21, rue d'Artois, F-75008 PARIS
http : //www.cigre.org

CIGRE US National Committee
2018 Grid of the Future Symposium

Analytics in Practice: Someone Has to Know What is Going On

S. RHOADS[1], J. WHITE[1], T. MCGRAIL[2], G. RAJAPPAN[2]
[1]National Grid, [2]Doble Engineering
USA

SUMMARY

It has become more common to apply tools which automate the analysis of data, combining data from multiple sources, and then reducing the data to an indication of condition or status of the asset. It is important that there is a clear audit trail between the raw data and any 'standard' diagnostics and subsequent decisions; if this trail is not clear there can be difficulty in justifying actions and interventions. The phrase 'Big Data' is commonly applied to almost any activity which can generate gigabytes of data, whether operational or condition-based applications: SCADA systems, for example, but also condition data in the form of partial discharge signals.

This paper reviews some examples of the application of analytics and the implications of those analytics: including relatively low data volume applications such as Asset Health Index applications, condition monitoring analysis and alert response, through to trending large volumes of partial discharge condition data. In each case, the analytic is well defined, but the implication of the result of the analytic, and the ability to extract value from the analysis, relies not only on the analytic itself but also on the context. In addition, in order to verify an analytic is performing correctly, human verification of the performance is occasionally required.

In addition, it is noted that analytics may be used to identify anomaly, but a response plan must be risk assessed to ensure that the response itself does not put people at risk.

KEYWORDS

Analytic, big data, justification

tmcgrail@doble.com

ANALYTIC EXAMPLE: ASSET HEALTH INDEX

In an example of asset health indices, analytics are becoming more common and support asset intervention decisions. However, in practical systems we have seen, the path between raw data, analyses, encoded data, failure mode diagnosis, and finally asset health assessment is often poorly laid out and difficult to justify. For a health index to be useful it must have an associated action and a timescale – the action may vary depending on the asset type and condition, but the timescale should be consistent across all assets.

In practice, an Asset Health Index does not tell us anything new about the asset: the data is understood, the analysis is understood, the conclusions should be understood. Consequently, the value of an index is in addressing the population: which assets ***most*** need attention, and when?

A useful health index must relate to the raw data, through failure modes, to a justifiable conclusion. Further, the index is should be both monotonic and calibrated:

- Monotonicity: a worse score always indicates a more urgent asset so comparisons can be made between assets
- Calibration: all identical scores have identical urgency to allow for consistent action planning

Figure 1 shows a health assessment for three transformers based on a number of individual factors. For each factor, 1 is good, 5 is bad. The values are *calibrated* so that all 3's, for example, have the same urgency. The assessments are built on raw data which identify the presence and severity of failure modes – again, calibrated in time. By simple addition of the scores in Figure 1, 'Trf 1' has the worst score and should be most urgent – the use of simple addition is an approach equivalent to uniform linear weighting of the assessments.

Factor	Trf 1	Trf 2	Trf 3
DGA Main Tank Score	2	1	1
Dielectric Score	1	1	1
Thermal Score	2	1	1
Mechanical Score	3	4	1
Oil Score	1	1	1
DGA LTC Tank Score	3	1	5
Operational Score	2	3	3
Design/manufacturer Score	1	4	1
Subject Matter Expert Score	3	1	2
Sum	**18**	**17**	**16**
Normalized Sum (%)	40.0	37.8	35.6

Figure 1: Asset health Indices for Three Power Transformers

It is clear in Figure 1 is that the most urgent transformer is 'Trf 3', which has a single score of 5 for 'DGA LTC Tank[1]' Score, but the transformer with the highest, or worst, Sum (overall) score is 'Trf 1'. We have lost the urgency required. It is interesting to apply asset health index approaches across a population. In Figure 2, a three factor system is used to rank transformers and the overall asset health of the population is found through an average of individual scores.

[1] DGA = Dissolved Gas Analysis, LTC = Load Tap Changer

I.D.	YoM	Age	DGA Score	indings Sco	Age Score	Weighted Sum	Replace?	New Age	New Age Score*	New weighted
S1	1991	25	3	1	2	2	*	1	1	1
S2	1969	47	1	1	4	1		47	4	1
S3	1957	59	2	1	4	1.5		59	4	1.5
S4	1966	50	1	1	4	1		50	4	1
S5	1954	62	5	1	5	3	*	1	1	1
S6	2007	9	4	1	1	2.5	*	1	1	1
S7	1998	18	1	1	2	1	*	1	1	1
S8	1994	22	1	2	2	1.5		22	2	1.5
S9	1964	52	1	5	4	3		52	4	3
S10	1965	51	1	1	4	1		51	4	1
S11	1974	42	1	2	3	1.5		42	3	1.5
S12	1959	61	2	1	5	1.5		61	5	1.5
S13	1955	61	1	1	5	1	*	1	1	1
S14	1982	65	1	1	5	1		65	5	1
S15	2014	2	3	1	1	2		2	1	2
S16	2002	14	5	1	1	3	*	1	1	1
S17	1955	61	1	1	5	1		61	5	1
S18	1988	28	2	1	2	1.5		28	2	1.5
S19	1983	33	1	1	3	1		33	3	1
S20	1974	42	1	2	3	1.5	*	1	1	1
Average	Average	40.2	1.9	1.35	3.25	1.63	*	29	2.65	1.28
					Check	1.625				
Weights	Age	0								
	Windings	50								
	DGA	50								

Figure 2: Random Replacement Improves the Overall Population

Replacement of transformers 'at random' is performed and the overall population score improves markedly from 1.63 to 1.38. This could be used as a justification for the validity of the approach! Any system which identifies transformer health needs to be demonstrated to work better than the 'placebo' of random replacment.

ANALYTIC EXAMPLE: SINGLE VALUE CONDITION MONITORING RESPONSE

In general, decisions made need to be justified by evidence, including data and experience. For example, in Figure 3 a recording of composite DGA data from an online monitor is shown. The data is an indicator of a possible rise in dissolved gases, each of which is a symptom of a possible incipient fault; a fault which could lead to catastrophic failure. The data is a 'single value' representation of the dissolved gases and indicates only that at least one of them seems to have risen. Small rises may be related to transformer loading and tempoerature, but in this case the data shows a step change from about 20 ppm (parts per million) to 120 ppm in about 3 hours. What is an appropriate response?

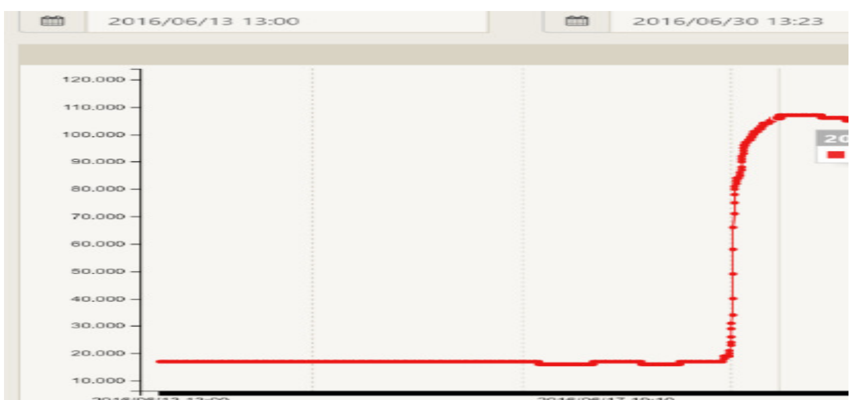

Figure 3: Step change in composite DGA response in an online monitor

A step change of this magnitude should generate some form of alerts: we would expect the analytics applied to the data to have both level and trend capability, based on standards such as IEEE C57.104 for the "Interpretation of Dissolved Gases in Large Power Transformers" and based on what is 'normal' at this location [1]. The key question is: What is the appropriate response?

In some organizations, the planned response to such a change in monitored DGA levels is to take a manual sample for ‘checking’ at a laboratory. There is a risk here to be managed: where does the responsibility lie if the transfomer fails while the sample is being taken? If the transformer fails catastophically and injures the crew taking the sample? This risk needs to be analyzed before sending a crew out to take a sample: after all, the monitor was put in place to indicate deterioration, and it has done so. There may be a false positive in the data, but that needs to be analyzed before putting someone at risk: and the risk analysis extends to all possible consequential issues: including safety.

ANALYTIC EXAMPLE: PARTIAL DISCHARGE DATA TRENDING

Partial discharge (PD) is an electrical breakdown of a portion of an insulation system which leads to overall insulation breakdown and failure. PD is an electrical process which leads to several distinct physical phenomena which can be detected: oil breakdown leading to DGA signatures, radio frequency interference (rfi) which can be picked up by an antenna in oil or air, or fast transient currents, detected in conductors using high frequency CT’s (HFCT’s), or in ground planes using transient earth voltage probes (TEV’s). Radio and electric signals propagate and may be detected at a long distance from the source. PD in one asset may be detected in the ground plane, or in another asset. Switching operations and tap changer movements generate PD signals which can be detected as ‘transient’ signals.

Consequently, detection of PD is very useful, but requires caution in detection, analysis and understanding. PD measurements can also lead to vast quantities of data (pulse magnitudes, phase angles, at over 1 GHz sampling rate), as per Figure 4,

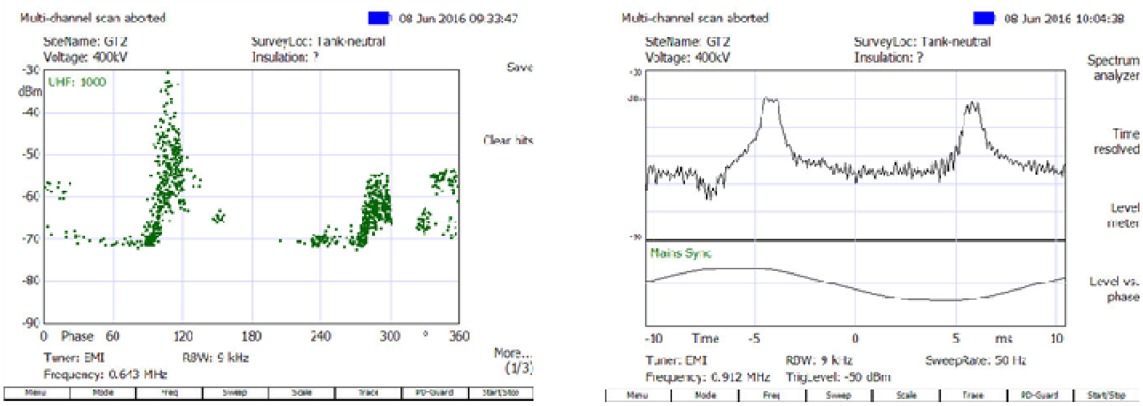

Figure 4 Partial Discharge Analysis tools: Phase resolve PD (left) and Time/Level Analyses

A PD spectrum records data across a broad frequency range, and allows for comparison of a signal from a PD free or background environment with other areas. Figure 5 shows a suspect signal (red) compared to a background (green), noting that there are siugnificant peaks present inthe environment which are not indicative of PD sources: radio, cell phone and communications signals, for example.

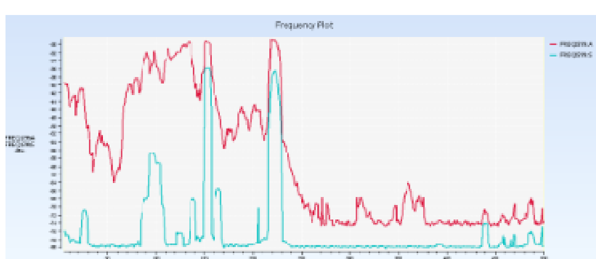

Figure 5 Partial Discharge Frequency Spectrum: Background and Suspect Signal

Interpretation of PD signals is a highly specialized skill, so there is value in application of statistical approaches which indicate PD severity and which allow for more detailed analysis when necessary, rather than continuously [2,3]. The 'root' data should be available, but the analysis should allow for trending and indicate both the presence and severity of a PD source and the rise in that severity over time. Such an approach is the "Peak-to-Average-Power Ratio" or "PAPR" which condenses the data in a PD spectrum to a single number: based on a short 100mS recording at each frequency, PAPR indicates the spread between the peak value at that frequency and the average value at that frequency, collated across the entire spectrum. Background noise and naturally has a low PAPR, while communications signals have a higher energy but also a low PAPR. PD signals have a higher PAPR due to the 'spikey' nature of the PD itself. We therefore seek signals which are high PAPR and which are sustained or growing as an indication of the presence of PD.

Figure 6 shows, in green, the leakage current magnitude from a power transformer bushing; the current drops to zero for two days when the transformer is taken out of service.

Figure 6 Bushing Leakage Current and PAPR Signals from Bushing and in Transformer Tank

The bushing PD, in red, shows some variation, but returns to its former PAPR level after the outage; the main tank PD, in blue, shows a significant rise in PAPR after the outage. This is an indication of the presence of a sustained source of PD: something worth investigating. The PAPR analytic is a means to condense a large amount of data, and indicates, here, a need for more detailed investigation

CONCLUSIONS

In the three cases presented here, analytic tools are used to indicate a 'condition' or 'health' status:

- for ranking via an asset health index
- for identification of general deterioration through DGA analysis
- for identification of insulation breakdown through PD trending

The analytics by themselves are useful, but to extract value there needs to be understanding of the context and the consequence of use of the analytic. Someone has to know what is going on…

BIBLIOGRAPHY

[1] IEEE C57.104 "Interpretation of Dissolved Gas Analysis of Large Power Transformers"
[2] "The 3 C's of Condition Monitoring: Control, Context, Conclusion", T. McGrail, K. Elkinson, R. Heywood, SEA Marconi "MyTransfo" Conference, Turin, Italy, 2016
[3] "Avoiding Metalclad Switchgear Failure Through Use of Partial Discharge Detection", T. McGrail, J. Garnett, M. Lawrence, National Electrical Testing Association (NETA) PowerTest Conference, Nashville, USA, 2015