	Voltage	Sub	Line #	Events
03-01-97	12:01:06:030	TURBIGO	380	FLERO	1	Started Protection Phase A
03-01-97	12:01:06:430	TURBIGO	380	FLERO	1	Tripped Protection Phase A

3. CAPABILITY OF FAULT DIAGNOSIS SYSTEMS

Fault diagnosis systems are designed to analyze the operations of protective devices, breakers, protection schemes, and power network configurations to identify:

- fault location: line, busbar, transformer internal or external faults, or their combinations,
- fault type: single-, double-phase, or three-phase faults.

Fault diagnosis systems are intended to identify the location and type quickly and accurately. As mentioned in section 1, the availability of event/alarm messages can affect the accuracy of a fault diagnosis system. The following requirements must be considered in fault diagnosis systems.

- Timing differences between substation clocks
- Repetitive massages
- Inconsistent information
- Missing information
- Multifunctions of protective devices

Timing differences between substation clocks. Protective devices or recorders at different substations should record the same time instant at the start of an event. However, it is likely that the time stamps of the events recorded at different substations are different due to clock errors. If Global Positioning Systems (GPSs) can send standard time stamps to all substations, then this timing error can be corrected. If GPSs are not available, fault diagnosis systems should have the capability to correct timing errors to the extent possible.

Timing errors can be synchronized based on relay settings and their coordination. For example, if relay-tripping times from two substations that are involved in the same fault are different, then the first relay-starting signal from the substations is taken as a reference time and the relay-tripping signals are synchronized based on the reference time stamp and the relay settings.

Repetitive massages. Event messages recorded by an SER may include repetitive relay operations due to a communication problem. These repetitive relay messages are redundant for fault diagnosis. To avoid confusion, these repetitive event messages should be removed.

Inconsistent information. Consider that a sequence of events as follows: relay-starting – relay-tripping – breaker opening. This logic sequence is identical for protection design even though the specific protective devices may vary from a power system to another. If there is any timing inconsistency in the sequence, the fault diagnosis system should be able to correct it. Most fault diagnosis systems rely on the breaker contact status acquired through the Energy Management System (EMS). The EMS data is usually more reliable and therefore it can serve as a basis for correction of inconsistencies.

Missing information. Due to a communication problem, some event messages can be missing. Fault diagnosis systems should be able to identify the missing information and fill the gap of information. For example, if a 'transition to zone 2' message is missing but all other subsequent messages (e.g., relay starting, tripping, and breaker opening messages) are present, then the missing zone-transition message can be reconstructed based on the known time interval between the relay-starting and tripping events. The time interval between relay operations (e.g., starting, tripping, and so on) varies from a power system to another. Power systems may not be equipped with a communication device or event recorder at every substation. For example, every substation in the Italian 380 kV power network has an SER, while this is not true for the 220 kV network. Thus, fault diagnosis systems need to have an ability to deal with a situation in which information is incomplete. Some results on missing information and missing SERs are reported in [13].

Malfunctions of protective devices. It is possible that protective devices and/or breakers do not operate as designed against a fault event. The general types of malfunctions are as follows:

- relays: wrong tripping zone, failure to trip, mis-operation, improper operating time, wrong direction of fault
- breakers: stuck breaker, failure to trip all three phases.

Malfunctions of these devices can lead to cascading outages, resulting in complex scenarios. Fault diagnosis systems should be able to identify the malfunctioning devices for complex scenarios.

4. THE STATE-OF-THE-ART TECHNOLOGY IN DEVELOPMENTS IN FAULT DIAGNOSIS

Over the last 15 years, much effort has been devoted to the development of fault diagnosis systems. Part of the effort was on the development of *'reasoning systems'* for the analysis of the events/alarms to determine the fault location and type. The state-of-the-art technologies are briefly discussed in this section. Note that only recent references are mentioned in this paper. The following technologies are discussed next:

- Rule-Based Systems (RBSs)
- Model-based systems (MBSs)
- Logic based systems (LBSs)
- Other intelligent system techniques

Rule-based systems (RBSs). A basic RBS contains a number of rules, an inference engine, and the working memory describing the state of the domain system and the status of reasoning. Each rule in the rule base consists of *'if-then'* structures. The inference engine is used to match rules from the working memory during the analysis of an event. There are two types of inference engines: 'forward' and 'backward'. 'Forward' chaining is data driven but 'backward' chaining is goal driven [14]. Rule-based fault diagnosis systems use rules to analyze fault events [1-3]. These rules are either acquired from human experts or developed based on protection logic. Earlier fault diagnosis systems are RBSs. However, constructing and updating the knowledge database may be time-consuming. The rule-based systems are known to be 'brittle' in that it may fail to handle a scenario that has not been encountered by the RBS.

Model-based systems (MBSs). Model-based reasoning is also called qualitative reasoning. The MBS relies on a model of the physical relationships between power system components (e.g., lines, transformers, busbars) and their protective devices. For a given set of alarms or events acquired from the EMS or SER, it is necessary to search the model base in order to identify the cause(s) of these alarms or events. If the model base completely describes the problem domain and the search method is sufficient to cover all possibilities, the MBS will be able to identify the cause(s) of a scenario. This is why the MBS is considered less brittle than RBSs. In practice, however, the model base is likely to be incomplete. For example, some devices may not be modeled and the cause-effect relations may not be covered exhaustively in the model base. Examples of model-based fault diagnosis systems are reported in [5-6].

Logic-based systems (LBSs). LBSs use the fact that protective devices operate to isolate the fault based on the designed protection logic. The result of these operations is a set of event messages received at the control center. Suppose S is the set of fault/malfunction scenarios and R is the set of event/alarm messages. In LBSs, one uses the event messages (R) to identify the fault/malfunction scenarios (S) based on the protection logic. In other words, the event/alarm messages are analyzed by protection logic. Thus, fault diagnosis systems using logic based reasoning can handle fault/malfunction scenarios that cannot be predicted in advance. The GAAM system is a logic based reasoning system.

Multiple Hypothesis Analysis Based on LBS

It may not be possible for a fault diagnosis system to identify the exact fault location and type when the fault event is complex. This can happen when the protective devices malfunction. In this case, a fault diagnosis system can generate multiple hypotheses for the scenario. The GAAM approach to multiple hypothesis analysis is to find the most 'credible' hypothesis. More details of the method are given in section 4.

Other Intelligent system techniques. Neural network and fuzzy logic have been utilized to identify the fault locations/types. Hybrid systems integrating neural networks with expert systems are also proposed [15]. An abductive reasoning technique with logic gates that handle the inherent functional and logical relationships between system components and the corresponding breaker contact status is reported in [4].

5. MULTIPLE HYPOTHESIS ANALYSIS

In this section, more details of the multiple hypothesis analysis are provided. This algorithm is implemented in GAAM. GAAM uses the status data from the EMS and event messages from SERs.

- Step 1: Identifying the de-energized zone

Power systems are protected by relays with settings that are coordinated off-line. If a fault is isolated by a set of open breakers, then the zone surrounded by the open breakers is classified as a de-energized zone. GAAM identifies this de-energized zone based on the SER messages and network topology. For fault analysis, every component in the de-energized zone is a candidate for the faulted component.

- Step 2: Identifying the protection net for each candidate component

Once a de-energized zone is identified, GAAM builds a protection net for every component (e.g., lines, busbars, or transformers) in the de-energized zone. The protection net consists of the set of all relays, breakers and the protection logic designed to protect the component. The protection net is built based on off-line information on the protection schemes, relays settings, and network configurations. For instance, if a line is a candidate for the faulted component, then the set of all relays/breakers that are designed to protect the line from zone 1 to zone 3 (or zones 4-5 if applicable) are included in the protection net.

- Step 3: Reasoning

After the identification of the protection net for every component in the de-energized zone, GAAM generates a set of hypotheses for the given data and messages. Based on protection schemes (e.g., relays settings, protection coordination, pilot relaying, and so on) and the set of SER messages, GAAM generates a set of hypotheses associated with candidate components. Each hypothesis represents a fault/malfunction scenario.

- Step 4: Ranking the hypotheses based on credibility

After step 3, GAAM generates a set of hypotheses representing possible fault/malfunction scenarios for a given set of SER messages. As mentioned in section 4, identification of the most credible hypothesis is a critical issue in fault diagnosis. GAAM ranks the multiple hypotheses based on their credibility. GAAM determines an index value for every hypothesis to indicate its level of credibility based on evidences (i.e., protection schemes and a set of SER / EMS messages). In validation, it is necessary that the *true* scenario be contained in the multiple hypotheses of fault/malfunction scenarios. GAAM has been tested with a number of actual fault scenarios but GAAM is also validated/verified with a rigorous method that is designed to cover the entire space of scenarios.

6. VERIFICATION AND VALIDATION

Validation and verification are important tasks during the development of software systems. A number of methodologies have been developed for maintenance of expert systems, e.g., relation checking algorithms to detect redundancies or conflicts in rule based systems [16], Petri Nets [17], equivalence class method for verification of rule-based systems [14].

A new contextual representation method, which used basic '*contexts*' to represent the generic types of domain scenarios, is discussed in this paper. The method is implemented in GAAM. The contexts represent the logic embedded in the design of power system protective relaying. *The concept of generic contexts is a powerful tool to avoid exhaustive testing of all possible scenarios during validation and verification.*

The following are the main components of the method:

- Contexts with SER
- Identification of a minimal set of test scenarios with SER

Contexts with SER. A context includes a combination of power system components such as lines, busbars, transformers, and SERs. One assumption in this context representation method is that all primary and/or back up devices within the designed protection net of a component are able to clear any fault on the component. In addition, the event recording devices (i.e., SERs) are able to transfer operations of protective devices to control centers. Within the protection net of a faulted line, the line together with the relays, breakers, and SERs located on both ends can be represented by an object (i.e., L1) in Figure 1. With the same concept, a busbar (i.e., B1) and transformer (i.e., T1) can be represented by objects in Figure 2 and 3. In the figures, 'A' and 'B' are the two ends of a line, 'Line', and a transformer, 'Transformer.' 'Rly_A' and 'Bkr_A' are the relay and breaker on the 'A' side of the line or transformer. 'SER_A' is the Sequence of Events Recorder located on side 'A'. The symbolic values for a line, busbar and transformer can be 'faulted' or 'not-faulted'. The symbolic values for an SER represent completeness of information and the values for protective devices are shown in Table 2. The combination set of values of the variables in a context can describe the fault and malfunction scenario. In addition to the number of objects, a context includes their associated symbolic variables with their values representing the configuration of a portion (e.g., protection network) of the power network.

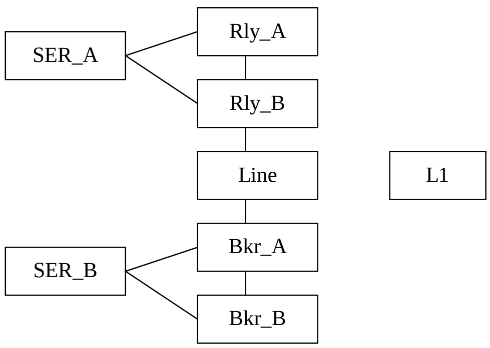


Figure 5.1. An object representing a transmission line

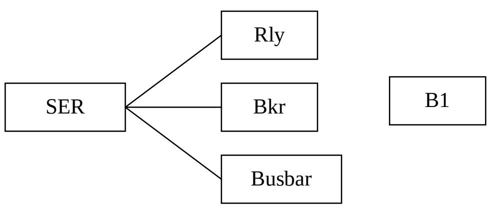


Figure 5.2. An object representing a busbar

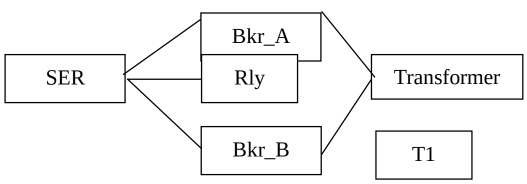
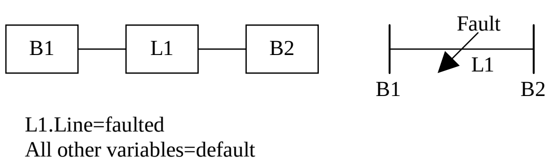


Figure 5.3. An object representing a transformer

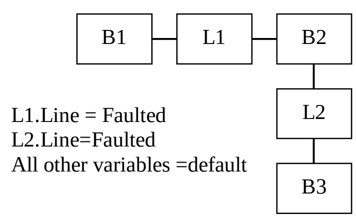
Figure 5.4-(a) depicts a contextual representation of the configuration in (b). In Figure 4.4, the operation '.' in the notation 'A.B' implies that variable B belongs to object A.



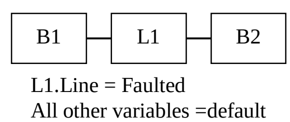
(a) (b)

Figure 5.4. A context and represented network

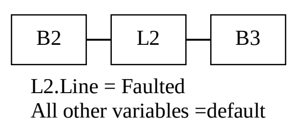
Identification of a minimal set of test scenarios with SER. For verification purpose, if two fault/malfunction scenarios use the same knowledge in the LBS, then these two scenarios are said to be '*equivalent*'. Thus, testing on one of the two scenarios is sufficient to ensure that the LBS provides accurate results for both scenarios. In fact, the common features shared by the two fault/malfunction scenarios were utilized in the development of the logic formulas. Identification of equivalent fault/malfunction scenarios for LBS is used for the development of the contextual representation. Any context can be decomposed into loop-less paths linking a protected component and a set of protective devices. These contexts are called *basic* contexts and a complete set of basic contexts is called a *basis*. The term '*complete'* means that any context can be represented by the basic contexts in the basis. It is possible to count all basic contexts for various configurations of a power system because the reach of the protective zones is limited to a small number of lines or substations. Figure 5.5 represents a context and its basic contexts. In Figure 5.5-(a), the context represents two multiple faults on line 'L1' and 'L2'. This context can be decomposed into loop-less paths as (b) and (c) representing basic contexts. However, the basic contexts in (b) and (c) are the same except for the name of the busbar and line. In other words, the same knowledge will be used for diagnosis of (b) and (c). Therefore, testing on context (a) is not necessary if GAAM is tested with the basic context (b) or (c). To identify the minimal set of test scenarios, the context (a) can be removed from the set of test scenarios. In this section, the method for the validation and verification of GAAM is discussed. Not all symbolic values are listed in Table 5-III. More detailed symbolic values and explanations can be found in [13,18].



(a) A context representing multiple faults on L1 and L2



(b) A basic context representing a fault on L1



(c) A basic context representing a fault on L2

Figure 5.5. A context and its basic contexts

7. A CASE STUDY OF ENEL

In this section, two test cases are used to demonstrate the capability of the fault diagnosis system, GAAM. The first fault/malfunction scenario illustrates GAAM's capability to analyze malfunctions of the protective devices. The second scenario shows how a fault/malfunction scenario can be represented by the contextual representation method. Both scenarios are based on the Italian power system. For validation and verification, a simulator was developed to generate EMS and SER messages resulting from a fault or malfunctioning scenario, which may also involve missing or repetitive messages specified by the user [19]. Now the discussion concentrates on:

- Case Scenario 1
- Case Scenario 2

Case Scenario 1. In Figure 5.6, the scenario includes a fault on the line between NAVE and FLERO with malfunctioning relays at TRAVAGLIANTO and OSTIGLIA respectively. The fault was isolated by the primary relays at NAVE and FLERO but the relays at TRAVAGLIANTO and OSTIGLIA recognized the fault in the first protection zone and tripped. This malfunction could be caused by improper relay settings or coordination. There are four candidate fault components: NAVE - FLERO, FLERO - OSTIGLIA, TRAVAGLIANTO - FLERO, and the busbar at FLERO. In other words, GAAM provides four hypotheses but only the first ranked hypothesis, which is the true fault/malfunction scenario, is shown here.

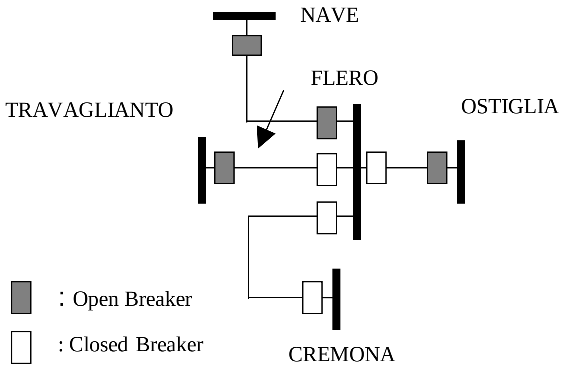


Figure 5.6. Test scenario 1

Hypothesis No. 1 (True Scenario)

Line No. 1 between FLERO-NAVE
Status = Permanent fault on Phase A
The malfunctioning relays / breakers are:
Impedance relay at TRAVAGLIANTO (line 1 between TRAVAGLIANTO-FLERO)

Status= malfunctioned. Briefly, the relay:
1: Tripped in wrong zone. The relay should not have tripped in the first zone.
Impedance relay at bus OSTIGLIA (line 1 between OSTIGLIA-FLERO)
Status=malfunctioned. Briefly, the relay:
1. Tripped in wrong zone. The relay should not have tripped in the first zone.

Case Scenario 2. This test scenario shows how a fault/malfunction scenario is represented by a context. In Figure 5.7-(a), a fault occurred on the line 366 between TRAVAGLIANTO-FLERO and the relay at FLERO failed to trip so the backup relays at OSTIGLIA, NAVE, and CREMONA tripped in the second zone to isolate the fault. The first ranked hypothesis provided by GAAM is shown and it is also the true scenario.

Hypothesis No. 1 (True Scenario)

Line No. 1 between FLERO-TRAVAGLIANTO
Status = Permanent fault on Phase B
The malfunctioning relays / breakers are:
Impedance relay at FLERO (line 1 between FLERO-TRAVAGLIANTO)
Statue=malfunctioned, Briefly, the relay:
1: Failed to trip

Since the two basic contexts are the same, if GAAM is tested with one of the basic contexts then the test on the context in (b) is not necessary. The symbolic values for the variables in the basic context are as follows:

L1.Line=faulted
L1.Relay=failed to trip

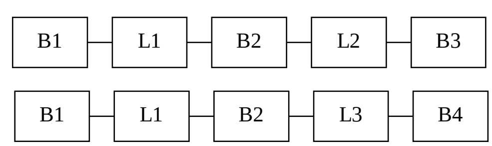
All other variables are default

8. CONCLUSION

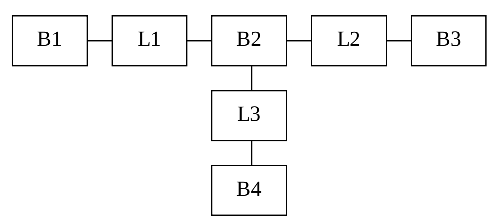
The capability of a fault diagnosis system depends on the data available for the task. The SER information provides valuable time tags for the events that are important for reconstruction of the fault scenario. GAAM is able to pinpoint the exact fault location/type if the information is complete and accurate. When information is missing or erroneous, the most credible hypotheses are determined by GAAM. SERs are not available at every substation. Utilities also use DFRs or other devices. Computer relays contain event information that may be able to replace the recorder messages. This has not been integrated in the current fault diagnosis systems. Current EMSs do not include SER or DFR information. However, as distributed computer systems become available for the power system operational environment, a fault diagnosis system can access event messages through the computer network.

Figure 5.7-(b) represents the context of the true fault / malfunction scenario in (a). The context in (b) can be decomposed into two basic contexts in (c).

(a) Test scenario



(b) A context for the scenario



(c) Two basic contexts for the context

Figure 5.7. Test scenario 2

9. ACKNOWLEDGEMENT

This research was sponsored by ENEL, the Italian Power Company and EPRI. The leadership and contributions from Dr. M. Sforna, ENEL, have been essential for this project. The authors would like to thank Dr. Dejan Sobajic and Dr. David Becker, EPRI, and Dr. Dario Lucarella, ENEL, for their guidance.

10. REFERENCES

[1] J.R. McDonald, et al., "Alarm Processing and Fault Diagnosis Using Knowledge Based Systems for Transmission and Distribution Network Control," *IEEE Trans. Power Systems*, Aug. 1992, pp. 1292-1298.

[2] Z.A. Vale, et al., "An Expert System with Temporal Reasoning for Alarm Processing in Power System Control Centers," *IEEE Trans. Power Systems*, Aug. 1993, pp. 1307-1314.

[3] C.Y. Teo, et al., "Conventional and Knowledge Based Approaches in Fault Diagnosis and Supply Restoration for Power Network," *IEEE Trans. Power Systems*, Feb. 1998, pp. 8-14.

[4] T.S. Sidhu, et al., "An Abductive Inference Technique for Fault Diagnosis in Electrical Power Transmission Networks," *IEEE Trans. Power Delivery*, Jan. 1997, pp. 515-522.

[5] C.C. Liu, et al., "Model-Based Synthesis and Suppression of Transformer Alarms in a Control Center Environment," *IEEE Trans. Power Delivery*, Jul. 1998, pp. 843-848.

[6] S.D.J. McArthur, et al., "The Application of Model Based Reasoning within a Decision Support System for Protection Engineers," IEEE Trans. Power Delivery, Oct. 1996, pp. 1748-1754.

[7] A.P. Alves da Silva, et al., "Neural Networks for Fault Location in Substations," IEEE Trans. Power Delivery, Jan. 1996, pp. 234-239.

[8] H. Miao, et al., "A New Logic-Based Alarm Analyzer for On-Line Operational Environment," *IEEE Trans. Power Systems*, Aug. 1996, pp. 1600-1606.

[9] Y.M. Park, et al., "A Logic Based Expert System (LBES) for Fault Diagnosis of Fault Diagnosis of Power System," *IEEE Trans. Power Systems*, Feb. 1997, pp. 363-369.

SECTION VI

IMPLEMENTATION USING SCADA SYSTEMS

Prof. J.R McDonald **Dr S.D.J. McArthur**
Centre for Electrical Power Engineering
University of Strathclyde
Glasgow, Scotland, U.K.

Abstract. This section discusses the issues, solutions and technologies underpinning the development of an intelligent decision support system (DDS) for protection performance analysis. This DDS automatically retrieves data from the SCADA (Supervisory Control and Data Acquisition) system, fault recorders and fault locators, interprets it and provides summarized information. It then discusses, in detail, the intelligent system technologies employed to offer engineering decision support functions as a basis for an advanced suite of data interpretation tools for analysis of power system disturbances.

1. INTRODUCTION TO PROBLEM

Over the past five years ScottishPower, in conjunction with manufacturers and universities, has been developing systems which facilitate the analysis of transient conditions. These systems are designed to provide improved data, in the first instance, for protection and control engineers. They include:

- Fault recorders: Over 70 fault recorders are installed on the transmission system and 120 on the distribution network. These are polled daily via modem links.
- Travelling wave fault locators: These utilize satellite technology to accurately determine the distance to a fault. Travelling wave fault locators presently cover all of the 400kV and 275kV networks, with some being installed on the 132kV network.
- System monitoring equipment: System monitoring equipment is installed which records power flows, system transients, frequency content and system damping.
- Modern protection relays: Modern protection relays are able to store data about the performance of the relays during fault conditions. This data can be downloaded through dial-up modem facilities.

All of the above systems are standalone and operate through their own dedicated user interfaces. Although each system can provide an insight into power system disturbances, the collation and analysis of the data provided by such systems is problematic. This problem is exacerbated during adverse weather conditions where the volume of data gathered increases dramatically. For example, on a day with fifty faults 2000 transmission fault records and 5000 distribution fault records can be triggered, due to voltage depressions and protection operations. Only 100 of these may be of interest. The following issues are discussed next:

- Data analysis problems
- The solution strategy

Data Analysis problems. As has been mentioned, the first problem encountered is access to the data capture systems. The engineers must use a number of different systems to collate all the data they are interested in. This is time consuming as much of the data is accessed via modem dial-up.

Once the data has been retrieved the second problem arises. The manual analysis and interpretation of the data is also time consuming. Coupled with the large data volumes gathered during storms and major disturbances, this task becomes both extremely time intensive and complicated.

Presently, the majority of the data capture systems are housed within the “Replay Facility” at ScottishPower’s corporate headquarters. This facility houses the new and developmental data capture systems which ScottishPower engineers use. As a result, only a minority of engineers who could make use of these systems have access to them. Nevertheless, there are a number of other engineers within the regional groups and other functional groups who could capitalize on the data available.

The solution strategy. ScottishPower, in collaboration with the Centre for Electrical Power Engineering (CEPE), has recognized the requirement to co-ordinate the data gathering process and to automatically extract the pertinent **information** provided by these systems, which subsequently can be delivered to relevant staff at the appropriate engineering and management level. This will be achieved through the use of improved communications and data gathering in conjunction with **intelligent** data interpretation systems.

There are some key aspects to the strategy which have been identified :

1. Data interpretation is required to convert data into meaningful **information**.
2. Different data capture systems must be integrated to collate data and allow for the best possible interpretation and extraction of information.
3. Automatic retrieval of data and automatic interpretation is required to remove the burden from engineers.
4. The information derived must be available to all those engineers who require it.
5. Different engineers require different information to be extracted from data sets.

Through the investigation of these requirements key developmental tasks/issues have been identified. These are:

1. The development of data interpretation systems.

2. Systems integration.
3. Data and information storage.
4. Corporate communications issues:

   a) On-line data such as SCADA system data, fault records, etc.
   b) "Corporate" information such as the output from data interpretation systems, e-mail, etc.

These are the key technical aspects behind the provision of an advanced suite of data interpretation facilities for the analysis of power system disturbances.

The conceptual architecture behind the strategy discussed is shown in figure 6.1.

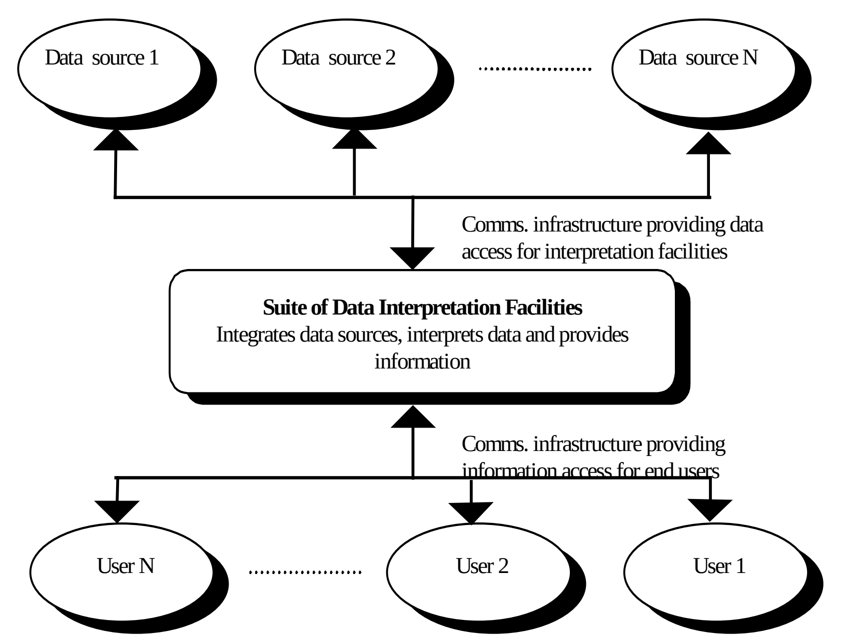


Figure 6.1 - Conceptual Architecture

In a drive towards this vision of comprehensive data interpretation significant research effort by CEPE has resulted in the development of an intelligent Decision Support System (DSS). Its function is to interpret data from a number of sources to produce comprehensive fault reports. This system extracts pertinent information which the protection engineers require. The research and development activity demonstrated that the technologies and techniques exist to allow widespread data interpretation to be achieved.

2. INTELLIGENT ANALYSIS OF PROTECTION PERFORMANCE

The DSS is designed to aid protection engineers, in the first instance, by interpreting the different data sources which are available to them. It provides on-line protection performance analysis following disturbances. The following DDS functions are discussed next:

- Functions of the decision support system
- Integration of fault recorder and fault locator system
- Information provided through Integration and Interpretation

Functions of the decision support system. The DSS interprets SCADA system alarms and fault records. It is comprised of three modules:

- APEX (Alarm Processing EXpert System) - This system summarizes SCADA system alarms and details the key events which are taking place. APEX is a knowledge based system. The knowledge within it has been elicited from domain experts through a structured knowledge analysis program [1-3].

- RESPONDD - This fault diagnostic expert system uses SCADA system data to determine the underlying faults and problems during a power system disturbance. RESPONDD can indicate possible protection failures and backup activity which has occurred as a result. This is a knowledge based system which also uses qualitative models of protection [1-3].

- Model Based Diagnosis (MBD) module - This uses model based reasoning to analyze fault records. Detailed models of protection behavior are used in conjunction with fault recorder data to validate and diagnose the protection operations. This approach can pick out protection failures, slow operations and false protection activity [1-3].

All three systems are required in order to provide a comprehensive analysis of protection performance. Although APEX and RESPONDD effectively interpret SCADA system data to provide meaningful information, there are limitations to the amount of information which can be extracted from SCADA system alarms. An alarm indicating that a protection relay has operated does not imply anything about whether the protection scheme should have operated. Also, if only the first main protection operated and the second main protection did not, then this calls the protection performance into question. Fault records are required to perform a detailed analysis of the protection activity. This is the task of the MBD module.

Integration of fault recorder and fault locator systems. As has been discussed, the DSS needs to access the data available from fault records. For the system to operate efficiently then this process should be automated.

A fault recorder control module has been designed which uses the conclusions generated by APEX and RESPONDD to decide which fault records must be retrieved. The fault recorders used have a "replay system" within ScottishPower's headquarters. The replay system controls the dial-up access of fault recorders (via modem) while providing viewing, analysis and printing capabilities.

The new module creates a control signal for the fault recorder replay system. This signal indicates the times and locations of fault records which must be retrieved. The manufacturer extended the replay system to accept this control signal and automatically dial-up and retrieve the fault records produced at the time and location intimated. This removes the burden of retrieving the fault records from the engineers.

The triggers for this process are configurable. For example, they may be switchout of lines and/or protection operations. Also, "fault recorder operated" alarms may be used.

Once the fault records have been retrieved the MBD module can extract the data it requires from the records using the standardized COMTRADE format. This makes the MBD system more open. It is able to process fault records from any fault recorder which supports this standard.

The same mechanism as above is used to automate access to the fault location system. ScottishPower makes use of travelling wave fault locators. As for the fault recorders, there is a replay system which is used to retrieve the distance to fault information. Retrieval of this information was performed manually by the engineers, but now it has been automated in the same way as the fault record retrieval.

The DSS can determine which fault location records are required based on the occurrence of events such as feeder switchout. A fault locator control module has been designed to send a control signal to the fault locator replay system. This works in an identical fashion to the control module for fault recorders. Once more the manufacturer has engineered their system to accept this signal.

Figure 6.2 is an example of the possible data interpretation performed by the DSS. It shows a portion of a SCADA system alarm stream and the conclusions generated by APEX. It then shows the control signals created by the fault locator and fault recorder control modules, which indicate the date, time and location of the records sought.

The system integration architecture to control the fault locator and fault recorder access is shown in figure 6.3.

SCADA Alarms

SRHN KETH SECOND MAIN PROT OPTD	ON	17:25:05.37
SRHN KETH FIRST MAIN PROT OPTD	ON	17:25:05.38
SRHN CB1 OPEN CLOSED	OPEN	17:25:05.41
SRHN KETH SECOND MAIN PROT OPTD	OFF	17:25:05.43
KETH SRHN FIRST MAIN PROT OPTD	ON	17:25:05.43
KETH SRHN SECOND MAIN PROT OPTD	ON	17:25:05.43
SRHN KETH FIRST MAIN PROT OPTD	OFF	17:25:05.49
KETH CB1 OPEN CLOSED	OPEN	17:25:05.50
KETH SRHN SECOND MAIN PROT OPTD	OFF	17:25:05.51
KETH CB2 OPEN CLOSED	OPEN	17:25:05.52
KETH SRHN FIRST MAIN PROT OPTD	OFF	17:25:05.55

APEX Conclusions

17:25:05.37
SUCCESSFUL PROTECTION OPERATION AT SRHN (KETH FDR)

17:25:05.41
SWITCHOUT OF (KETH SGT2, SRHN (KETH FDR))

17:25:05.43
SUCCESSFUL PROTECTION OPERATION KETH (SRHN FDR)

Fault Recorder Control Signal

17:25:05.41 14NOV1997 KETH SHRN FEEDER 1

Fault Locator Control Signal

17:25:05.41 14NOV1997 KETH SHRN FEEDER 1

Figure 6.2 -Data Interpretation Example

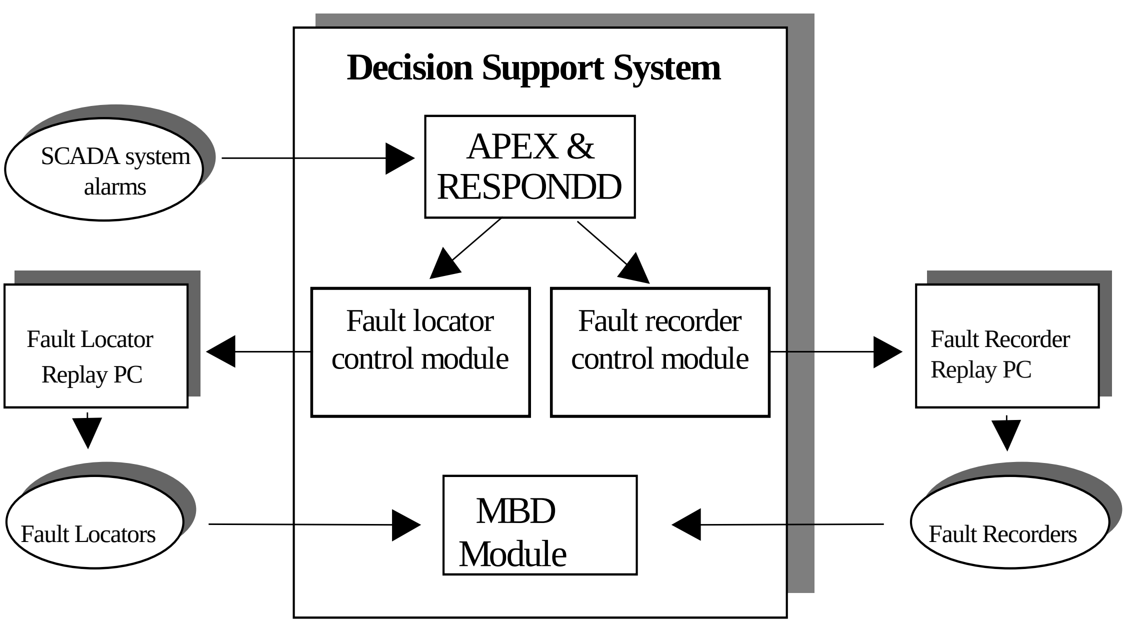


Figure 6.3 - System Integration

Information provided through Integration and interpretation. The integration of the DSS with the fault recorder and the fault locator systems allows relevant information to be provided to ScottishPower's engineers. The result of data interpretation by APEX, RESPONDD and the model based module should, when combined, produces a single report which provides the overall analysis of the protection activity. Importantly, it will highlight to the protection engineer if any further analysis is required.

An example fault report is shown in figure 6.4.

There was a transient fault on the KETH to SRHN feeder. All the relevant SCADA system alarms were received. The distance to the fault was 87.5km from KETH. The fault was cleared, however the first main protection at SRHN operated slowly and should be investigated

Figure 6.4 - Fault Report from the Integrated DSS

This represents an easily assimilated report providing information to the engineers. By providing such reports a priority list of actions/investigations can easily be assembled, reducing the burden on engineers during extreme operating conditions.

So far the structure and design of the integrated DSS has been described. To complete the discussion the utilization and access of this information must be considered.

3. UTILIZATION OF THE DSS

A number of users exist for the information produced by the DSS. The first set of users are the protection engineers. As these engineers are responsible for analyzing SCADA system data, fault records, fault location records, etc. following a disturbance then they will benefit from the decision support being offered. It reduces their manual data analysis task and pinpoints problems quicker. This represents a time saving, especially following a storm where the basic data analysis task can be ongoing over several weeks.

To alleviate some of the difficulties associated with the manual analysis ScottishPower developed an historical alarm database. This stores all the SCADA system alarms received. The database has a facility which permits searching for alarms by date, time, substation, type or any combination of these. When the appropriate alarms have been retrieved from the database they can be "replayed" through the DSS. The selected alarms are automatically sent, via the local network, to the DSS for interpretation. Such a capability is useful for a number of reasons:

- Analysis of historical events, as and when required.
- Training on the use of the DSS.
- Testing of the DSS, following knowledge base or software updates.

The information generated by the DSS would be particularly useful for engineers in the regional groups As they are responsible for the maintenance and upkeep of protection, quicker indication of problems would lead to reduced repair times.

Further users of the output from the DSS could be the engineers with responsibility for quality of supply issues. The DSS can provide detailed information regarding the nature and duration of faults.

This leads to a possible further development of the DSS. It could be configured to automatically generate fault statistics. Hence, removing the time consuming task of manual compilation of such figures.

Given the variety of possible users of the DSS, the facilities which provide access to the information are important.

A dedicated user interface exists for the decision support system. This is within ScottishPower's "replay facility", a centralized location for all the data capture systems available. This obviously restricts utilization of the DSS as access is restricted to a single location.

In order that the DSS can be used by the appropriate personnel at different locations, a standalone PC based user interface has been designed and implemented.

This remote user interface provides access to the conclusions generated by the DSS (i.e. conclusions produced by APEX, RESPONDD and the MBD module) across ScottishPower's corporate network. It also provides the ability to search the conclusions for particular events. The interface is fully configurable by the user, and also allows the engineers to attach notes to the reports, in order that they can add further explanation and description about the disturbance.

The DSS, as has been described, sends control signals to the fault recorder and fault locator replay systems. In addition, it receives on-line SCADA system data plus data from the historical alarm database. These communication capabilities are achieved via ScottishPower's corporate network. This network also facilitates access to the conclusions, generated by the DSS, through the remote standalone user interface.

The use of the corporate network has allowed a further function to be added to the DSS. When it determines that a significant event has taken place it will use electronic mail to inform the appropriate engineers that the diagnostic process has commenced. Thus alerting the relevant personnel to problems as they occur.

The architecture within which the DSS resides is shown in Figure 6.5.

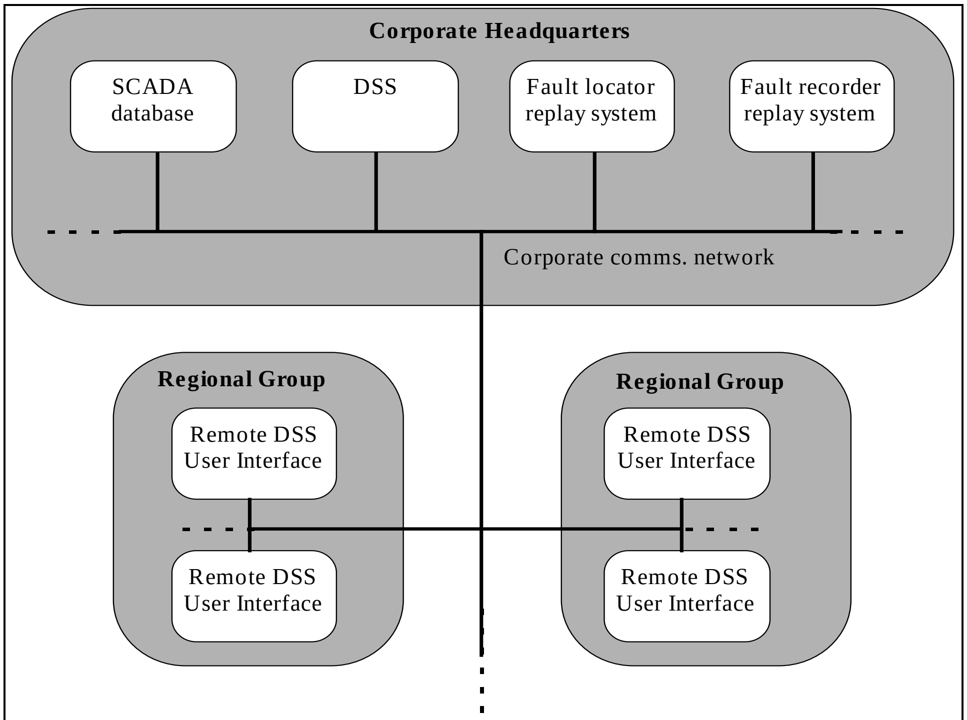


Figure 6.5 - DSS Environment

4. A CASE STUDY OF SCOTTISH POWER

This first section of the tutorial demonstrates the techniques and technologies which were applied in order to achieve automatic data interpretation for ScottishPower.

This system fulfills some of the key points of the strategy presented in section 3:

- The DSS interprets data and converts it into meaningful information.
- The DSS integrates with fault recorders and fault locators to enhance the information it can provide engineers with.
- The fault records and fault locator records are automatically retrieved by the DSS.
- Information is made available to all those who require it via the corporate network.

This research and development activity has been significant in that it proves that the strategy identified can be achieved. Intelligent system technologies are now mature and can be implemented within an industrial environment.

The DSS which has been developed is the first step towards a suite of advanced data interpretation facilities. It demonstrates the maturity of intelligent system technology and the capability to offer effective data interpretation. Nevertheless, a large collection of data sources (e.g. system monitors, modern microprocessor relays, plant condition monitoring equipment, metering equipment) which must be manually analyzed still exist. The necessary interpretation facilities must be developed and provided for these data capturing systems.

Global and open access to data and information is required. This must be available to all relevant ScottishPower personnel at all sites. Therefore, the logical approach is to utilize the corporate network, which interconnects all locations, as a channel for all data and information. The data interpretation systems would have access to this network which could provide them with data and make their interpreted information accessible (as in figure 6.5).

One key aspect is that there are functions which require access to data with minimum time delay. For example, alarm processing systems such as APEX. Hence, fast access channels or priority routes may need to be provided.

For data and information to be accessible to all, when it is required, it must be archived within some repository. Therefore a corporate database is required. In reality this will be an integrated set of distributed databases, which can only be achieved through a corporate communications network. The following system modules are discussed next:

- Alarm processing module
- Fault diagnostic module
- Model based reasoning module

Alarm processing module. The alarm processing expert system (APEX) interprets and summarizes SCADA system alarms and details the key events which are taking place using a knowledge based approach. The knowledge within it has been elicited from domain experts through a structured knowledge analysis program. Figure 6.6 (a) shows a process model describing the reasoning process within the alarm processor. This demonstrates the hypothesis, generate and test approach adopted [4]. The boxes indicate actual knowledge sources used at each stage and the ellipses indicate the reasoning, or inference, activity occurring. Also indicated are the interactions with the domain layer knowledge. From the domain layer the rulebase is used to input rules to the appropriate inference steps. In addition, a domain layer model of the physical electrical network is used at various stages. Figure 6.6 (b) schematically demonstrates how the alarm processor software functions. It was built using the "C" programming language to allow complete control and optimization of the alarm processing function.

The next sections describe the operation shown in Figure 6.6 and also describe the use of the domain knowledge (rules and model of the physical network).

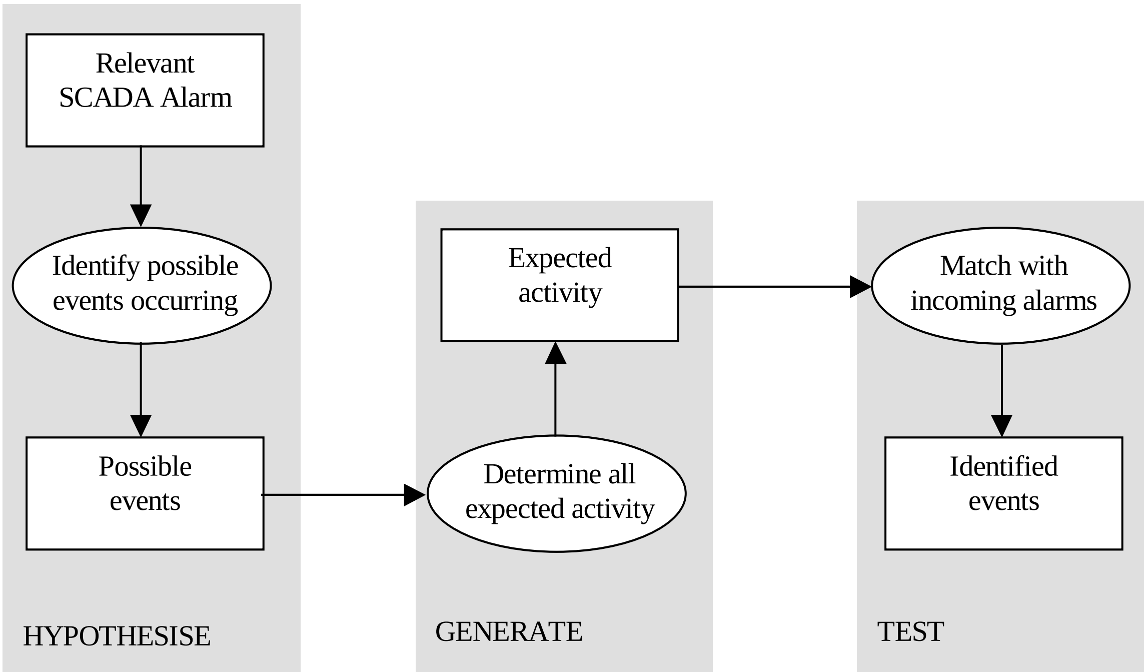


Figure 6.6 (a): Process Model of the KBS based Alarm Processor

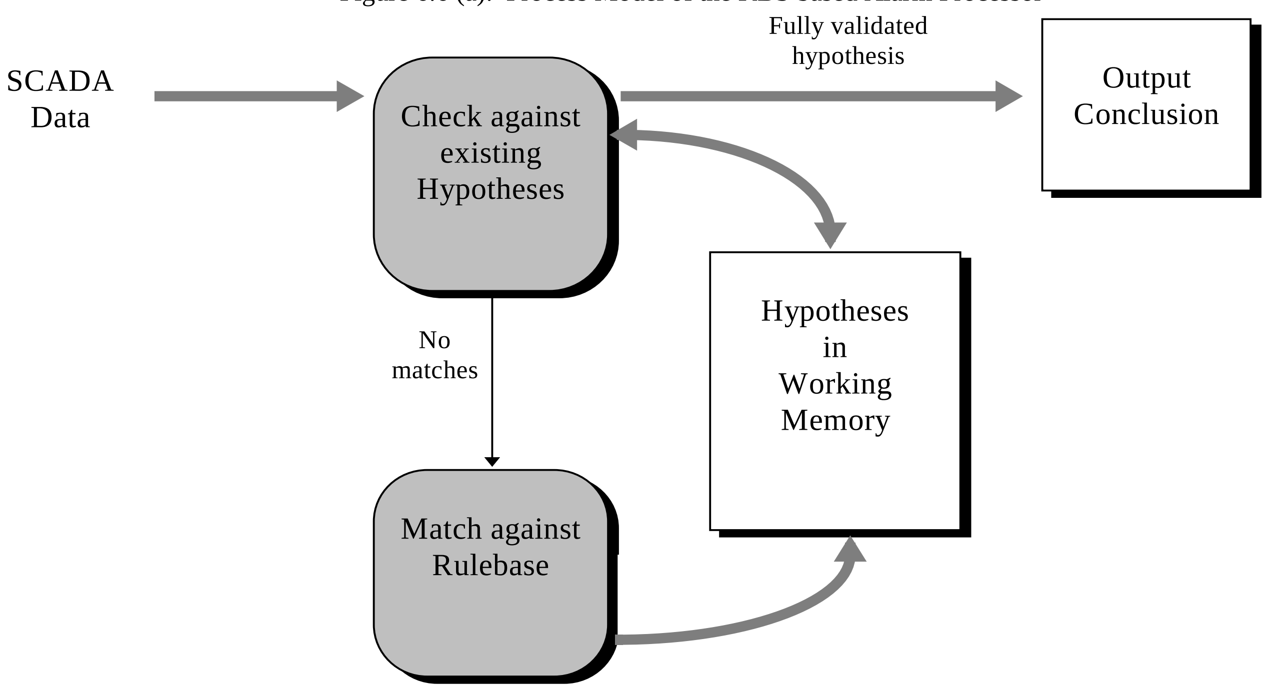


Figure 6.6(b): Reasoning Mechanism Underlying the KBS based Alarm Processor

Knowledge Capture, Validation and Representation

A thorough process of knowledge capture and validation was undertaken, as part of the research, which determined the rules implemented within the alarm processor. This involved three utility experts over a period of approximately eighteen months. A series of knowledge elicitation meetings were held where case studies of faults within the power system were discussed. These were audio recorded and transcripts of the discussions were produced and validated by the experts. From these transcripts, rules were constructed. The rules were implemented and the alarms from the case studies were then passed through the alarm processor for the experts to assess its operation. In this way, the complete rule base was constructed.

The rule base is designed to match events generically across the electrical network. That is, there is no requirement for a rule per event per physical location within the electrical network. Instead, each rule is independent of the location where the event might occur, and can introduce the relevant details dynamically. Furthermore, it uses a model of the electrical network to perform searches which indicate items of electrical plant and areas of the electrical network which have been taken off supply.

Three example rules are shown in Figure 6.7. The first indicates the rule for determining when an area of the network is “switched out”, or off supply (i.e. no electricity is flowing). The rule shows that for this to be the event which occurred it expects the alarm “OPEN CLOSED” to indicate the state OPEN for a “<CBSet>”. A “<CBSet>” is any set of circuit breakers (switching devices) which may open. When a circuit breaker opens it interrupts current flow. Once one of these opens, the alarm processor uses its model of the physical network to determine which others it expects to open (i.e. to fully isolate the faulted section of the power system). This defines the “<CBSet>”. Once they all open, then the conclusion given in the first line of the rule is output; “Switchout of <BlackOut>”. The identifier “<BlackOut>” leads to the use of a network model to determine the actual items of plant bounded by the <CBSet>, and output these within the conclusion. This rule demonstrates the generic nature of the rule base, and the method by which it can be applied throughout the network. The second rule works in an identical fashion, except it is triggered by circuit breakers closing and indicates the reconnection of the electricity supply.

The third rule pertains to protection device operation. This time the “<StationName>” identifier is used. Therefore if any of the alarms e.g. "FIRST MAIN PROT OPTD" with status ON, is received then “<StationName>” is replaced with the actual station that the alarm was received from. The expected alarms are then completed using this station identifier. When all are received, the conclusion is output. Once more, the form of the rule is seen to be generic in nature.

This high level discussion of how the rules operate misses some important aspects. First, to allow for traditional data problems associated with SCADA system data a complete set of alarms need not be received before the conclusion is generated i.e. the alarm processor copes with incomplete/missing data via a point scoring mechanism. Furthermore, the alarms can be received in any order. The generic and near natural language format of the rule base is intended to aid maintainability of the knowledge base.

Generic Applicability of the Alarm Processor Module

While the given application and example consider the analysis of alarms from an electrical power system, the core generate, hypothesis and test approach of the reasoning mechanism could be applied in other domains. By removing the use of the domain model of the physical electrical network and focusing upon the rulebase alone, the alarm processor can interpret alarms within any domain. The core reasoning engine has been proven to be able to interpret in the order of 30,000 alarms per minute.

Fault diagnostic module. The fault diagnostic module RESPONDD is a knowledge based system which uses SCADA system data to determine the underlying faults and problems during a power system disturbance. It can indicate possible protection device failures and backup protection and switchgear activity which has occurred as a result. Qualitative models of protection device behavior are used within the diagnostic process [5].

The reasoning process model describing the operation of this module is shown in Figure 6.8. It can be seen that the reasoning approach follows a similar generate, hypothesis

Event "Switchout of <BlackOut> "
Expect
{
Alarm " OPEN CLOSED" OPEN <CBSet>
}

Event "Restoration of <BlackOut> "
Expect
{
Alarm " OPEN CLOSED" CLOSED <CBSet>
}

Event "Protection operation at <StationName>"
Expect
{
Alarm "FIRST MAIN PROT OPTD" ON <StationName>
Alarm "FIRST MAIN PROT OPTD" OFF <StationName>
Alarm "SECOND MAIN PROT OPTD" ON <StationName>
Alarm "SECOND MAIN PROT OPTD" OFF <StationName>
}

Figure 6.7 : Example Rules from the Alarm Processor

and test paradigm to that used by the alarm processor. However, unlike the alarm processor the key input from the domain layer models comes from qualitative models of power system protection devices. These model unit and non-unit protection devices along with the associated trip relay and switchgear. Although there are well-known equations which can be used to model the operation of these protection devices, which use the current and voltages measured at the protection relay, it is the *qualitative* nature of the functionality which is modeled. In the first instance, these models are used when a protection device operates in order to identify the possible locations of a fault. Later in the diagnostic process they are used to simulate the expected behavior for all the possible fault scenarios.

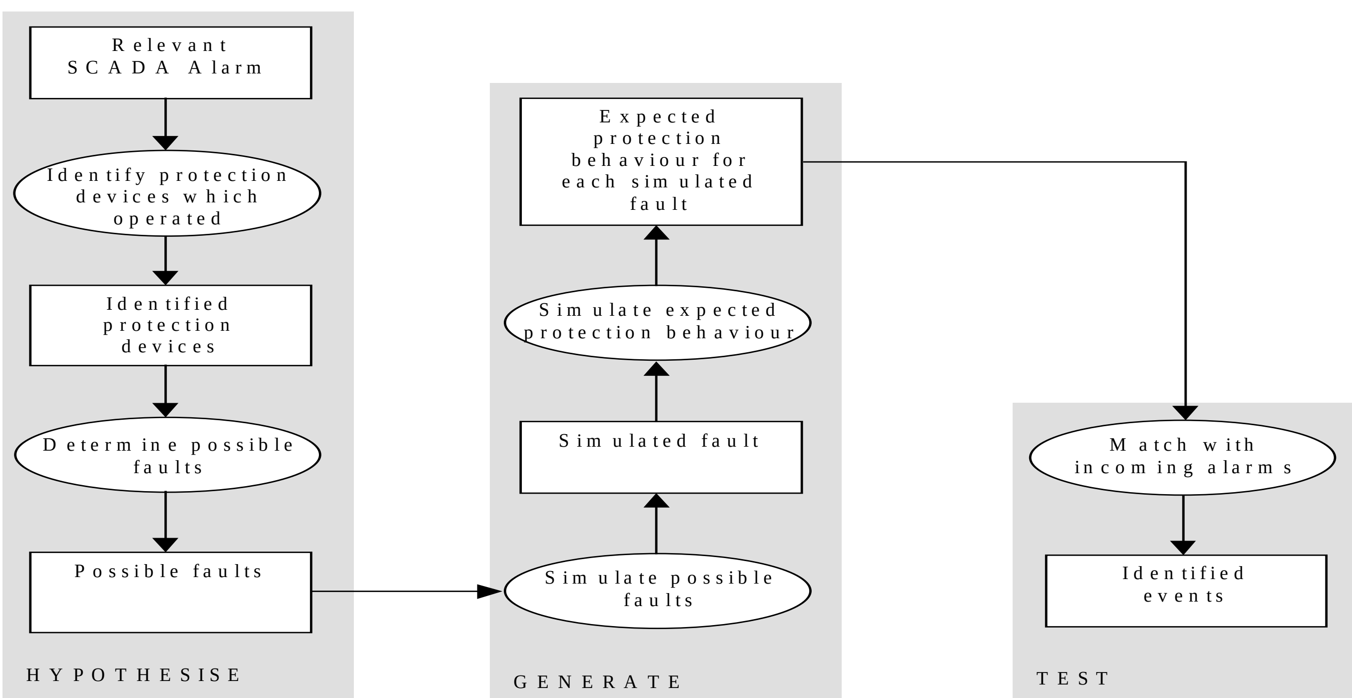


Figure 6.8: KADS Inference Layer Model for Fault Diagnosis

The complete reasoning paradigm is more fully discussed in a number of publications [1][5]. This intelligent module is coded in Prolog, which supported the form of diagnostic reasoning required. As it has a more detailed knowledge and models of the power system than the alarm processor, which allows a greater depth of reasoning.

Knowledge Capture, Validation and Representation

The process of knowledge elicitation undertaken for the alarm processor also supported the design of the fault diagnostic knowledge base. However, the key aspect of this knowledge base is its qualitative models of power system protection. Some forms of power system protection operate when they detect any high levels of current due to a fault. Others require the fault to be determined as being in a certain "zone" of the network before they operate. Also, some forms of protection use direct communication with other protection devices, for fault information, as part of their algorithm to determine whether to operate or not. All of the characteristics of the varying protection devices are modeled. The models are then used as the basis of the hypothesis, generate and test mechanism. Probable faults are simulated and the simulation results are compared with the actual data received.

Model based reasoning module. The model based reasoning module is controlled and triggered by the output from the alarm interpretation modules, as described in the first section of this tutorial and demonstrated graphically in Figure 5.9.

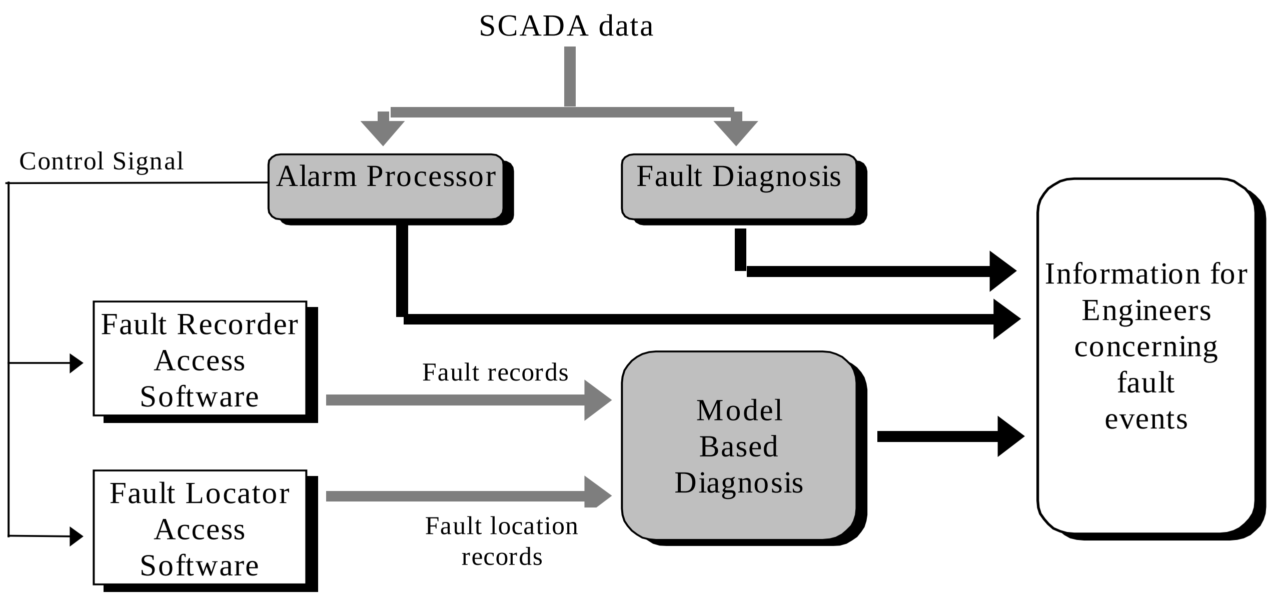


Figure 6.9: Integration of the Decision Support System

Initial research by McArthur et al [1] demonstrated that the General Diagnostic Engine (GDE) by deKleer and Williams [6] was suitable for analyzing protection scheme activity. However, the GDE requires to be extended to enable the temporal characteristics of the protection scheme to be taken into account.

This section investigates existing temporal MBD systems, identifying the main techniques adopted by each. The MBD module utilizes the most appropriate techniques identified for the analysis of fault recorder data and is described in detail.

Review of Existing MBD Systems

MBD systems utilize models of correct behavior to predict how the device being diagnosed should have operated. A comparison is made between the observed (actual behavior of the system being diagnosed) and the predicted behavior. If they are the same, assuming accurate models, the system is deemed to have operated correctly. A difference in the observed and predicted behavior is indicative of a device failure. Techniques such as the one based on the GDE, discussed in this paper, can then be used to identify the individual faulty device.

Several MBD systems have been developed to diagnose systems whose values change over time. This is essential when diagnosing protection scheme behavior as the components' state changes over time. Two such diagnostic systems are CATS [7] and MAGELLAN-MT [8][9] which adopt a strategy of "diagnosis at multiple times" and therefore diagnose the system on the basis of assumed discrete static periods. The CATS diagnostic system only considers values that hold at the same time instance and is not concerned with how they will develop over time. A similar approach is adopted by MAGELLAN-MT [8], with exceptions made for cases where actions have consequences at a later time. For example, component delays i.e. where a component requires a finite time to operate. However, in the subsequent paper by Dressler [9] no delays were considered, as all the relevant faults in their application domain, off-shore ballast tanks, can be identified without them.

Another approach adopted in MBD systems is the concept of time intervals. Examples of these include the Temporal Constraint Propagator [10], the Episode Propagator [11] and the diagnostic systems SIDIA [12] and FDE-2 [13]. They explicitly reference component values to a time (point or interval) known as an episode. An episode representing a *trip* value, for example, over the time period of 20ms to 50ms, produced using the circuit breaker {CB} component model is shown in Figure 10.

Value Time interval Environment

Trip(20ms, 50ms){CB}

Figure 6.10: An Episode.

The value relates to the measured or inferred state of a component in the system e.g. *trip*, *not trip*, *fault* etc. The time interval identifies the time period over which the value is present whilst the environment contains the components whose correct behavior is assumed when inferring the value.

Component delays have been incorporated into the Temporal Constraint Propagator, Episode Propagator and the diagnostic systems SIDIA and FDE-2. A delay of any duration is permitted in these systems except in the Episode Propagator, which is restricted to either a fixed delay period or no delay at all.

Modeling and measurement inaccuracies, or imprecisely defined components, can lead to inconsistencies being generated when the component has operated correctly. The diagnostic system CATS introduces tolerances to component values to allow inconsistencies relating to imprecision to be exonerated. A more recent diagnostic system FLAMES [14] extends this concept by introducing fuzzy logic to the intervals of component values permitting a measure of "faultiness" to be assigned to the diagnosis.

To overcome modeling uncertainty MAGELLAN-MT uses qualitative models as only significant deviations from the normal behavior are required to be recognized. However quantitative models are favored in the diagnostic systems FLAMES and CATS since they are required to detect subtle faults which would be undetectable using qualitative models.

5. PROTECTION SCHEME MODELING

The modeling techniques used to represent the protection scheme components are described in this section and build upon the approaches adopted by existing temporal MBD systems. The following issues are discussed:

- Temporal Representation
- Modeling issues
- Temporal tolerances
- Protection scheme model

Temporal Representation. In extending the GDE mechanism to enable temporal aspects to be taken into account an interval based approach was adopted. This is particularly suited to this application as the protection scheme components remain in a steady state for an appropriate length of time.

Typically, a fault recorder's sampling frequency is in excess of 2000Hz. A fault record storing a 0.5 second window of a power system fault would comprise of over 1000 samples. If a strategy of "diagnosis at multiple times" was adopted, as in the diagnostic systems CATS and MAGELLAN-MT, over 1000 individual static diagnostic cycles would be required. The fault record could be analyzed at every tenth sample, reducing the number of diagnostic cycles required, however this would also reduce the accuracy of the diagnosis.

As the diagnosis of the protection scheme is performed "post operation", all the relevant data can be collated before diagnosing the protection scheme's operation thus favoring an interval based technique.

The MBD technique used must also be capable of handling component delays as the operating time of the protection scheme components is of prime importance. As stated previously, component delays can be readily incorporated into an interval based methodology.

Modeling issues. Accurate quantitative models (e.g. physically based models) are required as both "soft" faults (the protection operating slightly slower than it was set for) and "hard" faults (the protection not operating) must be identified. As stated earlier the diagnosis is performed post operation and hence there is little to be gained by running qualitative models (e.g. logical models which are faster in execution).

The knowledge based systems APEX and RESPONDD identify which protection scheme to diagnose. Thus a model of the entire power system network is not required to diagnose one protection scheme. This approach of utilizing the SCADA system data to focus on the relevant protection scheme allows a detailed diagnosis to be achieved without the requirement for hierarchical models.

Fault models are not used as it is unrealistic to identify all component fault modes due to the enormous number of ways in which the components could fail to operate. A similar argument was put forward for not using fault models in the DEDALE diagnostic system [15].

Modeling a component's nominal operating characteristic is easier than modeling it precisely as discussed in [16] and [17]. A trip relay is a relatively simple component which operates, typically, in 10ms. This however can vary depending on contact wear, coil positions etc., which are never the same for two trip relays.

To overcome these modeling problems, temporal tolerances have been applied to the protection scheme component models.

Temporal tolerances. To overcome the uncertainty, introduced by modeling inaccuracies and measuring equipment, tolerances must be added to the time intervals. For example, if an observable is in the *trip* state over the time period (10, 30)±5 the component could have correctly operated between 5 and 15ms and reset between 25 and 35ms. Observables are episodes which have been measured.

Taking the operating condition first, it is assumed that the probability of the component operating is uniform over the time period 5 to 15ms. This can be represented as a continuous uniform probability density function $p_{op}(t)$. This identifies the probability of the component operating at time t.

The cumulative distribution function $P_{op}(t)$, which indicates the probability of the component having operated by time t, is defined as:

$$P_{op}(t) = \int_{-\infty}^{t} p_{op}(t')dt' \tag{1}$$

Similarly, a cumulative distribution function can be obtained for the reset condition $P_{re}(t)$, which represent the probability of the component having reset by time t. These two cumulative distribution functions can be combined to calculate the probability of being in the *trip* state at any time t. The cumulative distribution functions obtained for *trip*(10, 30)±5 are shown in Figure 6.11.

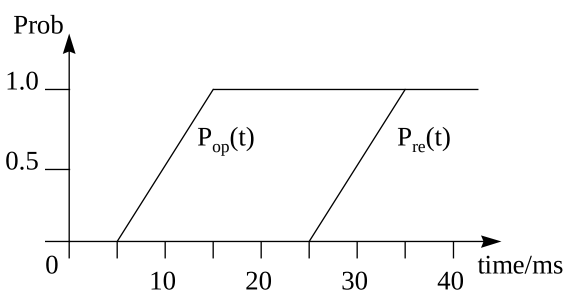


Figure 6.11: Cumulative Distribution Functions.

Let $P_T(t)$ be the probability of being in the *trip* state at time t. Then:

$$P_T(t) = \Pr ob(op(t) \cap \overline{re(t)}) \tag{2}$$

i.e.

$$P_T(t) = \text{Prob}\left\{\begin{array}{lcl} \text{of the component} & & \text{the component} \\ \text{having} & \text{and} & \text{not having} \\ \text{operated by} & & \text{reset by} \\ \text{time t} & & \text{time t} \end{array}\right\} \tag{3}$$

Assuming that the component must be in the operated state before it can reset:

$$\begin{aligned} P_T(t) &= \text{Prob}(op(t))\text{Prob}(\overline{re(t)}|op(t)) \\ &= \text{Prob}(op(t))\text{Prob}(\overline{re(t)}) \\ &= P_{op}(t)(1 - P_{re}(t)) \\ &= P_{op}(t) - P_{re}(t) \end{aligned} \tag{4}$$

The resulting probability of being in the *trip* state is depicted in Figure 6.12 for *trip*(10, 30)±5.

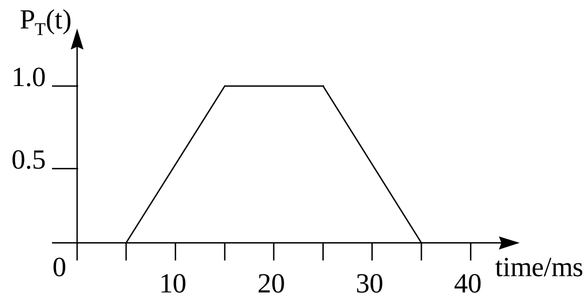


Figure 6.12: The Probability of Trip.

It should be noted that nothing can be assumed about the state of the value before 5 and after 35ms. The probability of being in the *trip* state increases linearly between 5 and 15ms. Between 15 and 25ms the value is definitely in the *trip* state. The probability of being in the *trip* state reduces linearly between 25 and 35ms to zero.

As a corollary, given a time interval over which a value is *not trip* and assuming the component must be in the reset state before operating, then the probability of being in the *not trip* state at time t is:

$$P_{NT}(t) = P_{re}(t) - P_{op}(t) \tag{5}$$

Protection scheme model. The protection scheme model in Figure 5.13 demonstrates the type of models currently utilized by the MBD module and how the component states are represented. In the interest of clarity, only the first main protection scheme at one end of the feeder is described here.

The first main protection scheme comprises a distance protection relay MP, which operates a trip relay TR, which subsequently trips the two circuit breakers CB1 and CB2.

Protection relay MP has an operating time of 15±5ms and a reset time of 30±5ms. The trip relay TR is modeled as operating in 10±2ms and resetting in 10±1 seconds. The circuit breakers CB1 and CB2 both have an operating time of 35±5ms and re-close in 19±1 and 17±1 seconds respectively.

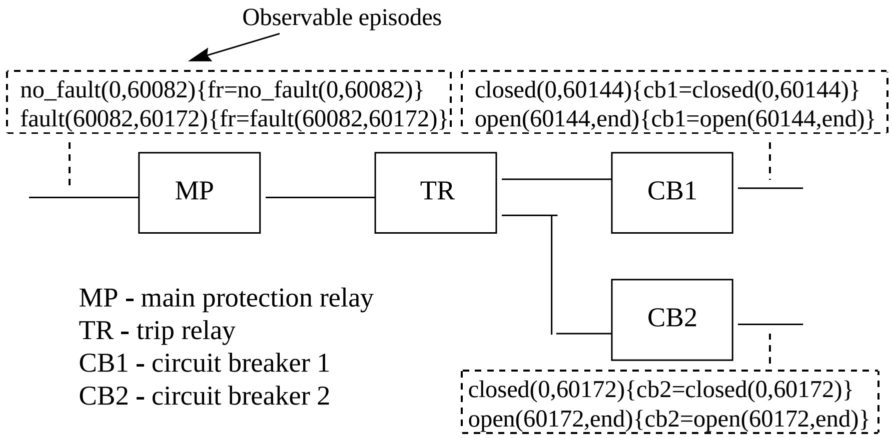


Figure 6.13: First Main Protection Scheme Model.

6. DIAGNOSTIC APPLICATION

In this section, the following issues are discussed:

- Diagnostic procedure
- Generating the diagnosis

Diagnostic procedure. As stated earlier MBD systems utilize the models of correct behavior to simulate how the system being diagnosed should have operated. Comparisons are then made with how it actually operated enabling any anomalous behavior to be identified.

Model Simulation

Taking the observable episodes shown in Figure 6.13. It is assumed that there is no power system fault within one minute (60000ms) of the fault being analyzed as an approximate buffer period to avoid the situation where the protection scheme is completing an auto-switching sequence.

The episodes should be interpreted as follows: *no_fault* (0,60082) {fr=no_fault (0,60082)} indicates that there was no network disturbance present over the period 0 to 60082ms. The components whose correct behavior is assumed in order to infer the episode are contained within the curly brackets. As the episodes in Figure 5.13 were observed they are only dependent on the actual measurement (fault recorder detected no fault between 0 and 60082ms) and not on any components. “End” is used to signify the end of the fault record.

Using the observables and the protection scheme models, all possible episodes are inferred. For example, the observables and the protection relay model shown in Figure 5.13 can be used to predict how the output of MP varies over time as shown in Figure 6.14.

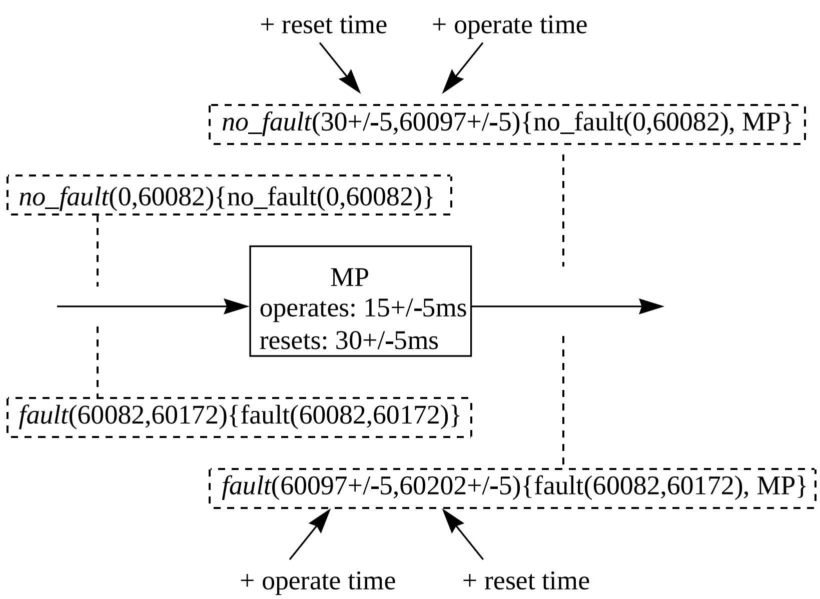


Figure 6.14: Propagating Forward Through MP.

After each episode has been inferred, checks are made for inconsistencies as described in the next section.

Identifying Inconsistencies

To establish if two episodes are consistent, a check is made between the probabilities of being in a different state, i.e. *trip* or *not trip*, over the same time period. The inconsistency is rated, based on the measure defined by Cano et al in [18], to identify the inconsistency between two probability envelopes.

However in this application all the states may not be inferred and hence the measure of inconsistency defined by Cano et al will not hold. For example a model and observable episode could be used to identify that a component has a probability of 0.6 of being in the *trip* state, at a certain time t. The remaining 0.4 probability undefined must also be taken into account when calculating the degree of inconsistency. This unaccounted probability is referred to as R_{obs}. The modified degree of inconsistency formulation is shown as follows:

$$\text{Degree of inconsistency} = \max_{B \in U} \{A_B\} \tag{6}$$

$$A_B = \begin{cases} P_{obs1}(B) - (P_{obs2}(B) + R_{obs2}) & \text{if } P_{obs1}(B) \geq P_{obs2}(B) \\ P_{obs2}(B) - (P_{obs1}(B) + R_{obs1}) & \text{if } P_{obs2}(B) > P_{obs1}(B) \end{cases}$$

Where:

B = State
U = Set of possible states
$P_{obs}(B)$ = Probability of state B at time t
A_B = Difference in the probabilities of being in state B, defined by two observables
R_{obs} = probability not accounted for at time t.

Subtracting R_{obs} ensures that only when an inconsistency will definitely occur, regardless of the other episodes inferred, with a degree of inconsistency greater than zero be generated. A degree of inconsistency between -1 and 0 indicates that the episodes from the two observable points are consistent, based on the presently inferred episodes. A value between 0 and 1 indicates a discrepancy, with the upper limit indicating a greater degree of inconsistency.

Inconsistency Example

Based on the model and observables in Figure 6.13 it is possible to infer the episodes shown in Figure 6.15 representing the operation of the circuit breaker.

open(60172 ± 12, end){CB1, CB2, cb2=open(60172,end), TR}
closed(2000 ± 3000, 60172 ± 12){CB1, CB2, cb2=closed(0,60172), TR}
open(60144 ± 0, end){cb1=open(60144,end)}
closed(0 ± 0, 60144 ± 0){cb1=closed(0,60144)}
open(60142 ± 12, end){fr=fault(60082,60172), MP, TR, CB1}
closed(29030 ± 2005, 60142 ± 12){fr=no_fault(0,60820), MP, TR, CB1}

Figure 6.15: Inferred Episodes for Circuit Breaker CB1.

Note the two episodes highlighted in Figure 6.15 are inconsistent because they represent different circuit breaker states over the same time period. This is shown graphically in Figure 6.8. The degree of inconsistency is quantified using equation 6 previously described.

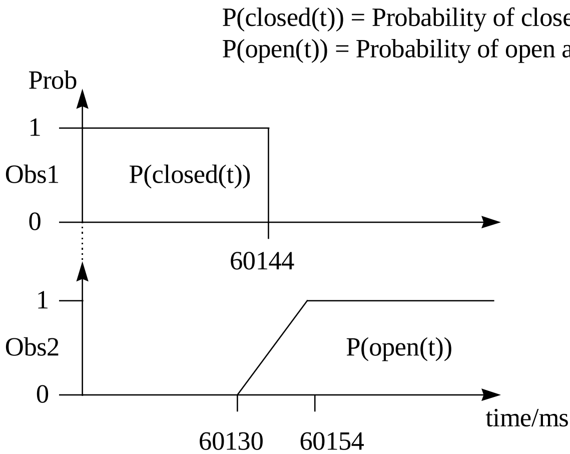


Figure 6.8: Inconsistent Episodes

Only the largest value of the degree of inconsistency produced between the two episodes is used. On the basis of this assumption, the degree of inconsistency is only calculated at times where discontinuities occur (i.e. episode start and end times). The associated probability values used and the degree of inconsistency calculated, using equation 6, at 60130, 60144 and 60154ms are shown in Table 6-II.

The maximum degree of inconsistency between the two episodes is 0.58 at 60144ms. That is, given these two episodes alone, the degree of inconsistency will be at least 0.58.

Other inconsistent episodes are generated throughout the protection scheme model and collectively they can be used to identify which component is most likely to have failed.

Table 6-II. Degree of Inconsistency

Time	60130		60144		60154	
	Obs1	Obs2	Obs1	Obs2	Obs1	Obs2
P(open)	0.0	0.0	0.0	0.58	0.0	1.0
P(closed)	1.0	0.0	1.0	0.0	0.0	0.0
R_{obs}	0.0	1.0	0.0	0.42	1.0	0.0
Degree of Inconsistency	0.0		0.58		0.0	

Generating the Diagnosis. The procedure for producing a diagnosis based on all the inconsistencies detected is explained fully in [6] and in relation to this application domain in [1]. In summary, the union of components on which the two inconsistent episodes rely is taken. This produces a list of components, of which at least one must be faulty, referred to as a conflict set. For the previously identified pair of inconsistent episodes the conflict set produced is < MP, TR, CB1 > with a degree of inconsistency of 0.58.

All conflict sets produced from the inconsistencies arising from the protection scheme model and observables in Figure 6.13, including their related degree of inconsistency, are shown in Figure 5.17.

< MP, TR, CB2 > 1.00
< TR, CB1, CB2 > 1.00
< MP, TR, CB1 > 0.58

Figure 6.17: Conflict Sets.

Combining the conflict sets in Figure 6.17 with a degree of inconsistency value of one, referred to as definite conflict sets, enables component(s) which are definitely faulty to be identified. The two definite conflict sets are combined in such a way so that at least one component from each is represented in all the diagnoses produced. For example component TR appears in both the definite conflict sets. That is, it could account for both inconsistencies defined by the definite conflict sets and is hence a possible faulty component. The resulting diagnoses are shown in Figure 6.18.

TR could have failed
or CB2 could have failed
or MP and CB1 could have failed

Figure 6.18: MBD Diagnoses.

Combining the remaining conflict set, with the degree of inconsistency value of 0.58, enables the diagnosis in Figure 5.18 to be further improved. That is, it takes into account inconsistencies which relate to components operating at their tolerance limits. The reasoning being that the component which can explain the most conflict sets, taking into account their degree of inconsistency, should be investigated first. The resulting diagnoses are shown in Figure 6.19.

TR could have failed
or CB1 and CB2 could have failed
or MP and CB1 could have failed
or MP and CB2 could have failed

Figure 6.19: Augmented MBD Diagnoses.

The first diagnosis in Figure 6.19 identifies that a maloperation of the TR component could be responsible for all inconsistencies detected. However should TR be deemed to have operated correctly, circuit breaker CB2 should be investigated as it could be responsible for all the definite conflict sets.

7. DISCUSSION OF THE MODEL BASED DIAGNOSIS APPROACH

MBD is ideally suited to utilizing the data available from fault recorders to identify if any of the components have failed. The approach adopted enables both "hard" and "soft" type faults to be diagnosed. The use of intervals of time to represent the temporal aspect of the data is appropriate for this application as the values remain in the same state for a number of sampling points.

Tolerances were also introduced in the temporal domain to overcome modeling inadequacies and inaccuracies in the actual measurements. The size of the tolerances are dependent on the model accuracy. For very accurate models a small tolerance would be required enabling subtle faults to be identified. If a coarser model was used then a larger tolerance would be required.

Presently, a prototype MBD system has been developed utilizing the techniques described in this tutorial. When the diagnostic system is fully implemented it will interface directly with the electricity utility's protection database. This database contains information relating to all the protection schemes fitted to the power system, including protection component types and settings.

In the MBD approach adopted the models and reasoning engine are separate, enabling many different types of protection schemes to be diagnosed by simply changing the models. A component model library is currently being constructed. Ideally all the protection scheme components installed on the utility's power system network should be modeled, enabling all the feeders to be analyzed. The model library could also include dynamic models [19] which would enable a more detailed diagnosis of the protection scheme's behavior to be made.

CURRENTLY, THE MODELS USE DISCRETE STATE VALUES, EITHER *TRIP*, *NOT TRIP*, *FAULT*, *NO FAULT* ETC. THE USE OF MORE ACCURATE MODELS WILL REQUIRE ANALOGUE VALUES TO BE REPRESENTED. SIMILAR PROBLEMS WILL OCCUR, AS IN THE TIME DOMAIN, WITH MODELING AND MEASUREMENT INACCURACIES CAUSING DISCREPANCIES TO BE INCORRECTLY IDENTIFIED. THUS TOLERANCES WILL ALSO HAVE TO BE ADDED TO THE ANALOGUE VALUES AND AN ASSESSMENT MADE OF WHETHER THE EPISODE IS OUTSIDE EITHER ITS VALUE OR TEMPORAL TOLERANCE.

8. FUTURE TRENDS

The following issues are discussed:

- Access to information
- Intelligent Agents
- Prototype Intelligent agent system

Access to information. A significant issue for the DSS is providing access to information. This information is of use to many engineers within the company from those based at the corporate headquarters to engineers based within regional offices. Information from the DSS is also of benefit to senior management personnel, for example, to provide information relating to a disturbance which affected a significant industrial customer.

The obvious solution to this problem is the use of a corporate intranet. Through this, information from the DSS could be accessible via "web pages" using an appropriate "browser". By adopting this approach information from the DSS would be widely available throughout the company.

As shown in figure 6.4 of this tutorial, and as discussed, it is appropriate to provide a summary event report as the first point of contact for engineers. This will summarize the salient information from all of the data interpretation systems. This could then be presented within an event summary web page. Figure 6.20 presents an example event summary web page.

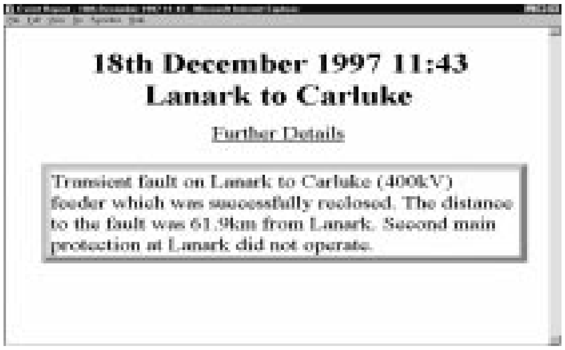


Figure 6.20: Event Summary Web Page

Given that information from the DSS (and possibly other data gathering/interpretation systems) is available through the intranet (using appropriate web pages), then it would be useful to link any additional information which is related or explanatory.

If the DSS determined that there had been a disturbance on a particular feeder, for example, it would be of value to link the web page detailing this disturbance with other supporting evidence. The waveform from the fault locator and a single line diagram of the appropriate area of the network could be linked to the event summary web page. Information from databases, containing details on protection could also be linked. The possibility of linking supporting evidence to the summary web page is shown in figure 6.20 through a "Further Details" link.

Some form of intelligent processing is required to compile the varying inputs to the event report and associated web page. It will need to co-operate and integrate with different systems to allow all the pertinent details to be compiled. Intelligent agent technology is seen as an appropriate technology to employ the intelligence and systems integration capabilities that are required.

Intelligent Agents. Software agents and agent based systems have been an area of interest within the research community for a long period of time. They address the issue of interaction between heterogeneous software and hardware systems across distributed platforms [20].

More recently, attention has been focused on the area of intelligent agents. These extend the capabilities of agent based software by exhibiting the following characteristics [21]:

- *Autonomy*: intelligent agents operate without the direct intervention of humans or others, and have some kind of control over their actions and internal state.
- *Social ability*: agents interact with other agents via some kind of agent communication language.
- *Reactivity*: agents perceive their environment and respond in a timely fashion to changes that occur in it.
- *Pro-activeness*: agents do not simply act in response to their environment but they are able to exhibit goal-directed behavior by taking the initiative.

Intelligent Agent Based Event Report Compiler

If the application in question is considered then the characteristics of intelligent agents are necessary to achieve the desired functionality. That is, automatic creation of an event report based upon various data sources, as described in this tutorial, and its formatting as a web page (figure 5.20) requires intelligent agent capabilities. The abilities described as autonomy, reactiveness and pro-activeness cover the fact that each of the interpretation systems must determine when their results or data may be appropriate to the report being produced. Furthermore, the agent creating the report must decide when it requires supporting information or data (e.g. single line diagrams, fault record traces, details from databases, etc.) for inclusion to help the eventual end user. In addition to these capabilities, social ability is required between each of the systems in order that they can communicate.

The use of intelligent agent technology to compile event reports for an intranet leads to the creation of easily accessible web pages which describe the events on the power system. This is shown in figure 6.21. The page labeled "Event Summary Web Page" would be of the format shown in figure 6.20. A web page will also exist which lists all the web page event reports which are available. This is labeled as an event list web page in figure 6.21. Therefore, a single point of access to the information required by engineers is provided. Overall, the intelligent agent approach is providing open communication and distributed Prototype Intelligent Agent System .

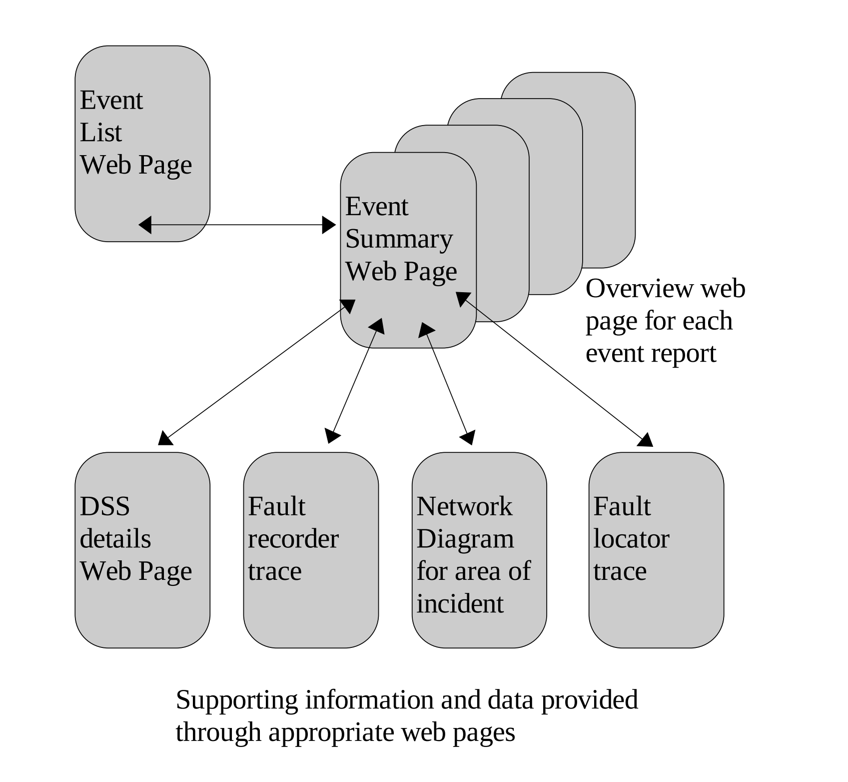


Figure 6.21 - Format of Event Report Web Pages

A prototype agent based processing capabilities.

system is being developed to provide the capabilities described within this paper. An intelligent event report compiler has already been coded as an agent using the Java programming language. This has made use of the Java Development Kit available from Sun Microsystems, inc. [22].

To allow this Java based agent to operate within a multi-agent environment the Java Agent Template (JAT) is used. The JATLite package (available from the Agent Based Engineering Group, which is a part of Stanford University's Centre for Design Research) is a package of programs written in the Java language which allow users to create software agents that communicate robustly over the Internet [23]. Java is the language of preference for creating agents as it allows them to run on heterogeneous platforms. JATLite provides a template for creating agents, therefore there is no restriction on their functionality. The developer can use the template to create the communication capabilities and then the intelligence and functionality can be coded freely.

JATLite provides an "agent router", which controls interaction in a multi-agent environment, and Java classes which control registration and connection to the router. In terms of agent architectures the use of an agent router is one model [20]. However, it is appropriate for the prototype system being discussed.

An added incentive to use the facilities offered by JATLite is the fact that it supports the emerging standard in agent communication languages (ACL). A standardized ACL is required to ensure that all agent based systems are able to easily communicate with one another. The first aspect of this standard is the Knowledge Interchange Format (KIF) which is a common language for expressing the content of a knowledge base [20][24]. Further to the KIF, the Knowledge Query and Manipulation Language (KQML) is a message format (and message handling protocol) which is designed to support run-time knowledge sharing among agents. KIF, KQML and other associated standards [20][24] are designed to simplify the communication between agents, therefore allowing interaction and co-operation. The prototype described within this paper makes use of this ACL standard.

Within the prototype intelligent event report compiler the underlying JATLite functions will allow it to access the appropriate data and information sources. This is shown in figure 5.22. The report compiler will interact with the systems shown in figure 5.9 to dynamically create a report.

The event report compiler will also create the appropriate web pages. To achieve this it will format the report with Hypertext Markup Language (HTML) keywords. Therefore, the environment shown in figure 6.21 is created dynamically following any power system events.

Figure 6.22 demonstrates the ideal situation whereby JATLite is used to allow each system to operate as an intelligent agent. This is easily accomplished for new software modules which can be coded in Java and can make use of JATLite's capabilities, as was the case for the intelligent event report creator. However, legacy software and software produced by third party vendors add further levels of complexity to this process.

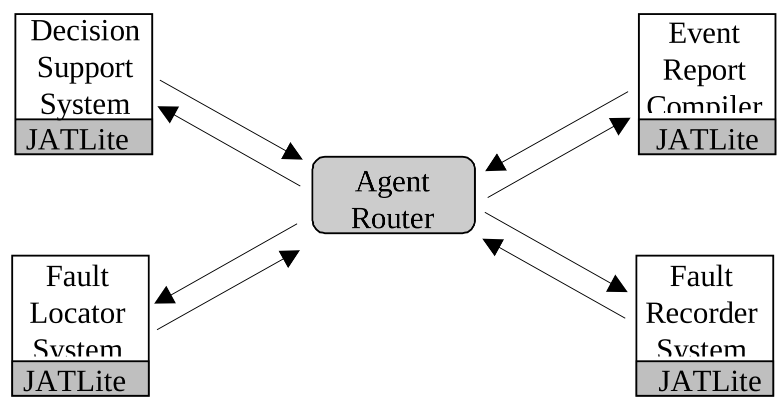


Figure 6.22 - Agent Architecture

If JATLite is to be the basis for converting all the software systems into intelligent agents then some form of interfacing code will be required to allow legacy code to make use of its capabilities. Various approaches have been used by other researchers to convert existing software into agent based systems. These approaches range from interface functions and "wrapper" code to complete re-coding of the systems [20]. An aspect of this research activity is the identification of the most effective mechanisms to convert existing software systems into a community of co-operating intelligent agents.

At present the communications between the fault locators, fault recorders and intelligent modules within the DSS is not agent based. This research is focused upon providing agent based facilities for data and information exchange following on-line interpretation. Obviously, once the legacy systems are converted to operate as agents, then this will become the communications protocol used for all the necessary communications, and will replace the mechanisms used currently.

9. REFERENCES

[1] S.D.J. McArthur, S.C., Bell, J.R. McDonald, R., Mather, S.M. Burt, "The application of model based reasoning within a decision support system for protection engineers," *IEEE Transactions on Power Delivery*, Vol. 11, No.4, October 1996, pp. 1748-1755.

[2] S.C. Bell, S.D.J. McArthur, J.R. McDonald, G.M. Burt, R. Mather, S.M. Burt, "On-line protection performance analysis using intelligent systems," *International Journal of Engineering Intelligent Systems - Special Issue on Intelligent System Applications on Power System Protection*, Vol. 5, No.4, December 1997, pp. 57-63.

[3] S.D.J. McArthur, S.C. Bell, J.R. McDonald, R. Mather, S.M. Burt, T. Cumming, "Decision support for the interpretation of power network data of relevance to protection engineers," *CIGRE 1996*, Group 34, paper 34-203, August 1996.

[4] D.J. Young, K.L. Lo, J.R. McDonald, R. Howard, J. Rye, "Development of a practical expert system for alarm processing," *IEE Proc. C - Gen, Trans. & Distrib.*,1992, 139, (5), pp. 437-447.

[5] J.R. McDonald, G.M. Burt, D.J. Young, "Alarm processing and fault diagnosis using knowledge based systems for transmission and distribution network control," *IEEE Trans. Power. Sys.*, 1992, 7, (3), pp. 1292-1298.

[6] J. DeKleer, and B.C. Williams, Diagnosing multiple faults, *Artificial Intelligence*, 1987, 32, (1), pp. 97-130.

[7] P. Dague and O. Jehl, "When oscillators stop oscillating," *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, IJCAI 91, 1991, pp. 1109-1115.

[8] O. Dressler, "Model based diagnosis on board: MAGELLAN-MT inside," *Working Notes of the Fifth International Workshop on Principles of Diagnosis*, 1994.

[9] O. Dressler, On-line diagnosis and monitoring of dynamic systems based on qualitative models and dependency-recording diagnosis engines, Proceedings of the 12th European Conference on Artificial Intelligence, ECAI 96, 1996, pp. 461-465.

[10] B.C. Williams, Doing time: putting qualitative reasoning on firmer ground, Proceedings of the National Conference on Artificial Intelligence, AAAI 86, 1986, pp. 105-112.

[11] R. Decker, "Modelling the temporal behavior of technical systems," *Proceedings of the German Workshop on Artificial Intelligence*, GWAI, 1988, pp. 41-50.

[12] T. Guckenbiehl and G. Schafer-Richter, "SIDIA - extending prediction based diagnosis to dynamic models," *Readings in Model Based Diagnosis*, Hamscher, W., Console, L. and DeKleer, J. (editors), Morgan Kaufman, 1992, pp. 309-317.

[13] S. Iwanowski, "An algorithm for model based diagnosis that considers time," Annals of Mathematics and Artificial Intelligence, 1994, 11, pp. 415-437.

[14] F. Mohammed, M. Marzouki, and M.H. Touati, "FLAMES: a fuzzy logic ATMS and model based expert system for analog diagnosis," *European Design and Test Conference*, EDTC 96, 1996.

[15] P. Dague, "Model based diagnosis of analog electronic circuits," *Annals of Mathematics and Artificial Intelligence*, 1994, 11, pp. 439-492.

[16] P. Dague, O. Raiman, and P. Deves, "Troubleshooting: when modeling is the trouble," *Readings in Model Based Diagnosis*, Hamscher, W., Console, L. and DeKleer, J. (editors), Morgan Kaufman, 1992, pp. 218-223.

[17] R. Davis, and W. Hamscher, "Model based reasoning: troubleshooting," *In exploring artificial intelligence: survey talks from the national conference on artificial intelligence*, edited by Shrobe, H.E., 1988, pp. 297-345.

[18] J.E. Cano, S. Moral, and J.F. Verdegay-Lopez, "Partial inconsistency of probability envelopes," *Fuzzy Sets and Systems*, 1992, 52, (2), pp. 201-216.

[19] A. Dysko, J.R. McDonald, G.M. Burt, J. Goody, and B. Gwyn, "Dynamic modelling of protection system performance," *Proceedings of the 6th International Conference on Developments in Power System Protection, DPSP 97*, March 97, pp. 381-385.

[20] M.R. Genesereth, and S.P. Ketchpel, "Software agents," *Communications of the ACM*, vol. 37, no. 7, 1994, pp. 48-53.

[21] M. Wooldridge, and N.R. Jennings, "Intelligent agents: theory and practice," *Knowledge Engineering Review*, vol. 10, no. 2, 1995, pp. 1-62.

[22] http://java.sun.com/products/index.html

[23] http://java.stanford.edu

[24] T. Finin, Y. Labrou, and J. Mayfield, "KQML as an age communication language," J. Bradshaw (editor), *Software Agents*, MIT Press, Cambridge, 1997.