University of Huddersfield Repository

Rescic, M., Seviour, Rebecca and Blokland,, W.

Accelerators and their ghosts

Original Citation

Rescic, M., Seviour, Rebecca and Blokland,, W. (2017) Accelerators and their ghosts. In: Proceedings of IPAC 17. JACoW, pp. 1975-1978. ISBN 9783954501823

This version is available at http://eprints.hud.ac.uk/id/eprint/32153/

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

http://eprints.hud.ac.uk/

ACCELERATORS AND THEIR GHOSTS

M. Reščič*, R. Seviour, University of Huddersfield, Huddersfield, UK
W. Blokland, Oak Ridge National Laboratory Spallation Neutron Source, Oak Ridge, TN, USA

Abstract

The issue of particle accelerator reliability is a problem that currently is not fully defined, understood nor addressed. Conventional approaches to reliability (e.g., RBDs) struggle due to a lack of data about specific component/system reliability and failure.

There is a large body of beam current data retrievable from operating accelerators that contains detailed information about the accelerator behaviour, both before and after a machine trip has occurred.

Analysing this data could provide insight and help develop a new approach to address accelerator reliability. In this paper, we propose a data-driven approach to detecting emergent behaviour in particle accelerators. Instead of attempting to identify every possible failure of a machine we propose an alternative approach based around a change in perspective, to knowing the normal default operational behaviour of a machine. Taking action when a "ghost in the machine" emerges that causes accelerator wide aberrant changes to normal machine behaviour.

INTRODUCTION

The reliability of particle accelerators has been identified as a key factor limiting the development of certain industrial applications, such as Accelerator-Driven Systems for nuclear waste-transmutation [1].

Current approaches to accelerator reliability focus around conventional reliability techniques developed, such as Reliability Block Diagrams (RBD), Fault Tree Analysis (FTA), etc [2–5]. Applying these approaches rigorously in accelerator systems is hampered due to the limited data on component failures [6]. In addition to the identification industrial applications require a real time response to failure, reduce both the number of trips per annum and decreasing the mean-time-to-recovery of the system.

In this paper we examine an approach to accelerator reliability modeling looking for emergent behaviour in the complex datasets, such as beam current/charge, created by the diagnostics systems during the operation of the accelerator. An example of emergent behaviour is the ghost in the machine. In Koestler's book [7] the human brain is built around earlier, primitive, structures, where the large number of subsystems all act together forming human conscientious. At times one, or several, of these smaller subsystems (the ghosts in the machine) can emerge dominate in the brains function, leading to extreme emotions, such as hate and anger which leads to observable changes in the behaviour of the larger system.

* miha.rescic@hud.ac.uk

For our research, the off-normal behaviour we are interested in is the machine trip of a particle accelerator - the shutdown of the machine due to a failure (or to prevent a failure).

This form of emergent behaviour is also found in electrical power networks, where the complete grid is formed from a large number of subsystems. Any change in behaviour of a component within the network (e.g., a generator) is added to and reflected in the frequency changes across the whole network. This variation in Electric Network Frequency (ENF) is used to detect off-normal behaviour in audio forensics. The ENF criterion is a procedure from audio forensics [8, 9] where an audio recording authenticity is validated by extracting low frequency mains hum and matching it to a reference database.

In this paper we look to apply this same approach from audio forensics, to particle accelerators. Using beam current measurements as the observable information on normal and off-normal accelerator behaviour, allowing us to search for predictive characteristics of the signal just prior to a machine trip.

Due to the high complexity of particle accelerators and availability of sample data we have focused our research on building a reference database of normal accelerator behaviour data where all the interconnected systems are operating in harmony.

With the data available to us we have concluded the alternative approach of building a reference database of off-normal behaviour is not feasible for our specific method. This is due to a large number of potential failures (and combinations of them) and sparse data associated with specific failures. Our selected approach enables us to identify unique off-normal behaviour which is detected by matching pulses to the reference database.

Although this proposed approach will not allow us to stop an accelerator from tripping it does allow operations to take steps to mitigate the trip. For example in some approaches proposed for commercial ADS the system consists of multiple accelerators all hitting the same target. Knowledge of when one accelerator is about to fail would allow operators to ramp-up the other accelerators to compensate as the trip occurs, hopefully increasing the up-time of the entire system.

METHODOLOGY

We are using a dataset from the Spallation Neutron Source (SNS) at Oak Ridge National Laboratory, which contains recorded current waveform in the linac section of the accelerator. The data contains three types of waveforms [10]: before the machine trip, during the machine trip and during normal operation after the machine has recovered from the trip. This allows us to extract unique patterns for each type of pulse and store them in separate datasets, either training sets or sample sets (see Fig. 1).

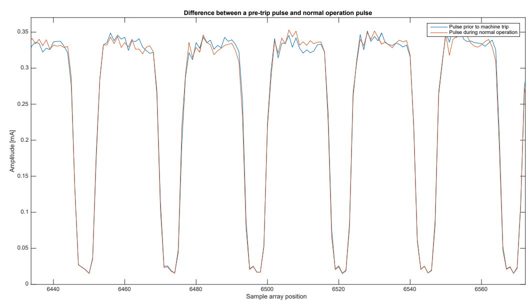

Figure 1: Zoomed section of a pre-trip pulse and normal operation pulse superimposed.

Each pulse current waveform is sampled at 100 MS/s for about 1 ms so the total pulse sample array length is 100000 sample points.

For each pulse array, we apply two procedures to generate the unique pulse signature. First procedure extracts the individual bunches from the pulse array. First detection threshold T_b is set at the mean value of all pulse values. Then, moving along the pulse array we extract bunches as the values pass over and under the T_b threshold

Second procedure is applied to remove the leading / trailing edges from the detected bunches by using a sliding average window method. The leading and trailing segments of the bunch array are removed to avoid relatively larger values that could potentially bias the matching process.

Using a sliding average with a window of $l_w = 10$ samples we detect and trip the leading and trailing edges of detected bunches.

Final step in preparation of the data is to concatenate extracted bunches into a single array that represents the pulses unique signature.

A **training** set is constructed by concatenating multiple pulse signatures into a longer array representing the reference database. A **sample** set is constructed by storing the pulse signatures independently for easier matching to the training set.

The training set was composed of 100 good pulses and 2 sample sets composed of 50 bad (pre-trip) pulses and 50 good (normal operation) pulses respectively recorded at SNS on 2015-05-05. After processing the data it yielded a sample set of approximately 5×10^6 training datapoints and 100 sample arrays of 5×10^4 datapoints each.

In order to determine if a pulse sample from a sample set matches or is similar to samples in the training set we define the measure of likeliness to be the Correlation Coefficient (CC). CC between two arrays x (a pulse from the sample set) and y (a section of the training set pulse) of length L is defined as:

$$CC(x, y) = \frac{\sum_{i=i}^{L}(x[i] - \bar{x})(y[i] - \bar{y})}{(L - 1)\sigma_x \sigma_y} \quad (1)$$

where $\bar{x}, \bar{y}$ represent array mean values and σ_x, σ_y standard deviations respectively. The coefficient has values in the range $[-1, 1]$.

Matching of a sample is performed by sliding the shorter sample array along the longer training set array. To determine if a given sample array s matches the training set S at a given position i an arbitrary matching threshold T_M must be set. A **match** is defined when $CC(s, S_i) \geq T_M$.

Given the matching process the number of matches a sample can have to a training set n_m is contained in the interval $[0, \lceil S/s \rceil]$. To determine the optimal detection threshold we performed a sweep of potential threshold settings and measured the matching results.

To classify the results of matching an arbitrary sample to a **good** training set we need to define the following possible outcomes:

- **True Positive** (TP) a good sample was identified as good as it matched the good training set *at least once*
- **True Negative** (TN) a good sample was not identified as good and has not matched the good training set
- **False Positive** (FP) a bad sample has been identified as good as it matched the good training set *at least once*
- **False Negative** (FN) a bad sample was not identified as good and did not match the good training set

RESULTS

For every sample set (good and bad) we measure the relative match results by dividing the number of samples with more than 0 matches with the total number of samples in the sample set, e.g., if 1 sample out of 10 would match the database then the match success race would be listed as 0.1. Table 1 lists the matching results for the experimental setup.

Table 1: Match Success Rates for Different Thresholds

T_M	TP	TN	FP	FN
0.200	0.7600	0.2400	0.6600	0.3400
0.206	0.7400	0.2600	0.4800	0.5200
0.212	0.7200	0.2800	0.4400	0.5600
0.218	0.6600	0.3400	0.3600	0.6400
0.224	0.6400	0.3600	0.2800	0.7200
0.230	0.6000	0.4000	0.2000	0.8000
0.236	0.5800	0.4200	0.1200	0.8800
0.242	0.5400	0.4600	0.0600	0.9400
0.248	0.4600	0.5400	0.0200	0.9800
0.254	0.3400	0.6600	0.0000	1.0000
0.260	0.2400	0.7600	0.0000	1.0000

After analyzing the initial results we have observed that for a given sample pulse the number of positive matches n_m detected varied between 1 and more than 100. This lead us to introduce a secondary detection threshold $T_s \in [1, 100]$ which then redefined a match as: $CC \geq T_M$ and $n_m > T_s$.

The introduction of the stricter secondary threshold improved the matching results. Figures 2 and 3 show the relative TP and FP matching success rates respectively plotted over a range of T_M and T_s threshold settings.

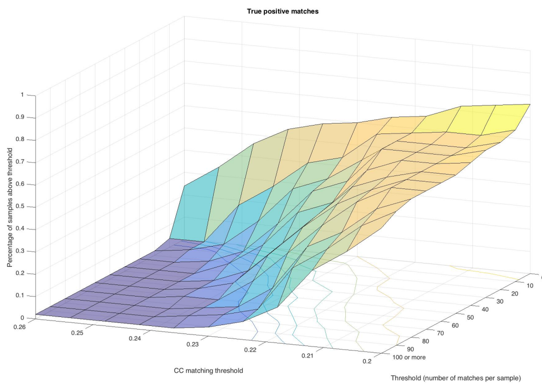

Figure 2: TP results, percentage of samples matching database for given threshold settings.

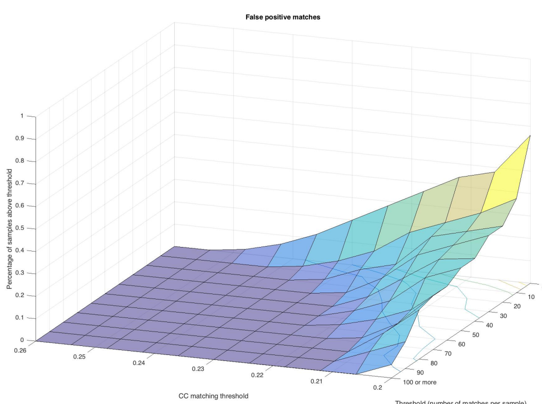

Figure 3: FP results, percentage of samples matching database for given threshold settings.

Figure 4 represents the difference between TP and FP match success plotted over different T_M and T_N settings.

We have further extracted the most favorable results into Table 2 where we list the threshold settings where the maximum TP success was reached together with the lowest FP success.

CONCLUSION

The results presented in Table 2 show that the proposed method successfully identified as much as 56% of good pulses as good while at the same time falsely identifying only 6% of bad pulses as good. Despite the big difference between good and bad pulse matching the method still misses to identify 44% of good pulses as good.

Further steps need to be taken to verify the results are not a statistical anomaly: the plan is to apply the method to a larger data set from the same or different accelerator and indeed

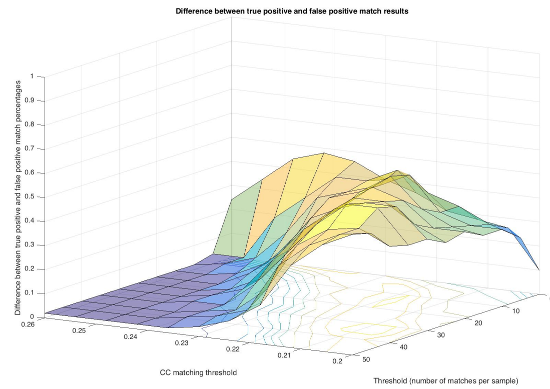

Figure 4: Difference TP−FP for a range of thresholds T_M and T_s.

Table 2: Match Success Rates with Second Threshold T_s Introduced

T_M	T_s	**TP**	**FP**
0.200000	90	0.600000	0.120000
0.200000	100	0.560000	0.060000
0.206000	60	0.620000	0.120000
0.206000	70	0.620000	0.100000
0.212000	50	0.580000	0.100000
0.212000	60	0.560000	0.060000
0.212000	70	0.540000	0.020000
0.218000	20	0.620000	0.120000
0.218000	30	0.620000	0.100000
0.218000	40	0.560000	0.060000
0.218000	50	0.520000	0.040000
0.242000	0	0.540000	0.060000
0.224000	10	0.600000	0.120000

verify that we can achieve a comparable or better success rate. Another next step is to compare pulse structure and perform matching on data acquired from different locations in the SNS accelerator and investigate the potential implications and differences.

The results we've presented are encouraging but they still leave us with significant room for improvement to achieve a success race feasible for industrial applications.

REFERENCES

[1] H. Abderrahim *et al.*, "Accelerator and Target Technology for Accelerator Driven Transmutation and Energy Production", DOE sponsored White Paper on Technology for Accelerator Driven Systems, 2010.

[2] L. Burgazzi and P. Pierini, "Reliability studies of a high-power proton accelerator for accelerator-driven system applications for nuclear waste transmutation", in *Reliability Engineering & System Safety*, volume 92, number 4, pp 449–463, 2007.

[3] A.E. Pitigoi *et al.*, "Modelling high-power accelerators reliability - reliability model of SNS linac", in *Proceedings, 39th*

Annual Meeting of Spanish Nuclear Society, Spain, September, 2013.

[4] C. Tapia *et al.*, "Exploration of reliability databases and comparison of former IFMIF's results", in *Fusion Engineering and Design*, volume 86, number 9-11, pp 2726–2729, 2011.

[5] E, Bargalló *et al.*, "Availability simulation software adaptation to the IFMIF accelerator facility RAMI analyses", in *Fusion Engineering and Design*, volume 89, number 9-11, pp 2425–2429, 2014.

[6] P. Pierini and L. Burgazzi, "ADS accelerator reliability activities in Europe", in *Proc. Utilisation and Reliability of High Power Proton Accelerators Workshop*, Daejeon, Korea, 16-19 May, 2004, pp. 137–147, OECD Publishing, Paris, `http://dx.doi.org/10.1787/9789264013810-14-en`

[7] A. Koestler, (1990-06-05) [1967]. "The Ghost in the Machine", Ed. 1990 reprint, Penguin Group, ISBN 0-14-019192-5.

[8] C. Grigoras, "Digital audio recording analysis–the electric network frequency criterion", in *International Journal of Speech Language and the Law*, vol. 12, pp 63–76, 2005.

[9] M. Huijbregtse and Z. Geradts, "Using the ENF criterion for determining the time of recording of short digital audio recordings", in *Computational Forensics: 3rd International Workshop, IWCF 2009*, The Hague, The Netherlands, August 2009, pp 116–124.

[10] W. Blokland and C. Peters, "A new differential and errant beam current monitor for the SNS accelerator", in *Proceedings of the 2nd International Beam Instrumentation Conference (IBIC'13)*, Oxford, UK, September 16–19, 2013, pp 921–924.