

Smart grid data analytics for digital protective relay event recordings

Tomo Popovic · Mladen Kezunovic · Bozo Krstajic

Published online: 8 June 2013
© Springer Science+Business Media New York 2013

Abstract Information systems and intelligent smart grid data analytics will have a critical role in managing the massive amount of data becoming available in power system substations. Digital protective relays are multi-functional intelligent electronic devices based on microprocessors, which are being installed in substations throughout the power grid. New digital relays are replacing old-fashioned electro-mechanical or solid-state relays, and besides their protective function, they are coming equipped with monitoring capabilities. These monitoring capabilities are creating potential for better observability of power systems, redundancy in measurements, and improved decision-making process when operating the system. This article discusses the implementation requirements for a fully automated data analytics solution that provides data integration, fault analysis, and visualization based on the event data recorded by digital protective relays.

Keywords Smart grid · Data analytics · Intelligent electronic devices · Digital protective relays · Power system fault analysis · Intelligent systems

1 Introduction

Contemporary developments in information technology, including information systems and communications developments, have changed business and organizational practices beyond all expectations (Sage and Rouse 1999). All around us, information is continuously being digitized, moved around, archived, and converted into usable knowledge. Protective relays are one of the key components in power system and have important role in maintaining security and reliability of the system (U.S.-Canada Power System Outage Task Force 2004). In recent years, due to development of information and communication technology, there was a dramatic increase in presence of digital protective relays (DPRs) throughout the power system. DPRs are being deployed in new substations, as well as being used to replace electro-mechanical or solid state relays. The modern DPR fall into the category of substation intelligent electronic devices (IEDs), which can provide users with abundant data recorded every time a relay experiences a fault or event. The recordings include samples from current and voltage waveforms, status of input and output contacts, status of internal protection and control elements, and relay settings (Costello 2000). This information is usually stored in event-triggered files containing event reports, waveform oscillography, and relay settings.

Large-scale deployment of DPRs and other substation IEDs resulted in a large volume of data that is becoming overwhelming for the personnel responsible for data collection and processing. The only feasible way to collect and analyze all this data is to automate the process (McDonald 2003; North American Electric Reliability Corp. 2006).

The automated analysis of data coming from other IEDs such as digital fault recorders (DFRs), sequence of event recorders (SERs), and SCADA remote terminal units

T. Popovic (✉)
XpertPower Associates, 1701 Southwest Parkway, Ste. 104,
College Station, TX, 77840 USA
e-mail: tomo@xpertpower.com

M. Kezunovic
Electrical and Computer Engineering, Texas A&M University,
TAMU 3128, College Station, TX, 77843 USA

B. Krstajic
Electrical Engineering, University of Montenegro,
Dz. Vasingtona bb, 81000 Podgorica, Montenegro

(RTUs) has been around for quite some time (Working Group D10 of the Line Protection Subcommittee - Power System Relaying Committee 1994; Kezunovic et al. 1994; MacArthur et al. 1995). Some work on utilization of DPR data has been reported in early 90s (Sun and Liu 1992). In practice today, DPR data is primarily used as non-operational data for after-the-fact analysis (Izylowski et al. 2007). With increasing monitoring capabilities of DPRs comes the opportunity to implement elaborate automated fault analysis. A good example of what can be done with DPR data is given in (Luo and Kezunovic 2005), but the given approach may be too specific and overly complicated to be practical.

This article focuses on the practical uses of DPR data as primary source for automated fault data analysis. The background section of the paper gives DPR characteristics, comparison to other IEDs, and outlines the goals for the DPR data analytics. The section discussing fault analysis based on DPR data addresses two practical options for automated processing of DPR data: a) by parsing the event reports; and b) by analyzing sample data obtained from DPR oscillography. The implementation section illustrates an automated solution realized in client/server paradigm. Finally, the testing and evaluation section provides the testing results and examples of the analytics reports.

2 Background

Besides their core control function, which is protection, the modern DPRs are often equipped with elaborate monitoring and recording capabilities that mimic digital fault and disturbance recorders (DFRs and DDRs). Some DPRs also have other functionalities such as sequence of event recorder (SER), fault location, and phasor measurement unit (PMU).

When triggered, substation IEDs capture signals in a small time window that typically contains a few cycles of the pre-fault and up to three dozen cycles of the post-fault data. These recordings consist of digital samples of multiple analog and status channels. A diagram of typical data sampling and processing in a modern IED is given in Fig. 1. An analog-to-digital (A/D) converter takes an analog signal and turns it into a binary number. Important criteria here is the accuracy of the A/D conversion related to the measuring range, and sampling rate, which can possibly affect the functionality that is based on the signal data (Brand et al. 2003). The vertical resolution of a n -bit A/D converter is a function of how many parts the maximum signal range can be divided into. The formula to calculate the resolution is 2^{n-1} . For example, an 8-bit A/D conversion has a resolution of $2^8 - 1 = 255$, which is 0.39 % of the full scale. Prior to the A/D conversion, the input signals are sampled using the sample-and-hold (S/H) circuit at the times

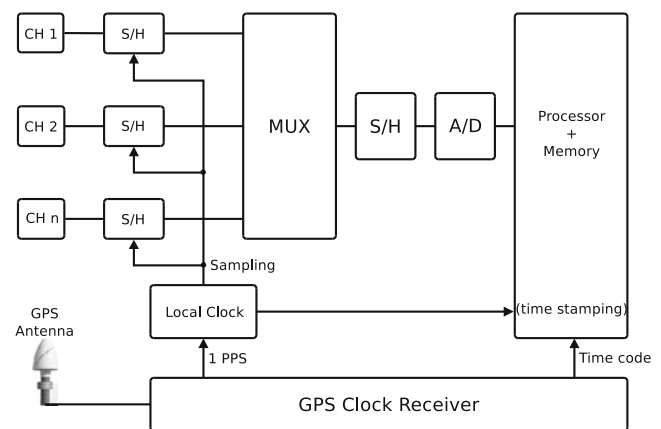


Fig. 1 Input channels data sampling and processing in a modern IED

defined by the sampling clock. Synchronous sampling of all the input signals allows correct determination and alignment of phase angles among different analog input signals. This can be accomplished either by using one A/D converter serving all channels but having separate S/H circuits on each channel and a multiplexer that feeds another S/H circuit in front of A/D conversion (see Fig. 1) or by using a separate S/H circuit and A/D converter on each channel. In addition, sampling can be synchronized between multiple IEDs by the external clock signal coming from GPS clock receivers (Lewandowski et al. 1999). Some older IED designs use a scanning A/D conversion method in which each channel is sampled and converted one at a time, causing a time skew among the corresponding samples on different channels. Besides the conversion process, the quality of the data is affected by wiring, input transformer characteristics, clock accuracy, internal signal propagation, non-linearity, anti-aliasing filters, and so on. When implementing and using the data analytics, it is important to understand the impacts of the quality of acquired measurements.

In this paper, the focus is on the event-triggered recording capability, which allows DPRs to produce oscillography recording. These recordings, together with event reports, can be automatically communicated, archived, and processed in the same fashion as it has been done for event recordings coming from DFRs (Kezunovic et al. 2000). There are some conceptual and performance differences that need to be taken into account when implementing solutions for automated data analytics based on DPR files. Typical DPR characteristics relevant to implementation of automated data analytics are:

- Signals are being continuously tracked, but the recording is triggered by events (pre-programmed conditions and thresholds).
- Sampling rate is relatively low, 4-16 samples/cycle, but in some instances may go even up to 32 samples/cycle.

- Synchronized vs. scanned channel sampling of internal A/D convertor is important feature allowing phase angle discrimination.
- Vertical resolution of the internal A/D conversion is at least 8 bits and in some instances 16 bits.
- DPRs typically monitor only the selected apparatus covered by the protection function but several DPRs can cover the entire substation.
- Signal conditioning and internal digital filters sometimes smooth the recorded waveforms too much, which may affect the processing and analysis results.
- Communication is either via serial or Ethernet link, which allows remote communication access to each device.
- Since the DPRs have a control function in the power system they are considered to be critical assets and require special attention with respect to cyber security (North American Electric Reliability Corp 2008).
- Time-stamping and GPS synchronization feature can be implemented if there is a need to correlate recorded files from different IEDs (Lewandowski et al. 1999).
- Applicable standards are well developed and include recommendations for settings format, data formats, communication, and cyber security.

The concept for automated substation data analytics systems is depicted in Fig. 2. Substation data integration is the foundation for systems automated data analytics (Popovic and Kezunovic 2012). The solution for automated DPR data analytics has to satisfy the following goals of this concept:

- Interfacing to DPR event files and their seamless integration with the data coming from other substation IEDs.
- Utilizing integrated DPR data and store the results back to the repository for easy access and re-use.

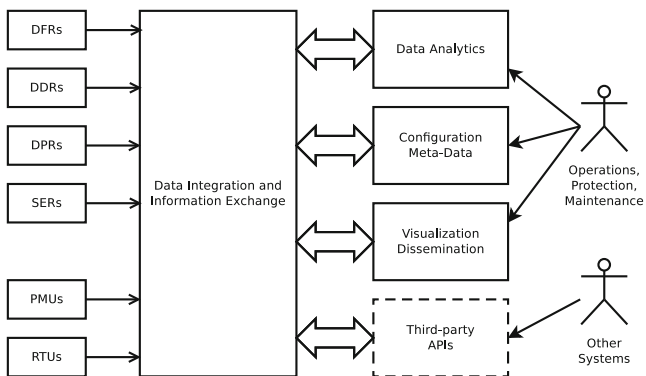


Fig. 2 Automated substation IED data analytics concept

- Handling of the configuration meta-data that enables correct semantic interpretation of the DPR data (channel assignments, scaling, transmission line parameters, etc).
- Providing for visualization and efficient dissemination of the analytics results.
- Allowing implementation of interfaces to third-party systems for efficient results transfer (for example to SCADA or GIS).

3 Fault analysis based on DPR data

DPR event files typically come in a vendor-specific file format. The waveforms are captured for a time window set around the event such as fault or disturbance. The event reports are based on the internal processing of the measurements captured during the event. In some cases, there is an option to automatically export the waveform files into COMTRADE file format (IEEE Inc 1999), but one still needs an access to report files to utilize results produced by the relays internal calculations.

There are two approaches to automated processing of DPR files as illustrated in the conceptual diagram in Fig. 3. The first approach is based on extracting the fault

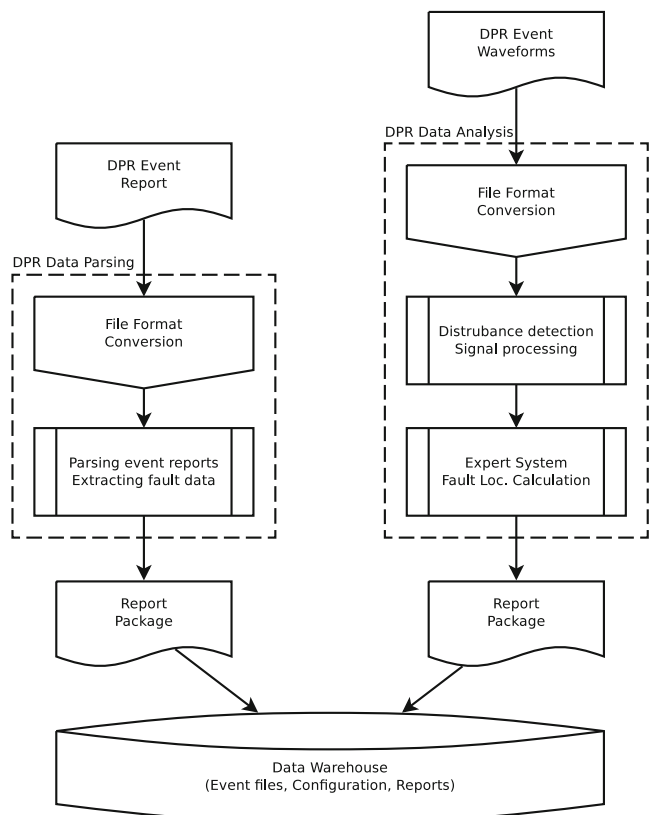


Fig. 3 Two approaches to automated DPR data analytics

event information by parsing DPR event reports. The second approach is to focus on the waveforms recording and process the recorded signals in the same fashion as it has been done for other IEDs such as DFRs. Both of these approaches end with producing a report package that contains data analytics report, typically a text file, and waveforms converted to standard file format such as COMTRADE. The following sections outline the implementation requirements for automated DPR data processing.

3.1 Communication and data collection

The communications protocols and data formats may be proprietary. Another challenge is the existence of multiple DPRs in a given substation, which may require integration of files across DPRs. Also, DPRs require elaborate cyber and physical security measures since they are considered to be critical assets. This makes the implementation of remote access functions such as automated event analysis more elaborate. While the implementation of communication solution is beyond the scope of this paper, it is important to have an automated data collection and DPR files stored on file servers. The file servers should be accessible to the solution using corporate network. It is desirable to get the files communicated quickly enough to satisfy target users. For on-line analysis, the target time for data collection should be few minutes within the event.

3.2 File format conversion

In order to implement meaningful automated processing and analysis of DPR files, they need to be automatically converted into a unified, non-proprietary, file format (IEEE Inc 1999, 2010, 2011). This process needs to be triggered by the occurrences of newly recorded files on the file servers. The conversion also needs to include extraction of the configuration data such as channel assignments and scaling, and setting parameters such as transmission line impedance and length. The converted files, stored in standardized format, will allow transparent implementation of the fault data analysis, visualization, and interfacing to other systems.

3.3 Signal processing

For each event the solution calculates signal features of ABC phases for current and voltage waveforms and relevant digital status signals. As the recordings contain pre-trigger, and post-trigger values, it is recommended to extract quantities that represent each time window. This can be done by calculating root mean square (rms) signal representative for each section using (1). N is number of samples in a 1-cycle wide data-frame, and $i(k)$ are discrete samples corresponding to continuous signal $i(t)$. For example, if we consider

phase-A current on a transmission line, the solution needs to extract and calculate the pre-fault, fault, and post-fault phasors that represent this signal Fig. 4. To achieve this, a mechanism for detection of the disturbance start and end times is needed. It can be implemented by checking the event reports generated by the relay or by processing the actual waveforms. When analyzing the waveforms, the system should check the changes in the current signals. Any dramatic jump in the phase or zero sequence currents should indicate the disturbance start. Similarly, towards the end of the file, the system looks for a significant drop in current signals to identify the disturbance end. DPR trigger time can also be used for the disturbance start. Once the disturbance start and end have been determined, the signal processing system calculates rms values and contact status for pre-disturbance, disturbance, and post-disturbance time regions. These quantities are to be passed into the expert system for further evaluation.

$$i_{rms} = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} i(k)^2} \quad (1)$$

3.4 Expert system

The expert system consists of a set of rules based on if-then logic and corresponding thresholds subsets. The rules and the thresholds map the knowledge about system faults into the patterns, which mimic the reasoning of human experts (Table 1). We successfully applied the approach described in (Kezunovic et al. 2000) for digital fault recorders on to the oscillography files obtained from digital relays. The rules utilize analog quantities, in this case calculated rms values, corresponding to pre-fault, fault, and post-fault for each signal. For example, an A-to-ground fault is expected to have jump in the phase A current, dip in the phase A voltage, unchanged B and C phase currents and voltages, and jump in the zero impedance current signals. The expert system evaluates all rules to determine if there was a fault, if the fault was a ground or line fault, and which phases were involved.

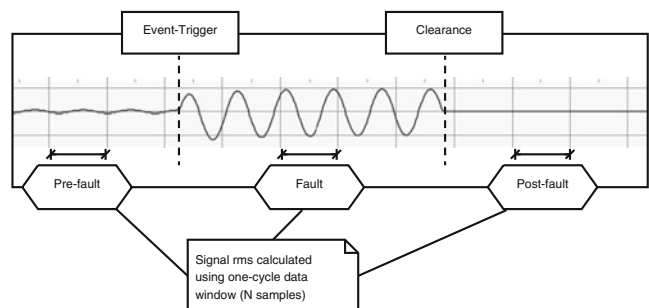


Fig. 4 Extracting phase current features with respect to the fault start and end

Table 1 Behavioral patterns of the current and voltage signals during faults in transmission lines

Fault Type	0 Sequence Current	Faulted Current	Unfaulted Current	0 Sequence Voltage	Faulted Voltage	Unfaulted Voltage	Line Voltage
A-G	$I_0 > 0.1I_{a,b,c}$	$I_a > 1.3I_{ap}$	$I_{b,c} < 0.5I_a$		$V_a < 0.95V_{ap}$	$V_{b,c} > 0.9V_n$	$V_{ab} \approx V_{ca}$ $V_{bc} \approx V_{bcp}$
AB	$I_0 < 0.05I_{a,b,c}$	$I_a > 1.5I_{ap}$ $I_b > 1.5I_{bp}$ $I_a \approx I_b$	$I_c < 0.25I_a$	$V_0 < 7V_{0p}$	$V_a < 0.9V_{ap}$ $V_b < 0.9V_{bp}$ $V_a \approx V_b$	$V_c > V_{a,b}$	$V_{ca} \approx V_{bc}$ $V_{bc} > V_{ab}$
AB-G	$I_0 > 0.1I_{a,b,c}$	$I_a > 1.4I_{ap}$ $I_b > 1.4I_{bp}$ $I_a \approx I_b$	$I_c < 0.5I_a$	$V_0 > 35V_{0p}$	$V_a < 0.9V_{ap}$ $V_b < 0.9V_{bp}$ $V_a \approx V_b$	$V_c > V_{a,b}$	$V_{ca} \approx V_{bc}$ $V_{bc} > V_{ab}$
ABC	$I_0 < 0.05I_{a,b,c}$	$I_f > 1.4I_p$ $I_a \approx I_b \approx I_c$		$V_0 < 9V_{0p}$	$V_f < 0.9V_p$ $V_a \approx V_b \approx V_c$		$V_{ab} \approx V_{bc} \approx V_{ca}$
ABC-G	$I_0 > 0.05I_{a,b,c}$	$I_f > 1.4I_p$ $I_a \approx I_b \approx I_c$		$V_0 < 9V_{0p}$	$V_f < 0.9V_p$ $V_a \approx V_b \approx V_c$		$V_{ab} \approx V_{bc} \approx V_{ca}$

With that knowledge the solution can automatically select the proper fault location calculation algorithm.

3.5 Fault location calculation

Once the fault has been detected and its type identified, it is fairly straightforward to implement single-end fault location calculation (Takagi et al. 1982). However, when using two-end algorithms it is essential that DPRs utilize GPS time synchronized time-stamping so that measurements from both ends of the line can be properly aligned. Since the relays provide pre-calculated phasor values, the two-end fault location can easily be implemented as long as the angles are corrected with respect to the time-stamps (2). The (V_R) is vector of voltage phasors at the remote end of the line. The δ is calculated based on the difference in GPS time-stamps referring to the start positions of the recordings obtained from DPRs at both ends of the transmission line.

$$V'_R = V_R e^{\delta} \tag{2}$$

After aligning the voltage and current measurements to precise time-reference, both single-end and two-end calculation algorithm can be successfully used (IEEE Inc 2005). When using the relay oscillography data, one needs to look for the relays with highest available sampling rate and vertical resolution of the internal analog-digital conversion.

3.6 Non-functional requirements and general considerations

There are some general considerations and recommendations that need to be observed:

- Time synchronization of recordings coming from multiple DPRs (and IEDs in general) is critical to an efficient analysis.

- Sampling rate for the recordings should be at least 16 samples/cycle or more. In modern DPRs this is typically a part of settings.
- The recordings used for automated signal analysis should contain raw and not filtered data.
- High speed communication should be used so that the analysis results can possibly be used in the decision making process. DPRs should be configured to use Ethernet or fastest serial option.
- The output results should use readable formats such as text files (XML, ASCII, HTML).
- To facilitate automated data analytics, power system component description should be provided using Substation Configuration Language (SCL) from IEC 61850 standard (International Electrotechnical Commission 2003).
- When designing a solution the data and configuration modeling should consider harmonization between 61850 and 61970 (CIM) standards (International Electrotechnical Commission 2002, 2003).

4 Implementation of the automated solution

The implementation utilizes well known client/server paradigm. Such an approach allows for various deployment configurations where there can be one or multiple clients and servers depending on the application needs. The processing client for automated processing of substation IED event data is illustrated in Fig. 5. The client scans DPR file storage for new event files. New files are being converted into unified file format. If the configuration settings metadata is available, the client performs automated processing and analysis of the signals. If the fault has been detected

and identified, the client performs the fault location calculation. In some instances the configuration meta-data may be included in DPR files themselves, depending on the model, vintage, and vendor. For such cases where the configuration is missing, there is still value in finalizing file format conversion and sending the data packet, without the analysis report, to the server.

The server side implements a universal data management as depicted in Fig. 6. The server scans the incoming folder for packages coming from processing client(s). When a new package is received, it is unpacked, parsed for results, and the results are then stored in to the event table in the centralized database. The data manager also checks for availability of DPR (or other IED) data from the remote end, and if data is available, performs two-end fault location calculation. For DPRs, if possible, it is good to pair same model and vintage of DPRs on both ends of transmission lines so that equal quality of the data can be expected. The outcome of the server-side processing and two-end fault location calculation is stored into the centralized data warehouse. For critical events, depending on the settings, the data

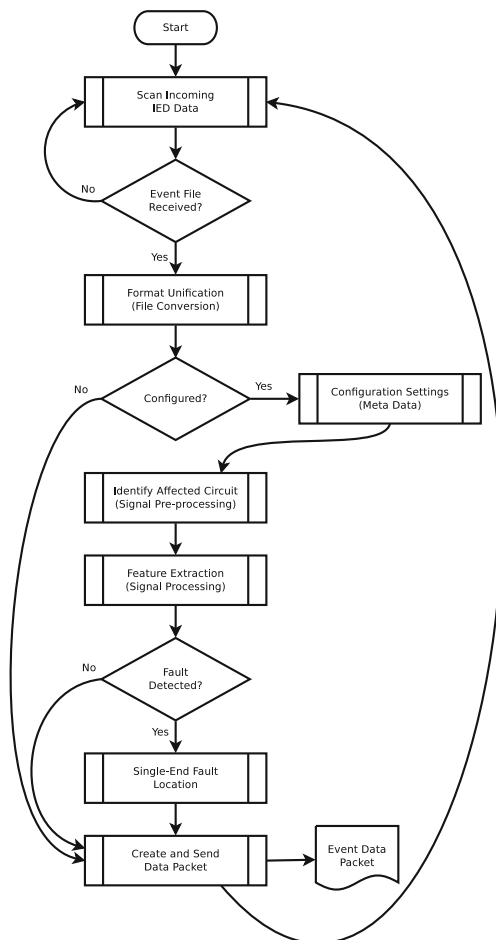


Fig. 5 Client implementation for the automated analysis of DPR files

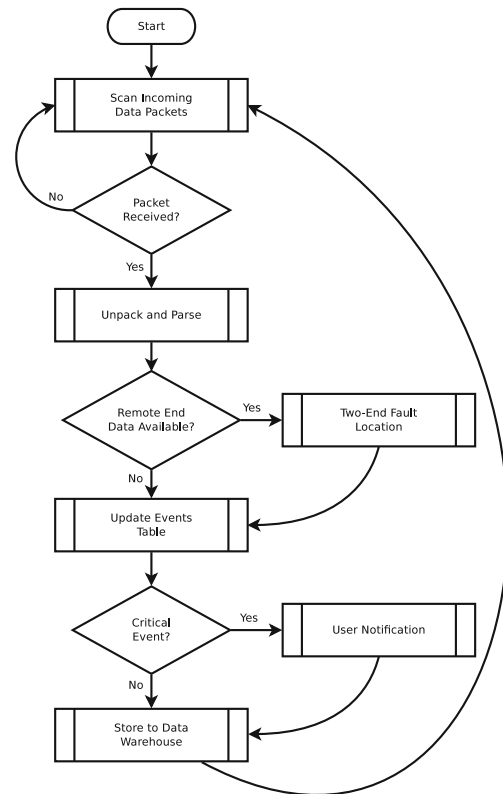


Fig. 6 Data manager implemented on the server side

manager sends out user notifications in form of emails, pager messages, and printouts.

The solution satisfies the requirements to achieve: a) transparency in substation IED data obtained from DPRs; b) data analytics functions agnostic to event data source; and c) universal access to output results that can be used and re-used. Satisfying these conditions enables a seamless integration with other IED data analytics. A good illustration of value in the transparent approach is the use of a universal waveform and report viewer as illustrated in Fig. 7. The viewer can open any COMTRADE file or package containing waveforms and analytics results regardless of what model, type, or vintage of the DPR is being used. The results are archived in the same fashion as it has been done with DFR and other IED data.

5 Testing and evaluation

The solution described in this paper has been tested in various scenarios. All of these scenarios fall into two groups: a) in-house testing using simulated fault data, and b) evaluation using the data recorded in substations. For in-house testing an electromagnetic transient program (EMTP) called ATP has been used (Alternative Transient Program 2012).

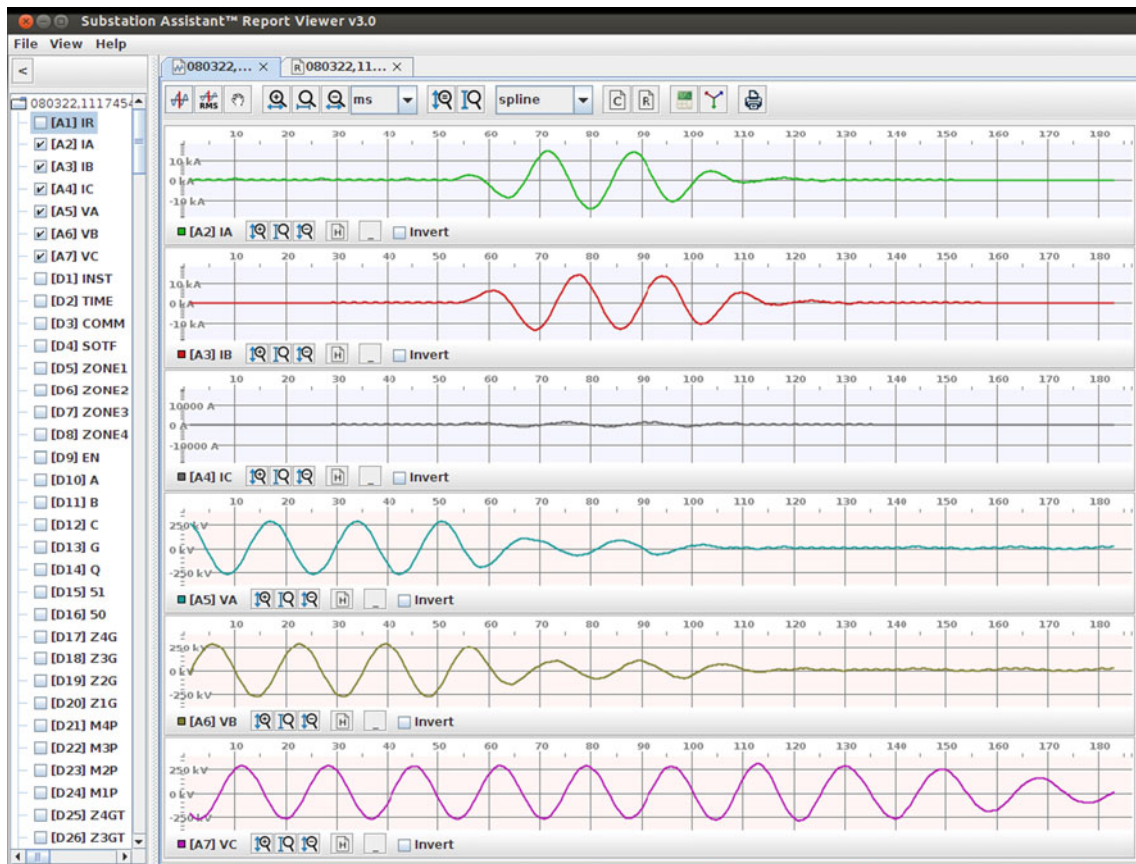


Fig. 7 Transparent access: DPR oscillography in the universal viewer

An example 9-bus system model that corresponds to a section of a real power system is shown in Fig. 8. This model has been verified and calibrated using field measurements and DFR recordings. It has been used to simulate four sets

of fault events. The parameters varied in the simulation were fault type (A-G, AB, AB-G, ABC), location (50 to 95 % in steps of 5 %), and fault resistance (0.01, 1, and 10 Ω). The output files from EMTP simulation were converted to

Fig. 8 An example 9-bus system model used for EMTP simulations

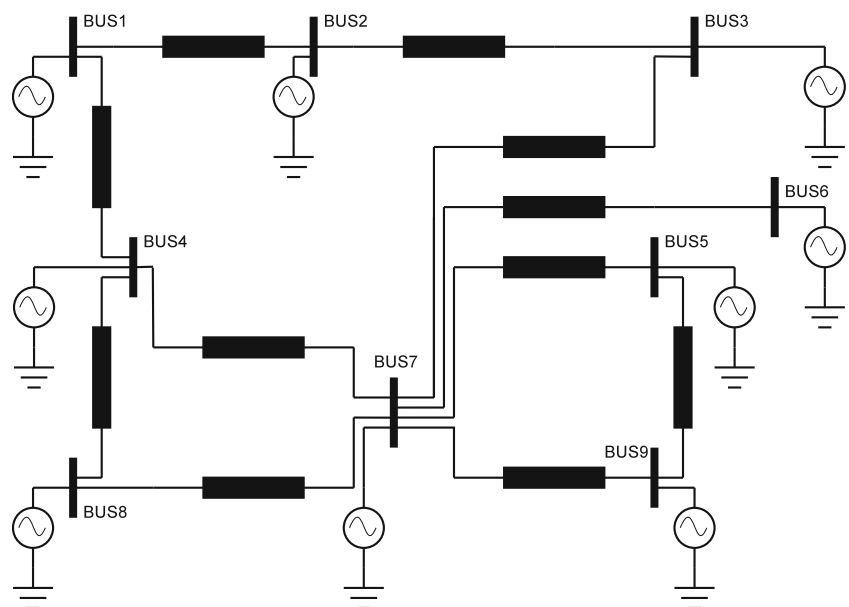


Table 2 Evaluation with EMTP simulated fault data

#	Fault type	Num. of cases	Detection/Classif.	Single-end error [%]
1	A-G	30	30/30	0.61–3.75
2	AB	30	30/30	0.59–2.57
3	AB-G	30	30/30	0.61–2.57
4	ABC	30	30/30	0.42–2.57

COMTRADE and presented to the data analytics. In all cases the fault detection correctly identified the fault type and disturbance start and end times. The calculated fault locations were fairly accurate considering a single-end algorithm was used. The test cases and fault location calculation results are summarized in Table 2.

For few simulated cases, COMTRADE files were run through an open-loop simulator, and fed via D/A conversion and amplifiers into the wired relays. It was verified that relays operate as expected, and produce waveform recordings and event reports. These files were used for another pass of the analytics evaluation and allowed us to perform both parsing and event analysis approaches.

Besides simulated files, the evaluation has also been done using actual field recordings collected from SEL-421 and GE D60 distance relays (Schweitzer Engineering Laboratories 2012; General Electric 2012). The arrangement was that on each end of the transmission line there is one GE and one SEL relay (primary and backup). The event files were retrieved manually for the purpose of testing.

In few instances, there was two-end waveform data available so the testing included two-end fault location as well. Table 3 illustrates examples of such results using SEL-421 digital relays. The table shows a comparison between single or two-end calculation, performed by automated data analytics, and fault information parsed from digital relay reports. In all cases the fault type was correctly detected and matching with the information from the relays. It is important to note that the two-end data obtained from SEL relays in this cases shows very good alignment due to the use of GPS synchronized clocks on both ends of the transmission

Table 3 Example evaluation with field data (SEL relays)

#	Substation Bus	Single-end Analytics	Two-end Analytics	SEL Relay Parsing
1	Bus A	8.7	8.2	8.75
	Bus B	0.3	0.2	0.16
2	Bus A	8.4	8.6	8.61
	Bus B	0.1	−0.2	0.05
3	Bus A	8.4	8.6	8.63
	Bus B	0.0	−0.2	0.05

```

*** Relay Event Summary ***
DPR Assistant Client: Demo
Substation: Sub2 DPR Line 3 BU
DPR Native File Name: Relay BU, XX0720.EVE
Affected Circuit: Line 3 BU
Trigger Date and Time: 20XX-07-20 11:59:34.307
Start Date and Time: 20XX-07-20 11:59:34.207
Event Description: AGND_FAULT
Fault Location: 24.93 [Miles]
Event Outcome: -

*** Native File ***
...
    
```

Fig. 9 Data analytics results by parsing DPR event reports

line. The time difference for the event trigger was under 10 ms. The highest available sampling rate was used and a very good match between calculated values and SEL reports is observed. The universal report and event viewer is utilized to enable easy inspection and manual analysis of records. Combining substation data provides additional redundancy and improves reliability of the decision making process.

The illustration of what kind of DPR data analytics results can be obtained using the two approaches are shown in the report examples. Figure 9 provides a report generated by parsing the DPR event report file, which is an actual A-G fault recorded in a real system. Figure 10 illustrates another field example, but in this case the results are obtained using the analysis of the DPR oscillography file.

```

*** Expert System Log ***
The bus breaker is not monitored!
The middle breaker is not monitored!
Line breaker(s) open after the disturbance!
The event is a ground fault!
The event is a line A to line B to ground fault!
The fault is cleared by the protection at this substation!
Primary relay is not monitored!
Backup relay is not monitored!

*** Event Origin ***
DPR Assistant Client: Demo
Substation: Sub4 DPR Line 1 PRI
DPR Native File Name: Sub4 PRI
Affected Circuit: Line to Sub5

*** Event Summary ***
Trigger Date and Time: 20XX-03-22 11:17:45.431
Start Date and Time: 20XX-03-22 11:17:45.431
Disturbance Start, End: 72, 99 [ms]
Duration: 27 [ms]
Event Description: ABGND_FAULT
Fault Location: 6.9 [Miles]

Event Outcome: EV_CLR_LOCAL
Breaker Operation: 1st, CB_OK
Breaker Operation: 2nd, CB_OK
Relay Operation: PRIM, RL_NOT_MONITORED
Relay Operation: BACK, RL_NOT_MONITORED

*** Analog signal Values ***
Prefault Values:      Fault Values:      Postfault Values:
IO = 0.0151 [kA]     IO = 8.5445 [kA]   IO = 0.0011 [kA]
Ia = 0.1711 [kA]     Ia = 8.6328 [kA]   Ia = 0.0010 [kA]
Ib = 0.1523 [kA]     Ib = 9.4936 [kA]   Ib = 0.0007 [kA]
Ic = 0.1286 [kA]     Ic = 0.6440 [kA]   Ic = 0.0014 [kA]

V0 = 0.4069 [kV]     V0 = 113.4833 [kV] V0 = 190.9335 [kV]
Va = 199.1539 [kV]   Va = 48.5382 [kV]  Va = 3.4481 [kV]
Vb = 199.9901 [kV]   Vb = 62.7262 [kV]  Vb = 4.4166 [kV]
Vc = 200.5269 [kV]   Vc = 195.0992 [kV] Vc = 191.1003 [kV]

Vab = 345.0298 [kV]  Vab = 74.7612 [kV] Vab = 0.9702 [kV]
Vbc = 347.7242 [kV]  Vbc = 243.6859 [kV] Vbc = 191.3641 [kV]
Vca = 345.9049 [kV]  Vca = 234.9727 [kV] Vca = 191.2469 [kV]
...
    
```

Fig. 10 Data analytics results by analyzing DPR event oscillography

6 Conclusions

The use of automated intelligent systems for processing data obtained from the power system substations is the key to addressing the challenge of the large-scale deployment of various IEDs. Such solutions when combined with an efficient substation data collection scheme can be successfully used to quickly, within minutes, provide additional knowledge needed to operate the system in real-time. Traditionally considered non-operational data such as DPR event records can be used as operational and help the decision-making process for the system operations. The following is the list of key contributions of this paper:

- The article outlines main characteristics of DPR data and how it may be used to facilitate an automated fault data analytics. It provides the requirements for basic automated data analytics and it discusses similarities and differences between DPRs and other substation IEDs.
- The issues related to communication, sampling rate, internal filtering, and time-stamping are identified as posing challenges to the use of DPR data for automated analysis.
- Design specifications for the following solutions for automated fault data analytics are provided: a) parsing of relay event reports; and b) performing intelligent event analysis on waveform recordings. Both approaches can be combined together, as well as with other substation IED data analytics to achieve better redundancy.
- The presented solution allows for seamless integration with automated data analytics of the data coming from other IEDs, namely the DFRs. The integrated data and results can be viewed using the universal report viewer.
- In-house testing and evaluation results for both approaches to DPR data analytics are illustrated by providing the actual analysis reports. The solution was tested with the simulated and field DPR data. The simulated data was generated using digital power system simulator and transient simulation program.

Acknowledgments The authors wish to acknowledge support from the individuals providing the field data examples and enabling us to perform initial tests with on-site pilot installations:

Mr. Scott Sternfeld who was at the time with FirstEnergy, Akron, Ohio; and Ms. Manjula Datta-Barua and Mr. Brian Clowe at Center-Point Energy, Houston, Texas.

References

Alternative Transient Program (2012). Resource document. <http://www.emtp.org>. Accessed 18 Dec 2012.

- Brand, K.-P., Lohmann, V., Wimmer, W. (2003). *Substation automation handbook* (pp. 97–101). Bremsgarten: Utility Automation Consulting Lohmann.
- Costello, D. (2000). Understanding and analyzing event report information. Resource document, Schweitzer Engineering Laboratories Inc., Pullman. <http://www.selinc.com>. Accessed 18 Dec 2012.
- General Electric (2012). D60 Line Distance Protection System. Resource document. GE Digital Energy. <http://www.gedigitalenergy.com>. Accessed 18 Dec 2012.
- IEEE Inc. (1999). Common format for transient data exchange (COMTRADE), *IEEE Standard C37*, 111–1999.
- IEEE Inc. (2005). IEEE guide for determining fault location on AC transmission and distribution lines. *IEEE Standard C37*, 114–2004.
- IEEE Inc. (2010). Common format for event data exchange (COMFEDE) for power systems. *IEEE Standard C37*, 239–2010.
- IEEE Inc. (2011). Common format for naming time sequence data files (COMNAME). *IEEE Standard C37*, 232–2011.
- International Electrotechnical Commission (2002). Common interface model (CIM). *IEC Standard*, 61970–301.
- International Electrotechnical Commission (2003). Communication networks and systems in substations. *IEC Standard* 61850-5, ed1.0.
- Izylowski, J., Rosolowski, E., Saha, M.M. (2007). Postfault analysis of operation of distance protective relays of power transmission lines. *IEEE Transactions On Power Delivery*, 22(1), 74–81.
- Kezunovic, M., Rikalo, I., Fromen, C.W., Sevcik, D.R. (1994). Expert system reasoning streamlines disturbance analysis. *IEEE Computer Applications in Power*, 7(2), 15–19.
- Kezunovic, M., Liu, C.C., McDonald, J.R., Smith, L. (2000). Automated fault analysis, IEEE PES tutorial. Resource document. *IEEE Power and Energy Society*.
- Lewandowski, W., Azoubib, J., Klepczynski, W.J. (1999). GPS: primary tool for time transfer. *Proceedings of the IEEE*, 87(1), 163–172.
- Luo, X., & Kezunovic, M. (2005). An expert system for diagnosis of digital relay operation. In *13th international conference on intelligent systems application to power system* (pp. 1–6). Washington.
- MacArthur, S.D.J., McDonald, J.R., Bell, S.C., Burt, G.M. (1995). Expert systems and model based reasoning for protection performance analysis. *IEE Colloquium on Artificial Intelligence Applications in Power Systems*.
- McDonald, J.D. (2003). Substation automation, IED integration and availability of information. *IEEE Power and Energy*, 1(2), 22–31.
- North American Electric Reliability Corp. (2006). Protection and Control, PRC-002-1 and PRC-018-1. *NERC Standard*, 22–44.
- North American Electric Reliability Corp (2008). Cyber Security CIP-002 through CIP-009. *NERC Standard*.
- Popovic, T., & Kezunovic, M. (2012). Measures of value: data analytics for automated fault analysis. *IEEE Power and Energy*, 10(5), 58–69.
- Sage, A.P., & Rouse, W.B. (1999). Information systems frontiers in knowledge management. *Information Systems Frontiers*, 1(3), 205–219.
- Schweitzer Engineering Laboratories (2012). SEL-421 Protection, Automation, and Control System. Resource document, Schweitzer Engineering Laboratories Inc., Pullman. <http://www.selinc.com>. Accessed 18 Dec 2012.
- Sun, Y., & Liu, C.C. (1992). RETEX (Relay, Testing, EXpert): an expert system for analysis of relay testing data. *IEEE Transactions on Power Delivery*, 7(2), 986–994.
- Takagi, T., Yamakoshi, Y., Yamaura, M., Kondow, R., Matsushima, T. (1982). Development of a new fault locator using the one-terminal

voltage and current data. *IEEE Transactions on Power Apparatus and Systems*, 101(8), 2892–2898.

U.S.-Canada Power System Outage Task Force (2004). Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations. Resource document. <https://reports.energy.gov/BlackoutFinal-Web.pdf>, Accessed 18 Dec 2012.

Working Group D10 of the Line Protection Subcommittee - Power System Relaying Committee (1994). Potential applications of expert systems to power system protection. *IEEE Transaction on Power Delivery*, 9(2), 720–728.

Tomo Popovic received his Dipl. Ing. degree in 1994 from the University of Novi Sad, Serbia, and the Master's degree in 2005 from the Texas A&M University, College Station, Texas, both in electrical engineering. He has been with Texas A&M University since March 2012. He is a consulting associate with XpertPower Associates, where he was a Sr. Systems and Software Engineer since 1998. His research interests include software development and application of smart information systems to power systems monitoring and event analysis. His main professional interest are developing and implementing software and hardware solutions for industrial applications, especially in the field of electric power system engineering. Mr. Popovic is a Senior Member of IEEE and member of CIGRE.

Mladen Kezunovic received the Dipl. Ing., MS and PhD degrees in electrical engineering in 1974, 1977 and 1980, respectively. Currently, he is the Eugene E. Webb Professor at Texas A&M University, Site Director of NSF I/UCRC Power Engineering Research Center, PSerc, and Deputy Director of another NSF I/UCRC Electrical Vehicles: Transportation and Electricity Convergence, EV-TEC. His main research interests are digital simulators and simulation methods for relay testing, as well as application of intelligent methods to power system monitoring, control, and protection. He has published over 400 papers, given over 100 seminars, invited lectures and short courses, and consulted for over 50 companies worldwide. He is the Principal of XpertPower Associates, a consulting firm specializing in power systems data analytics. Dr. Kezunovic is a Fellow of the IEEE, a member of CIGRE and registered Professional Engineer in Texas.

Bozo Krstajic received his Dipl. Ing. degree, the MS and PhD degrees from the University of Montenegro, Montenegro, all in electrical engineering, in 1992, 1996 and 2002, respectively. He is Associated Professor on Department of Electrical Engineering and Director of Center for Information Systems on University of Montenegro. He is also a Manager at the local CISCO Academy. His main research interests are adaptive signal processing, system modeling and control. His professional interests include design, deployment, and maintenance of information systems and computer networks. Dr. Krstajic is a member of IEEE.