

Automated Fault Location Analysis – Data Analysis Method

Introduction

FirstEnergy Service Company on behalf of its transmission owning affiliates determined there was a need to improve the accuracy and speed of its fault location process especially for the 138 and 69 kV system. This system is heavily tapped with industrial customers and substations. The main objective of this project was to see if the need to call out staff, usually overnight, to run the fault-location program could be eliminated for the majority of faults that occur in a pilot/test area.

The overall objective of this effort is to reduce the time to determine where a fault has occurred with sufficient certainty to route field crews to the location of the fault quickly and improve restoration times. At some locations, there may be the ability to sectionalize the 69 kV or 138 kV transmission network so that operations staff can begin restoring customers in areas not directly affected by the faulted line section.

The utility uses a “pull” model and is based on a system operator initiating the process after an event has occurred by enabling a point in SCADA at the relevant stations enabling the process to remotely connect to each site and gather the data from each device based on information contained in a reference table. The first step in the automation plan was documenting the existing manual process and then systematically develop the automation process.

Problem Statement

The utility on behalf of its transmission owning affiliates determined there was a need to improve the accuracy and speed of its fault location process especially for the 138 and 69 kV system. This system is heavily tapped with industrial customers and substations. The main objective of this project was to see if the need to call out staff, usually overnight, to run the fault-location program could be eliminated for the majority of faults that occur in a pilot/test area.

When a fault occurs on a major transmission line it is critical that a protection engineer rapidly assess the situation and accurately determine the location of the fault to minimize the outage duration. Unfortunately, this is a very challenging task for a number of reasons. First, the source of the trip and the associated line must be identified by the SCADA operators. Then the appropriate sensor/instrumentation data associated with the tripped line needs to be evaluated. This data can be in the form of relay event records and/or DFR – Digital Fault Recorder files, which can sometimes be in a proprietary format. In addition to the different recording formats, the DFR records may only implement a few of the current or voltage measurements of the faulted line. Once the protection engineer has gathered all the appropriate data they must analyze the waveforms to determine the nature of the fault (phase -> ground, phase -> phase), the “to and from” bus numbers and the maximum fault current in RMS. All of this is then manually entered into a fault analysis program and a fault study is initiated to determine the approximate location of the fault. The goal of this project is to automate most of these steps to significantly reduce the amount of time required to determine the fault location and impact on the protection engineer.

Objective

The overall objective of this effort is to reduce the time to determine where a fault has occurred with sufficient certainty to route field crews to the location of the fault quickly and improve restoration

times. At some locations, there may be the ability to sectionalize the 69 kV or 138 kV transmission network so that operations staff can begin restoring customers in areas not directly affected by the faulted line section. The expectation is that by automating the analysis, the time previously spent in the off-hours shifts to locate a protection engineer, provide him or her with the needed information, and then perform the analysis would be greatly reduced. Streamlining this process will help improve restoration times and decrease the length of outages for our customers.

Approach

Today they use a “pull” model for data retrieval of relay records. For some dial-up connected locations, this is based on a system operator initiating the process after an event has occurred by enabling a point in SCADA at the relevant stations to allow remote access. This enables the engineer to remotely connect to each site and gather the data from each device. For DFR records, there is an automated polling of records to get the information. The relationship between the relays and the lines that are associated with them is contained in a cross reference table. Once the data is acquired, the process to extract the relevant data and locate the fault begins.

The first step in the automation plan was documenting the existing manual process. Figure 1 shows the relevant process steps for the “pull” model. The process flow has two paths through it and requires the creation of a new table. The first path identifies the parameters of the fault such as fault type and magnitude. The other path determines what network changes are required to accurately represent the grid elements that are out of service in the vicinity of the fault. The new table was developed to essentially “connect” the operations’ field data historian with the fault study program. [1]

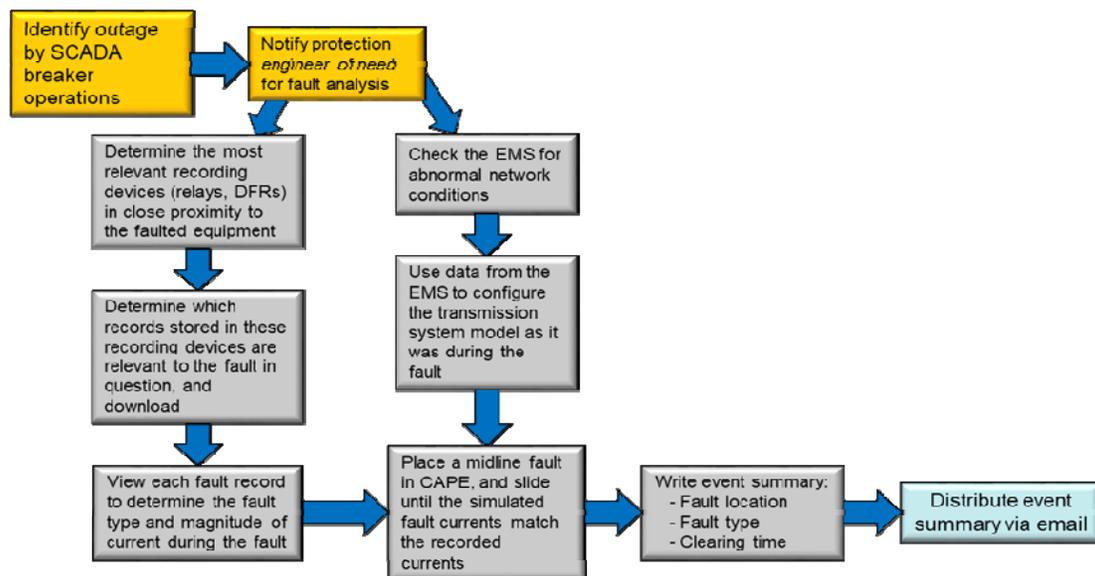


Figure 1
Master fault record table – Pull Model

Data Sources and Format

The two primary sources for fault information are from the relay fault records and the DFR records. These files typically contain waveform data sampled at 16 samples/cycle for all the key electrical parameters (IA, IB, IC, IR, VA, VB, VC) as well as digital/trip state within the relay. The SEL fault

challenges in using the DFR data is that, in an effort to reduce cost, utilities will implement a minimum number of channels to monitor a line, typically 1 phase current and the residual current. Also, a single location in the substation is often used to measure voltages that maybe several buses away from the fault and can even be isolated due to switching as a result of the fault. When determining the fault, additional logic will need to be provided to build the appropriate fault study program macro parameters. Figure 4 shows an example of the channels implemented in a typical DFR

channel analogue	1	"kV"	80.00	254.500	0	0.000	AC
channel analogue	2	"kV"	80.00	254.500	0	0.000	AC
channel analogue	3	"kV"	80.00	254.500	0	0.000	AC
channel analogue	4	80.00	15155.000	0	0.000	AC	
channel analogue	5	"A"	240.00	45418.000	0	0.000	AC
channel analogue	6	"A"	800.00	30226.000	0	0.000	AC
channel analogue	7	"A"	800.00	31593.000	0	0.000	AC
channel analogue	8	"A"	240.00	45472.000	0	0.000	AC
channel analogue	9	"A"	1200.00	47536.000	0	0.000	AC
channel analogue	10	"A"	1200.00	45460.000	0	0.000	AC
channel analogue	11	"A"	240.00	45533.000	0	0.000	AC
channel analogue	12	"A"	1200.00	45411.000	0	0.000	AC
channel analogue	13	"A"	1200.00	47311.000	0	0.000	AC
channel analogue	14	"A"	160.00	30274.000	0	0.000	AC
channel analogue	15	"A"	800.00	30266.000	0	0.000	AC
channel analogue	16	"A"	800.00	31558.000	0	0.000	AC
channel analogue	17	"A"	120.00	22724.000	0	0.000	AC
channel analogue	18	"A"	120.00	22700.000	0	0.000	AC

Figure 4
Typical DFR record

In an effort to standardize on a single file format thereby reducing the complexity of the program and eliminating the need for proprietary software modules, the industry standard COMTRADE (Common Format for the Transient Data Exchange for Power Systems) format was chosen. It is a file format for storing oscillography and status data related to transient power systems that is supported/maintained by the PSRC (Power Systems Relaying Committee) of the IEEE. The file format consists of 2-3 files:

- 1) CFG file containing signal names, number of samples, min/max values, conversion coefficients
- 2) DAT containing digitized waveform data
- 3) HDR – optional file containing additional configuration/metadata about the trip

All the fault records regardless of their native formats will be converted to COMTRADE file format. In this case there are three relevant files to consider:

- CFG – Configuration data about .DAT file
 - signal names, start time of the samples, number of samples, min/max values, scaling equations ($y=mx + b$)
- DAT – Digitized sample (waveform) data
 - Rows and columns, sample #, μ sec timestamp, integer values
- HDR – Machine location, Time of recording
 - Standard format but the data includes sample rates, number of channels captures and time accuracy varies depending on underlying device

Some observed differences between the relay based event records and those capture by a dedicated digital fault recorder are shown the table below.

Description	Relay	DFR
-------------	-------	-----

Reporting Rate (samples / cycle)	16	96
Units of measure	None	Defined
Location	Not included	Included
Channel Information	Minimal	More informative
Typical channels measured	All	Partial

Data Cleansing

For a variety of historical reasons, most naming conventions are not standardized between applications and devices. The name of the line in fault analysis program and the corresponding COMTRADE file line/location names may be different. Sometimes the names are abbreviated or a dash replaces a blank or the suffix “LINE” is appended to the name. The fault analysis automation program will need to be provided with a translate table, a form of Rosetta Stone to deal with the inconsistencies between the various data sources. Figure 5 shows the process for automatically generating the master list of all possible EDOA (Energy Delivery Outage Application) line names and using REGEX (regular expression) rules to match the fault analysis names. The result is a data join with the following column headings:

- Line Name
- From Bus Number
- From Bus Name
- To Bus Number
- To Bus Name
- Circuit Number Line Length
- From Sub Number
- From Sub
- To Sub Number
- To Sub

The resulting table forms the basis for the EDOA Rosetta Stone data join that includes the EDOA line name, fault analysis name and line length in meters.

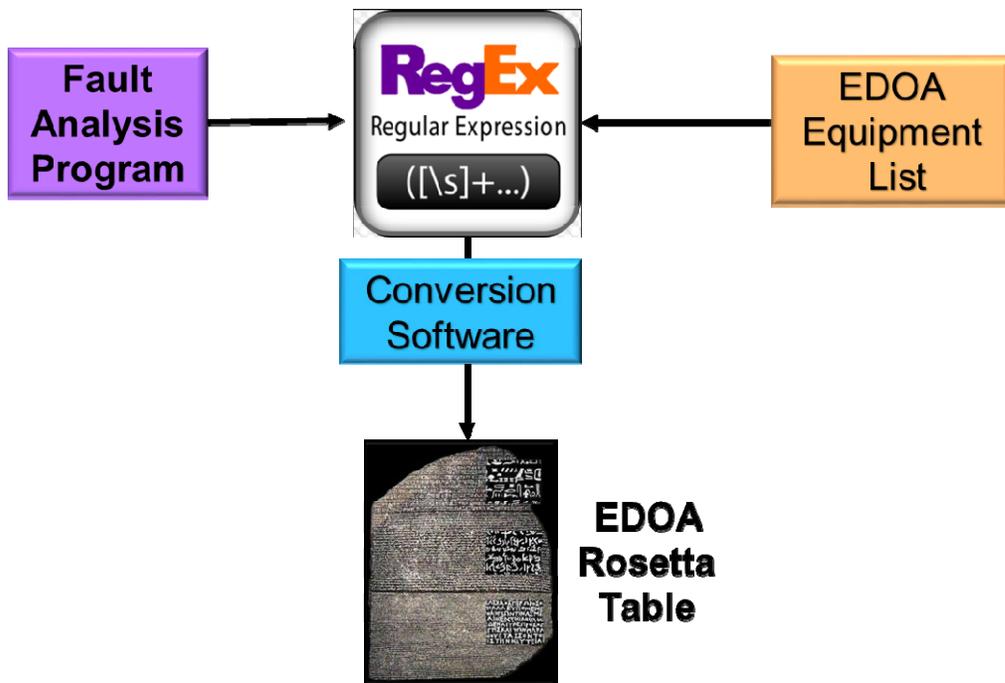


Figure 5
Build EDOA Rosetta Database – “Fuzzy” (regex) search

The DFR Rosetta Stone of DFR records is more challenging because a single DFR can monitor over 100 analog sensors. A special software program was written to digest all the DFR configuration files and produce the DFR Rosetta Stone. The columns include: FAULT STUDY NAME, DFR Location, IA, IB, IC, IR, VA, VB, VC.

An anomaly was discovered during the development process that indicated very high fault currents were recorded. However, further investigation showed that these were typically spurious readings and not an actual fault condition. Upon investigation, in most cases the transient spike existed on all channels at a time when the line was not in service.

Also, some DFR devices generated many fault records during times when there is no fault activity. Further investigation revealed that channel noise and low current trip limits can cause the generation of false trip recordings.

To accommodate the anomalies presented it was decided to initiate the fault study automation process from an event generated by the Energy Delivery Outage Analysis (EDO) system to avoid processing of spurious event records. [3]

Fault analysis program database and automation

The fault analysis tool is based on an extensive internal database of over 250 tables that describe in detail line, relay, breakers and substation physical and electrical properties. It is constantly maintained and updated to accurately represent the current state of the electrical transmission grid.

Internal to the fault study program is an open source database that is used to hold all the system information including line names, connectivity, electrical and physical (length) properties. The entire open source database was exported into a commercial database so that specific tables and fields could be used to build the automation. Figure 6 is an example of using fault study program tables to build name translate tables and provide electrical and physical data (line mileage) to CTA. [2]

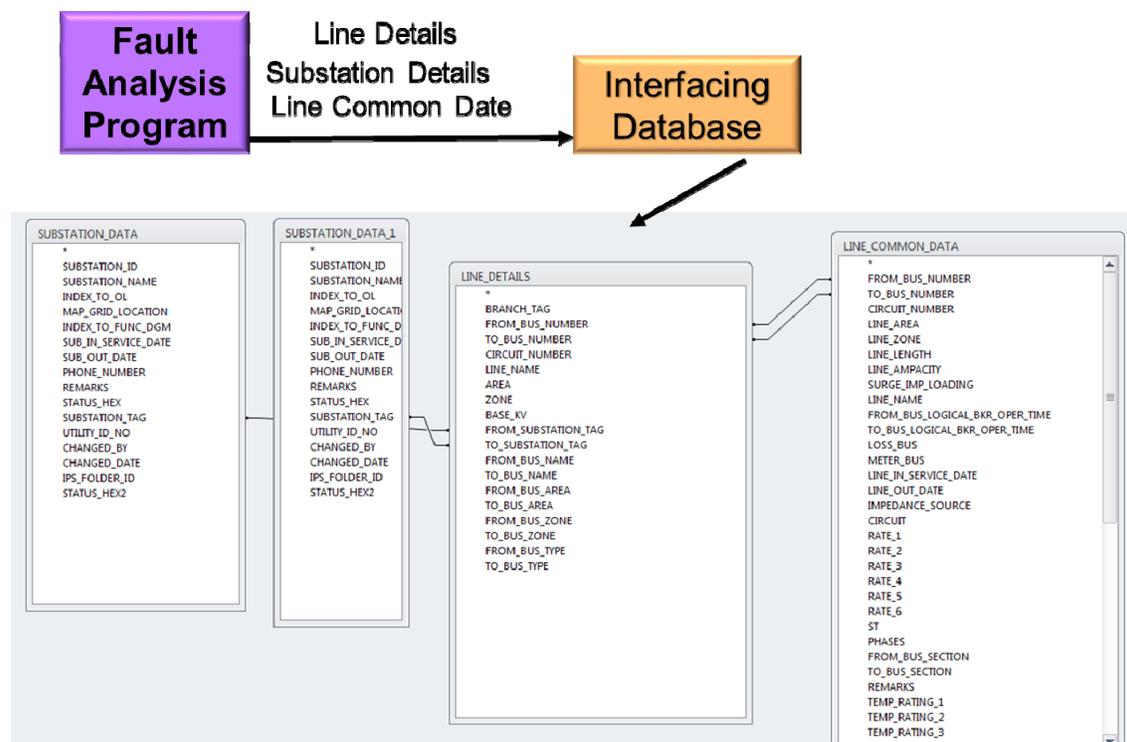


Figure 6

Build EDOA Rosetta Database – fault analysis program data “join”

The fault analysis program can be run using the standard graphical user interface (GUI) that gives the protection engineer the ability to select the appropriate faulted lines and the faults electrical parameters. Alternatively, a macro can be executed to automate the fault location process using the windows command line.

To achieve the desired automation, a way to run the fault study program without the standard GUI was needed. Since a command line interface to the fault study program already existed, the ability to create a batch file to launch the fault study program via. this interface was included in the CTA program.

In order to allow a particular asset to be included in the analysis, the fault study program session to include or not include certain power system elements (such as a future transmission line) they use the Equipment Category feature of the fault study program. Beyond a baseline model, a fault study program user can choose which power system elements to include in any particular fault study program session. In response to our request, the vendor added the capability to choose Equipment Categories for a fault study program session which was launched via the command line interface.

Data Flow

Figure 7 shows the high-level data flow for the COMTRADE Analysis application (CTA). COMTRADE files from both relays and DFR records are converted to COMTRADE format and placed into folders on the CTA server. When a fault occurs, a file is sent from the FE EDOA system and placed in the EDOA directory. The CTA program is periodically polling the EDOA directory for new fault files. Once CTA detects that a fault has occurred the following steps are taken:

- 1) EDOA file is parsed and the name of the faulted line is extracted
- 2) All COMTRADE files in the COMTRADE directory are inspected to determine if any are associated with the faulted line.
- 3) If an appropriate COMTRADE file is found, all 3 of the COMTRADE file (.dat,.cfg,.hdr) are loaded and analyzed.
- 4) The oscillography/waveform data is analyzed to determine the maximum fault current and the type of fault
- 5) The fault analysis program macro is modified to reflect the appropriate electrical parameters for the fault
- 6) The fault analysis program is launched using the macro parameters. Normally CTA would wait for fault analysis program to complete. However, when fault analysis program is finished with its analysis it returns to the interactive GUI mode waiting for user input. To “work around” this problem CTA waits for a predefined period (currently 30 seconds) and then automatically kills/aborts the fault analysis program task in windows.

- 7) The output from the fault analysis program run is analyzed and the fault analysis program results file is timestamped and moved to the fault analysis program History directory.

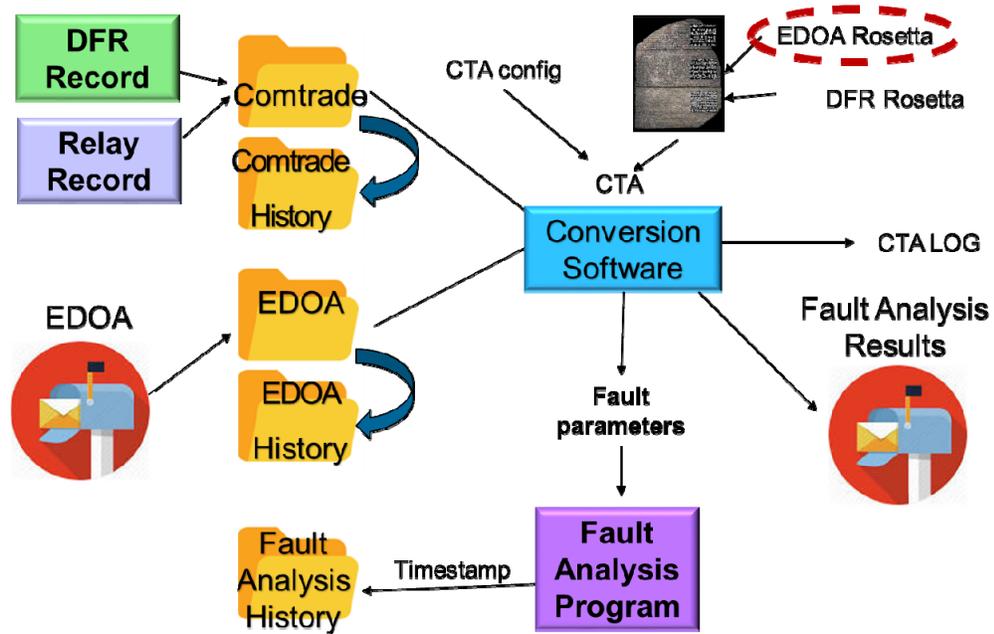


Figure 7
Overall High Level Data Flow CTA – COMTRADE analyzer automation

- 8) The EDOA file that started the entire process is moved to the EDOA history directory to avoid creating duplicate fault analysis program runs. It is possible that COMTRADE files associated with a given fault will arrive after the EDOA notification. Therefore, an EDOA time window was implemented (Figure 8) to keep the EDOA request “active” for a period after its arrival.
- 9) CTA returns to its polling mode waiting for another EDOA request to be generated
- 10) All the above steps and the results of the analysis are all logged to the CTA log file.

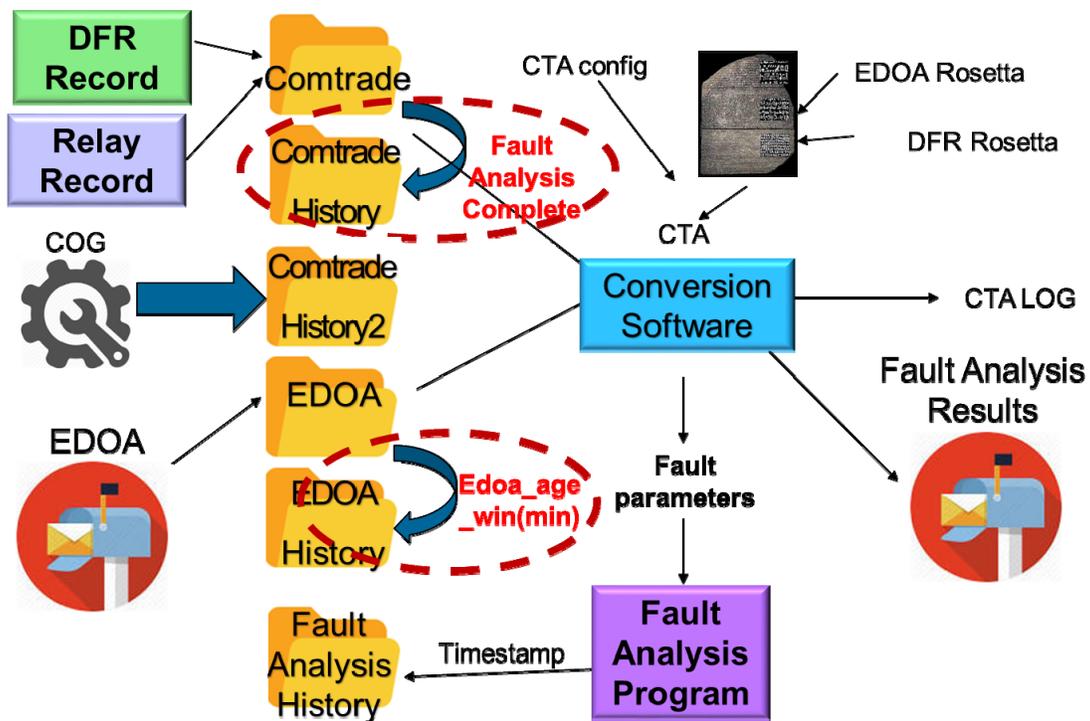


Figure 8
EDOA request AGE Window

Waveform Analysis

The oscillography / waveform data is analyzed to determine the maximum fault current and the type of fault (phase to ground, phase to phase...). To determine the correlation between the various phase currents a Euclidian Distance algorithm calculation is performed between all 4 line currents (depending on sensor implementation). As can be seen in Figure 9, the Euclidian Distance algorithm is both shape and magnitude sensitive making it a suitable algorithm for phase to phase correlations. Once the correlation is determined, the appropriate fault analysis macro parameters are updated and logged to the CTA log file

- I entire waveform Euclidian dist A,B,C, currents
- A-B 54875.9
- A-C 3054.6
- B-C 54727.2

- IR entire waveform Euclidian dist A,B,C, currents
- A-IR 55022.6
- B-IR 4287.8
- C-IR 55870.3

- B phase correlated with residual currents
- B – ground fault



Figure 9
Euclidian Distance – Entire Waveform

Conclusion

This paper has presented how one can automatically extract data from three sources of event data. Then take that data and adjust it to be able to “connect” across different data tables through the use of a number of translation tables (Rosetta Stones) using Fuzzy (regex) search tools. The resultant data can then be joined with other data to establish a data source that can be automatically run using a fault analysis program macro to locate the possible fault location. Lastly, presented a summary process flow and waveform analysis approach. While still being finalized a very solid foundation has been built that can be used to fully automate the entire process.

BIBLIOGRAPHY

1. C. Black & G. Wilson Southern Company P. Myrda EPRI CIGRE GOTF October 2017, Cleveland, OH – “Automated Fault Location Analysis – Data Analysis Method”
2. C. Black & G. Wilson Southern Company P. Myrda EPRI M. Gehrs Gehrs Consulting PAC World, June 2017 Wroclaw, Poland – “Benefits of Modern Database Tools Applied to Event Record Data Extraction”
3. P. Myrda - (EPRI), C. Black, G. Wilson, Southern Company, M. Behe FirstEnergy CIGRE GOTF October 2016 Philadelphia, PA – “Automated Fault Location Analysis - Data Gathering Assessment”