

# Automated analysis of power systems disturbance records: smart grid big data perspective

Ukil, Abhisek; Zivanovic, Rastko

2014

Ukil, A., & Zivanovic, R. (2014). Automated analysis of power systems disturbance records: smart grid big data perspective. IEEE Innovative Smart Grid Technical Conference., ISGT ASIA 2014. Kuala Lumpur, Malaysia.

<https://hdl.handle.net/10356/96767>

<https://doi.org/10.1109/ISGT-Asia.2014.6873776>

---

© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [<http://dx.doi.org/10.1109/ISGT-Asia>]

*Downloaded on 01 Mar 2021 00:53:08 SGT*

# Automated Analysis of Power Systems Disturbance Records: Smart Grid Big Data Perspective

Abhisek Ukil, *Senior Member, IEEE* (corresponding author)  
School of Electrical & Electronic Engineering  
Nanyang Technological University, Singapore  
Email: aukil@ntu.edu.sg

Rastko Zivanovic, *Member, IEEE*  
School of Electrical & Electronic Engineering  
University of Adelaide, Adelaide, Australia  
Email: rastko@eleeng.adelaide.edu.au

**Abstract**—Analysis of faults and disturbances play crucial roles in secure and reliable electrical power supply. Digital fault recorders (DFR) enable digital recording of the power systems transient events with high quality and huge quantity. However, transformation of data to information, expectedly in an automated way, is a big challenge for the power utilities worldwide. This is a key focus for realizing the ‘Smart Grid’. In this paper, the architecture and specifications for the primary and the secondary information for the automated systems are described. This provides qualitative and quantitative guidelines about the information to derive out of the disturbance data. A quantified estimate of big data for the substations, has been estimated in the paper. Possible ways of reducing the big data by utilizing intelligent segmentation techniques are described, substantiated by real example. Utilization of centralized protection and remote disturbance analysis for reducing big disturbance data are also discussed.

**Index Terms**—Analytics, big data, centralized protection, fault analysis, high voltage, HV, medium voltage, MV, radial network, remote relay testing, smart grid, web service.

## I. INTRODUCTION

For secure and reliable electrical power supply, analysis of faults and disturbances is crucial. This is becoming even more important in the ‘Smart Grid’ initiative, with focus on blackout prevention, fast system restoration [1], better estimation of fault location, integrating communication in protection, condition monitoring and diagnostics, etc.

Application of the digital recording technology in the power systems transient events results in high quality and huge quantity of disturbance records. Information from the analysis of digital records are required to get an insight into the behavior of the power system as well as the performance of the protection equipments. However, manual analysis of the records is both time-consuming and complex. One of the major challenges in the ‘Smart Grid’ initiative is to automatically convert data into knowledge. This will allow more time and flexibility of the human resources to implement corrective or preventive action [2].

Secondly, the transformation of the power systems disturbance data to information in an automated way is an even bigger challenge for the power utilities worldwide. The huge amount of disturbance data are practically never analyzed. Therefore, with intelligent techniques if this huge amount of data could be reduced to meaningful information at source,

that is of great interest. This paper will highlight the nature and source of such big data, and possible ways to reduce it.

The remainder of the paper is organized as follows. Section II describes the current best practices in power systems disturbance analysis, e.g., protection devices, characteristics of those, major manufacturers and common mode standard. Section III describes the automated analysis system’s architecture, specifications for primary and secondary information, needed for deriving a complete and meaningful information. Section IV discusses about the quantified source of the big disturbance data and presents a signal processing-based method on reducing it, with discussions on centralized protection and remote disturbance analysis. Discussions are mentioned in section V, followed by conclusions in section VI.

## II. POWER SYSTEMS DISTURBANCE ANALYSIS

### A. Protection Devices

Historically, the following types of equipments 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)
- Fault locators (FLs).

However, most of these different protection devices are now increasingly being merged into the DFRs, as different functionalities into one device. Major power utilities are equipped with the DFRs (also referred as IEDs) on the feeder bays, with an additional few installed on the static var compensators (SVCs).

The DFRs trigger due to reasons such as power network fault conditions, protection operations, breaker operation, etc. Most of these devices are remotely accessible via X.25 communication systems [3]. In future, communication standards like IEC 61850-9-2 [4]–[6] are expected to enhance the communication, data availability more flexibly. Besides standard Internet-based communication, potential utilization of the wireless LAN with IEC 61850 [7] for substation and distribution automation applications is under investigation.



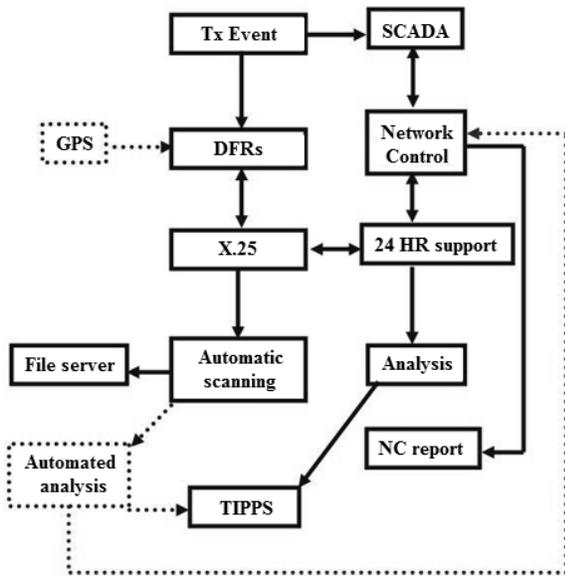


Fig. 2. Typical disturbance analysis scheme employed at network control centers.

As per Fig. 2, when a transient event (Tx-event) takes place in the power transmission network, the NC gets informed via a SCADA (Supervisory Control And Data Acquisition) network and also the DFRs will be triggered. The protection engineers provide a 24-hour support through the analysis of these records.

It is possible to download the fault record via the X.25 communication network. The protection engineer has to do manual analysis to create a report for the NC, which is cumbersome. It takes a lot of time, typically from 1 hour to 10 hours or more, depending on the complexity and severity of the event. Additionally, an automatic scanning PC downloads the records. The records need to be saved manually at a file server. Results from the analysis of the digital fault records are also captured in the transient information processing and protection systems (TIPPS) database [2]. Due to large amount of historical data available, trend analysis on the performance of primary and secondary equipment can be done.

An automated analysis system (AAS), shown in dotted lines in Fig. 2, would be much faster. It could be placed in two different ways. The more efficient way is to do the automated analysis directly at the substation, transmitting only a report to the NC. This is the fastest way to do it but additional PCs and communication infrastructure are needed. Standards like IEC 61850-9-2 [4] would facilitate more communications in future. The other way is to run the automated analysis on a central PC. The automated analysis system should also be able to write the analysis results to the TIPPS database.

### B. Specification: Primary Information

The first requirement for the AAS is that it must be able to read a COMTRADE file. The AAS must be able to automatically import the COMTRADE file from a user defined directory, do the analysis and produce the results in a format

which can be viewed by any text editor. The distribution of the results should be done through any of the following media: email, SMS, print or Web. The AAS must extract from the COMTRADE file the following information.

- Faulted phase(s),
- Fault type,
- Total fault duration,
- Main 1 protection operating time,
- Main 2 protection operating time,
- Fault location,
- Fault resistance,
- DC offset,
- Breaker operating time,
- Auto re-close time.

### C. Specification: Secondary Information

Besides the above mentioned parameters, the AAS must also determine and report on the following information.

- Was the fault on the specific feeder?
- Did the main 1 relay operate?
- Did the main 1 relay operate correctly?
- Did the main 2 relay operate?
- Did the main 2 relay operate correctly?
- Was the main 1 permissive carrier signal sent, and was this correct?
- Was the main 2 permissive carrier signal sent, and was this correct?
- Was the main 1 permissive carrier received, and was this correct?
- Was the main 2 permissive carrier received, and was this correct?
- How did the breaker operate, 1 pole or 3 poles?
- Did the breaker auto re-close (ARC)?
- What was the magnitude of the fault current?
- What was the magnitude of the neutral current?
- What was the magnitude of the healthy phase currents during the fault?
- What was the depression in voltage on the faulted phase/s?
- What was the depression in voltage on the healthy phase/s?
- What were the dominant frequencies before and during the fault?
- Did any of the breaker poles re-strike?

## IV. BIG DATA: AMOUNT & REDUCTION

### A. Size of Disturbance Big Data

The total size of the disturbance data basically depends on the typical COMTRADE file size. The file size mainly depends on the sampling frequency and length of the records. Current power systems disturbance data communication standard [4] allows for maximum of 256 samples/cycle. For the maximal systems frequency of 60 Hz, this will translate to 15.36 kHz for possible maximum sampling frequency. For typical data recording for 5s, it would provide 76800 samples. There will be additionally overhead information as described in section

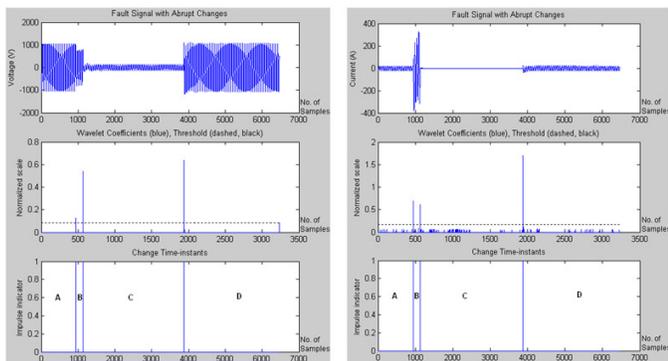


Fig. 3. Wavelet transform-based segmentation of the voltage (left) and current (right) signals during power systems disturbances.

II.D. This typically leads to about 1–3 megabytes (MB) of file size for each data record. There could be several thousand disturbance records in a particular substation. Depending on the protection settings, many relays are placed. If the relays typically record about 100 records in a particular time period, and there are 1000 relays (which will cover relatively reasonable area), this will lead to  $3 \times 100 \times 1000 = 300,000$  MB, i.e., about 300 GB of data for a typical substation area (please note this is an empirical rough estimate). This figure multiplied by the exact number of DFRs in a substation and total number of substations would provide an idea about the total amount of data being generated.

### B. Steps to Reduce Big Data

The first step in such disturbance analysis is to identify the fault. Here, a fault generally means an overcurrent [1], mainly from a short circuit.

In power systems, the protection follows sequential steps. For example, there is normal prefault condition, then fault inception takes place, which is followed by fault identification. If the fault is persistent, CB opening is done. As many disturbances in power systems are rather temporary and self-clearing, an auto-reclosing of the CB is done after pre-defined period, which might result in successful system restoration or opening of CB again due to permanent fault. The DFR recordings typically cover the entire sequence of events.

The first step in automated analysis is to perform a segmentation of the disturbance signals according to the sequential protection events. There are different approaches to achieve the segmentation, using abrupt change detection methods [20],[21]. Amongst these, the wavelet transform-based method [2], [10]–[18], and system identification-based method [19],[21] are particularly useful. However, in power systems disturbance analysis, the speed of operation is very important. From that perspective, the wavelet transform performs much better than the system identification approach [10],[11],[19].

Once the segmentation is done, as in Fig. 3, the fault section can be considered instead of the whole signal. For example, in Fig. 3, the total signal length is 6500 samples. However, for subsequent analysis, one would be interested in the fault section, which is segment B, about 200 samples.

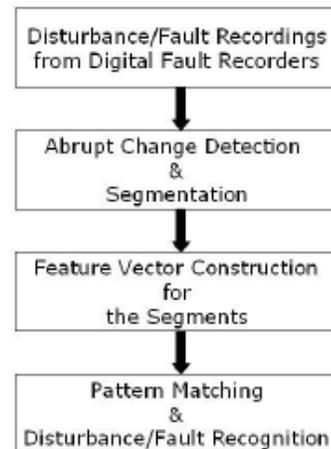


Fig. 4. Architecture of the Automated Analysis System.

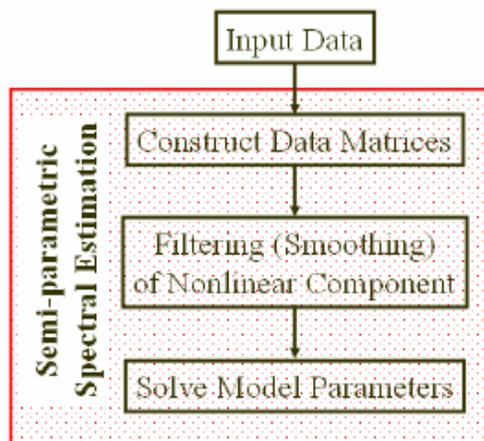


Fig. 5. Semi-parametric approach based feature vector construction.

Therefore, instead of saving the whole signal, the segment information and sections of particular interest could be saved. For example the long part C (after CB opening), which is about 1100 samples, could be removed and only the beginning and end point of segment C could be kept. This will reduce the disturbance file size by 20%. Further intelligent reductions could be achieved.

Then, we would construct the appropriate feature vectors for the different segments, e.g., using semi-parametric method [22],[23]. And finally the pattern-matching algorithms would be applied using those feature vectors to accomplish the fault recognition and associated tasks [2],[3]. The flowchart is shown in Fig. 4, and the semi-parametric feature vector construction in Fig. 5. These steps would contribute to derive the required information, as described in section III.B–C.

### C. Centralized Protection

In power distribution systems, radial and ring-main sub-transmission systems are commonly used topology. Typically, a nondirectional time-delay protection logic is used to protect

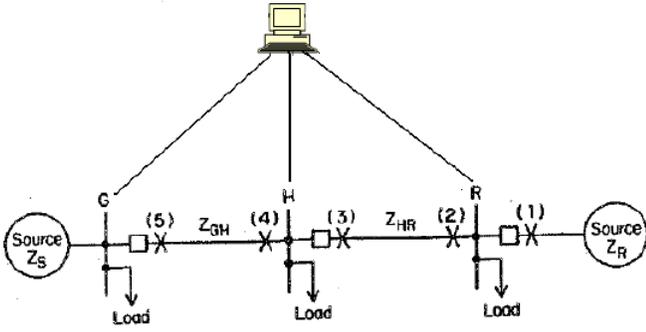


Fig. 6. Centralized protection in distribution systems.

the line. But it will be difficult to identify effectively the affected zones only and accordingly cut out minimal sections, without interrupting customers who are outside the faulty zones.

In comparison, effective protection in terms of optimal section breaking could be offered with current-only directional relays [24]–[26]. Fig. 6 shows the distribution topology with centralized protection scheme. In case of faults, the relays at G, H, R can communicate the directional judgment relative to their own position to a central computer (substation) or controller unit. The decision levels are ‘Forward as 1’ (fault between source to relay point), ‘Reverse as -1’ (fault between relay point to line), ‘Neutral’ (e.g., outside directional sensitivity) or Normal as 0.

The central computer or unit has to identify which of the contiguous relays provide contradicting decisions. The fault section would be between those. This way, we can ensure optimal fault sectionalization which is only possible with the position-aware sensing by the novel current-only directional relays [24]–[26] along with the decision logic.

The centralized protection scheme would leverage the IEC 61850 standard [4] for digital substation. Transmitting the disturbance data to the central computer would increase the system reliability. It would also reduce the big data, as data accumulation and analysis could be done concurrently, with the help of the central computing resource. This is currently not possible with the nodal and local computations using the relays, which have limited computational power.

#### D. Remote & Distributed Disturbance Analysis

The big data from power systems disturbances could be effectively reduced by performing the analysis remotely in a distributed manner. This could be effectively realized using the application service provider (ASP) technologies [27],[28]. Fig. 7 shows the schematics of the remote disturbance analysis using the ASP service.

The automatic analysis system could be available for use via the ASP technology over the Internet in two modes. In the Indirect mode of use, users will need to upload their fault signals via the Internet into the analysis system and will get back the analysis report via the Web and e-mail. In the Direct mode of use, users will be able to use the analysis system

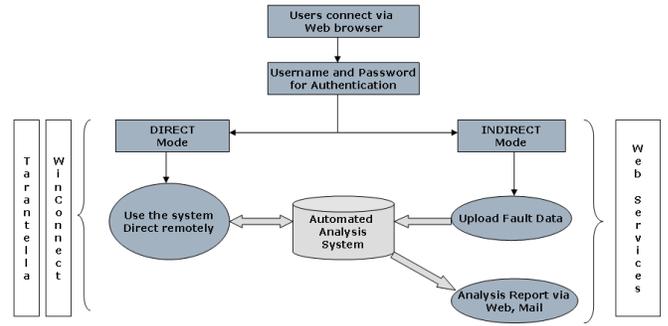


Fig. 7. Remote disturbance analysis using application service provider (ASP) backbone.

remotely (hosted in a central server computer) via the Internet after logging into the system using their login id and password [27],[28].

It will be required to web-enable and integrate various systems in the indirect mode, e.g., the fault analysis engine, the web interface, feedback system etc which are cross-platform from the point of view of operating system and programming languages. So, the Web service would be the perfect technology to integrate and implement the indirect mode of the ASP of the automatic fault analysis system [27],[28].

Using the ASP, low level relay testing could be done. Test results are reported in [28]. As the disturbance analysis is done remotely and in a distributed manner, this scheme would effectively reduce the large volume of disturbance data. Besides, it will effectively create valuable information from data, which most of the power utilities are lacking today.

## V. DISCUSSIONS

The following comments are cited.

- 1) The intelligent signal processing techniques, e.g., the segmentation and selection would need to be applied in an embedded manner in the DFR, so that the disturbance record file size can be reduced.
- 2) Further file size reduction could be done by computing the parameters required in the automated analysis as specified in the sections III.B–III.C. Those values could be stored in the DFR and communicated accordingly.
- 3) For communicating reduced and processed information, some adjustments in the standard would be necessary. This is because, currently standard like IEC 61850-9-2 [4] defines requirements for the sample values only, not for processed or modified information. However, the nature of the processed information, as demonstrated in this paper, is not critically different. It is just rejecting unnecessary parts of the complete records. Therefore, existing standard like IEC 61850-9-2 [4] is well applicable. The reduced data would require reduced communication bandwidth as well.
- 4) As the change in modus operandi would require significant computation, the DFRs would need more computational resources, e.g., processor speed, etc. On the other

hand, less storage would be required. And subsequently, the problem of huge amount of data residing at the network control centers, which never practically gets analyzed, could be improved.

- 5) Besides storing the computed information, those could be communicated to the operating and the maintenance personnel via sms, email, etc.
- 6) Besides the signal processing-based methods presented in this paper, data compression techniques could be applied to compress the data [29]. However, for disturbance analysis that might not be preferred. Also, the interoperability of protective devices by different vendors with the compressed data might be a problem.

## VI. CONCLUSION

The analysis of faults and disturbances plays a central role for a secure and reliable electrical power supply. This is becoming even more important in the ‘Smart Grid’ initiative, with focus on blackout prevention, fast system restoration, better estimation of fault location, etc. DFRs enable digital recording of the power systems transient events with high quality and huge quantity. The transformation of the power systems disturbance data to information, expectedly in an automated way, is a big challenge for the power utilities worldwide, but a key focus for realizing the ‘Smart Grid’.

The architecture and specification for the primary and the secondary information for the automated systems are described in this paper. This will provide a guideline what to derive out of the big data. Furthermore, a quantified estimate of big data, typically in the order of 100–300 GB for a particular substation, has been estimated empirically in the paper. Finally, possible ways to reduce the big data by utilizing intelligent segmentation techniques are described with real example.

Big data in the distribution side could be reduced by adopting the centralized protection scheme. Transmitting the disturbance data to the central computer, e.g., using the IEC 61850 standard, would increase the system reliability. It would also reduce the big data, as data accumulation and analysis could be done concurrently, with the help of the central computing resource. This is currently not possible with the nodal and local computations using the relays, which have limited computational power.

Furthermore, the big data from power systems disturbances could be effectively reduced by performing the analysis remotely in a distributed manner, using the application service provider (ASP) technologies. This scheme would effectively reduce the large volume of disturbance data. Besides, it will effectively create valuable information from data, which most of the power utilities are lacking today.

## REFERENCES

- [1] P. M. Anderson, *Power System Protection*, McGraw-Hill, NY, 1999.
- [2] A. Ukil, R. Zivanovic, “Application of Abrupt Change Detection in Power Systems Disturbance Analysis and Relay Performance Monitoring,” *IEEE Transactions Power Delivery*, vol. 22, no. 1, pp. 59–66, 2007.
- [3] A. Ukil, *Intelligent Systems and Signal Processing in Power Engineering*, Springer, Heidelberg, 2007.
- [4] Int. Elet. Comm. (IEC), “Communication networks and systems in substations Part 9-2: Specific communication service mapping (SCSM) Sampled values over 390 ISO/IEC 8802-3,” IEC 61850-9-2:2004(E), 2004.
- [5] G. Manassero, E. L. Pellini, E. Senger, R. Nakagomi, “IEC61850 based systems - Functional testing and interoperability issues,” *IEEE Transactions Industrial Informatics*, vol. 9, issue 3, pp. 1436–1444, 2013.
- [6] Y. Chen, D. Liu, B. Xu, “Wide-Area Traveling Wave Fault Location System Based on IEC61850,” *IEEE Transactions Smart Grid*, vol. 4, issue 2, pp. 1207–1215, 2013.
- [7] P. Parikh, T. Sidhu, A. Shami, “A Comprehensive Investigation of Wireless LAN for IEC 61850 Based Smart Distribution Substation Applications,” *IEEE Transactions Industrial Informatics*, vol. 9, issue 3, pp. 1466–1476, 2013.
- [8] EMTP-RV Software, 2013 [Online]. <http://www.emtp.com>
- [9] IEEE, “IEEE Standard Common Format for Transient Data Exchange,” Std. 37.111-1991, ver. 1.8, 1991.
- [10] A. Ukil, R. Zivanovic, “Abrupt Change Detection in Power System Fault Analysis using Wavelet Transform,” Int. Conf. Power Systems Transients–IPST’05, Montreal, Canada, 2005.
- [11] A. Ukil, R. Zivanovic, “Abrupt Change Detection in Power System Fault Analysis using Adaptive Whitening Filter and Wavelet Transform,” *Electric Power Systems Research*, vol. 76, issues 9–10, pp. 815–823, 2006.
- [12] M. Kezunovic, L. Chen-Ching, J. R. McDonald, L. Smith, “IEEE tutorial on automated fault analysis,” 2001.
- [13] M. Kezunovic, “Smart Fault Location for Smart Grids,” *IEEE Transactions Smart Grid*, vol. 2, issue 1, pp. 11–22, 2011.
- [14] D. E. Allen, A. Apostolov, D. G. Kreiss, “Automated analysis of power system events,” *IEEE Power and Energy Magazine*, vol. 3, no. 5, pp. 48–55, 2005.
- [15] T. Popovic, M. Kezunovic, “Measures of Value: Data Analytics for Automated Fault Analysis,” *IEEE Power and Energy Magazine*, vol. 10, no. 5, pp. 58–69, 2012.
- [16] E. M. Davidson, S. D. J. McArthur, J. R. McDonald, T. Cumming, I. Watt, “Applying multi-agent system technology in practice: automated management and analysis of SCADA and digital fault recorder data,” *IEEE Transactions Power Systems*, vol. 21, no. 2, pp. 559–567, 2006.
- [17] D. DeSilva, Y. Xinghuo, D. Alahakoon, G. Holmes, “A Data Mining Framework for Electricity Consumption Analysis From Meter Data,” *IEEE Trans. Industrial Informatics*, vol. 7, no. 3, pp. 399–407, 2011.
- [18] A. Ukil, R. Zivanovic, “Adjusted Haar Wavelet for Application in the Power Systems Disturbance Analysis,” *Digital Signal Processing*, vol. 18, issue 2, pp. 103–115, 2008.
- [19] A. Ukil, R. Zivanovic, “The Detection of Abrupt Changes using Recursive Identification for Power System Fault Analysis,” *Electric Power Systems Research*, vol. 77, issues 3–4, pp. 259–265, 2007.
- [20] A. Ukil, R. Zivanovic, “Abrupt Changes in Power System Fault Analysis: A Comparative Study,” In proc. S. African Power Engg. Conf.-SAUPEC, Johannesburg, South Africa, 2005.
- [21] M. Basseville, I. V. Nikoforov, *Detection of Abrupt Changes Theory and Applications*, Prentice-Hall, NJ, 1993.
- [22] J.A. Jordaan, R. Zivanovic, “Non-parametric Time-varying phasor estimation in power systems by using a non-quadratic criterion,” *Transactions SAIEE*, vol. 95, pp. 35–41, 2004.
- [23] R. Zivanovic, P. Schegner, O. Seifert, G. Pilz, “Identification of the resonant-grounded system parameters by evaluating fault measurement records,” *IEEE Tr. Power Deliv.*, vol. 19, no. 3, pp. 1085–1090, 2004.
- [24] A. Ukil, B. Deck, V. H. Shah, “Smart Distribution Protection Using Current-Only Directional Overcurrent Relay,” IEEE PES Conf. Innovative Smart Grid Tech. Europe-ISGT10, Gothenburg, Sweden, 2010.
- [25] A. Ukil, B. Deck, V. H. Shah, “Current-Only Directional Overcurrent Relay,” *IEEE Sensors Journal*, vol. 11, no. 6, pp. 1403–1404, 2011.
- [26] A. Ukil, B. Deck, V. H. Shah, “Current-Only Directional Overcurrent Protection for Distribution Automation: Challenges and Solutions,” *IEEE Transactions on Smart Grid*, vol. 3, issue 4, pp. 1687–1694, 2012.
- [27] A. Ukil, R. Zivanovic, “Practical Approaches in Application Service Provider: A Comparative Study,” 2nd World Congress on Lateral Computing-WLCO5, Bangalore, India, Dec. 2005.
- [28] M. Musaruddin, M. Zaporoshenko, R. Zivanovic, “Remote Protective Relay Testing,” Australasian Univ. Power Engg. Conf.-AUPEC’08, 2008.
- [29] J. Ning, J. Wang, W. Gao, C. Liu, “A Wavelet-Based Data Compression Technique for Smart Grid,” *IEEE Transactions Smart Grid*, vol. 2, issue 1, pp. 212–218, 2011.