Two-day advance prediction of a blade tear on a steam turbine of a coal power plant

Using only automated mathematical data-analysis techniques and no prior knowledge or human experience

Dr. **P. Bangert**, algorithmica technologies GmbH, Bremen

Abstract

A blade tear on a steam turbine of a coal power plant is a major failure causing large financial losses. Knowing that such an event is going to occur in one or two days time, allows the operational staff to shut down the turbine before the damage is done. The affected blade may then be removed and normal operations resumed. To this end, we must verify two hypotheses: (1) It is possible to predict two days in advance that a blade tear is going to occur and (2) it is possible to determine which blade(s) is affected. Moreover, we wish to do so in an automated way not involving human experience or knowledge so that the prediction is fully objective, can be continuously run and is relatively inexpensive to implement. This paper describes the verification of both hypotheses on an actual example of a steam turbine in a coal power plant in Germany.

Kurzfassung

Ein Schauffelabriss an einer Dampfturbine in einem Kohlekraftwerk ist ein großes und teures Ereignis. Zwei Tage vorher zu wissen, dass ein solches Ereignis eintreffen wird, erlaubt dem Betreiber ein kontrolliertes Abfahren bevor der Schaden eintritt. Die betroffene Schaufelreihe kann nun entfernt und die Anlage wieder hochgefahren werden. Um dieses Ziel zu erreichen, müssen wir zwei Hypothesen verifizieren: (1) Es ist möglich zwei Tage im Voraus einen Schauffelabriss zu prognostizieren und (2) es ist möglich zu bestimmen welche Schauffelreihe betroffen ist. Darüberhinaus möchten wir dies automatisieren, so dass wir keine menschliche Erfahrung oder Wissen benötigen und damit die Prognose objektiv, kontinuierlich und finanziel günstig gemacht werden kann. Dieser Artikel beschreibt die Verifikation beider Hypothesen an einer realen Dampfturbine in Deutschland.

This paper was published in: Link, M. (2010): Schwingungsanalyse & Identifikation -- Proceedings of the Conference of 23./24.03.2010 in Leonberg, Germany. (VDI-Berichte No. 2093) ISBN: 978-3180920931. pp. 175--182

1. Statement of the Problem

A coal power plant essentially works by creating steam from water by heating it by a coal furnace. This steam is passed through a turbine, which turns a generator that makes the electricity. The most important piece of equipment in the power plant is the turbine. It is large, expensive and custom made. Any damage to it causes long delays before damaged components can be replaced. Moreover, any shutdown of the turbine necessarily yields a total shutdown of the entire facility. See figure 1 for a diagram.

As the turbine is so important, its operations are monitored by a variety of sensors installed in key locations. The most crucial information regarding the health of the turbine is contained in the vibration measurements. All sensor output is logged in a data historian and therefore available for study.

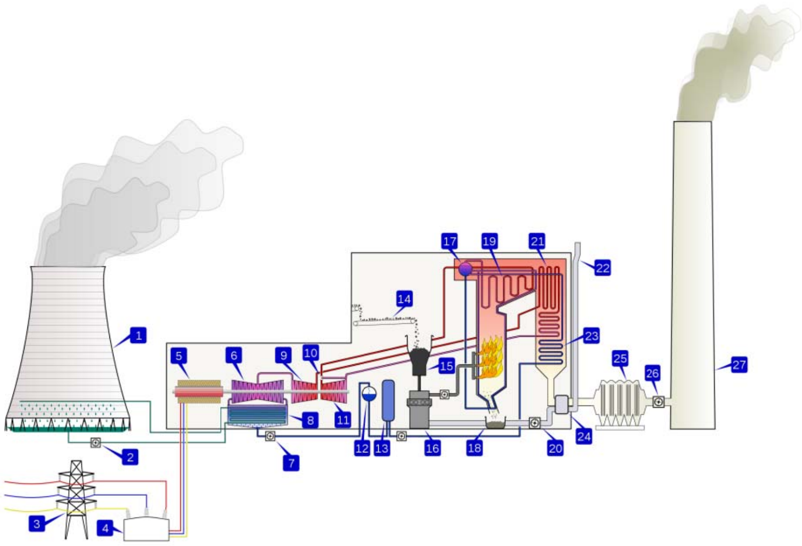

Fig. 1: The main constituents of a coal power plant.

Let us assume that we have a total of *N* measurements on and around the turbine that we wish to look at. We may represent the state of the turbine at time *t* by an *N*-dimensional vector, $\mathbf{x}^{(t)}$. Via the data historian, we may obtain a set of such vectors for past times. If we order this set with respect to time, then this set is called a time-series, $\mathbf{H} = (\mathbf{x}^{(-h)}, \mathbf{x}^{(1-h)}, \mathbf{x}^{(2-h)}, \ldots, \mathbf{x}^{(0)})$ where time $t = 0$ is the current moment and time $t = -h$ is the most distant moment in the past that we wish to look at. Thus the time-series **H** is effectively a matrix with $h+1$ columns and *N* rows.

Given that the system under question (the turbine) is governed by physical laws that do not change over the history and that h is sufficiently large, then it follows that the function $\boldsymbol{f}$ exists: $\mathbf{f}(\mathbf{H}) = \mathbf{H} \prod \mathbf{x}^{(1)}$, where the symbol $\prod$ indicates concatenation of the vector $\mathbf{x}^{(1)}$ to the right side of the matrix **H**. This function may be applied recursively so that $\mathbf{f}^{n}(H) = \mathbf{x}^{(n)}$. In this way, we may compute the n-th state of the system, i.e. the state that the system will have in n time steps from the current time.

Thus, the mathematical challenge is to design methods to find $\boldsymbol{f}$. The computational challenge is to determine $\boldsymbol{f}$ such that it gives reliable (i.e. $\mathbf{x}^{(1)}$ matches reality) and stable performance (i.e. $\mathbf{x}^{(n)}$ for large n does not diverge).

2. Theoretical limitations

Of course, whatever methods we choose, they cannot have arbitrary accuracy or stability. Every time the function $\boldsymbol{f}$ is applied, a bit of accuracy is lost. Thus, every $\mathbf{x}^{(n)}$ has an inherent model induced uncertainty $\Delta\mathbf{x}^{(n)}$ attached to it. This means that the true value of the state vector is somewhere in the range $[\mathbf{x}^{(n)} - \Delta\mathbf{x}^{(n)}, \mathbf{x}^{(n)} + \Delta\mathbf{x}^{(n)}]$. By and large, the size of this range grows with n and so there is a maximum possible n to which it is reasonable to perform a predictive computation.

The initial source of the uncertainty is the measurement uncertainty $\Delta\mathbf{x}^{(0)}$ of the current state. Please note that no measurements made in the real world are ever completely precise. There are random and structured errors associated with the measurement process, also physical sensors drift with age and environmental effects. All of these must be taken into account to determine a reasonable measurement uncertainty $\Delta\mathbf{x}^{(0)}$, which then leads, via $\boldsymbol{f}$, to the predictions' uncertainties $\Delta\mathbf{x}^{(n)}$.

A further limitation is the length of the history h. The history must contain a record of the variations that are to be expected in the future so that these variations, correlations and other structures may be included in the function $\boldsymbol{f}$. It is thus desirable that h be as large as possible and also the time unit (governing the frequency of measurements) be as small as possible. Together these two define a history that contains the maximum available knowledge about the system.

Our efforts are thus limited by three fundamental factors: (1) The number and identity of the measurements made, (2) the length, frequency and variability of recorded history and (3) the inherent accuracy of a measurement itself. Together these three factors will determine whether a reliable and stable model ***f*** can be found and, if it can, for how long into the future it can deliver useful predictions.

3. Methods

There exist many methods to find the mathematical model ***f*** that we desire and many computable representations of it. A typical approach is to create a model based on first principles (i.e. physical laws derived from mechanics, thermodynamics, etc.). These models are usually created by human engineers and implemented in a generic software program. Manually created models have to be simplified in order to be tractable by humans. Such models are also very difficult to adapt to new situations (e.g. a rebuild) and thus have a finite lifetime associated with them. Because of the necessary effort to create such models and the above limitations, these are expensive.

Mathematically, the field of machine learning has focused on creating such models automatically without human participation [1]. The advantages are that the model is produced within a very short time (usually days), that it is adaptive (i.e. it learns continuously as it experiences more data), that it can change to match new situations (the new data is learnt) and that the entire problem can be modeled (and not a simplified version as in the manual approach). Thus, this method is economical.

For example, the technology of neural networks is often applied to problems of this type. Here we must differentiate classificatory neural networks [2] from recurrent neural networks [3]. The first can tell the difference between a finite number of types of objects while the second can represent the evolution over time.

The method applied later in this paper is an advanced state-of-the-art proprietary machine learning method developed by algorithmica technologies for these forms of problems.

4. Application

The specific turbine in question has over 80 measurements on it that were considered worthwhile to monitor. Most of these were vibrations but there were also some temperatures, pressures and electrical values. A history of six months was deemed long enough and the frequency depended upon each individual measurement point – some were measured several times per second, others only once every few hours. In fact, the data historian only stores a new value in its database if the new value differs from the last stored value by more than a static parameter. In this way, the history matrix contained a realistic picture of an actual turbine instrumented as it normally is in the industry. No enhancements were made to the turbine, its instrumentation or the data itself. A data dump of six months was made without modification.

The data stopped two days before a known (historically occurring) blade tear on that turbine. During time leading up to the blade tear and until immediately before it, no sign of it could be detected by any analysis run by the plant engineers either before or after the blade tear was known. Thus, it was concluded that the tear is a spontaneous and thus unpredictable event.

Initially, the machine learning algorithm was provided with no data. Then the points measured were presented to the algorithm one by one, starting with the first measured point $\mathbf{x}^{(-h)}$. Slowly, the model learned more and more about the system and the quality of its predictions improved both absolutely (the interval $\Delta\mathbf{x}^{(n)}$) and in terms of the maximum possible future period of prediction. Once even the last measured point $\mathbf{x}^{(0)}$ was presented to the algorithm, it produced a predication valid for the following two days of real time. The result may be seen in figure 2.

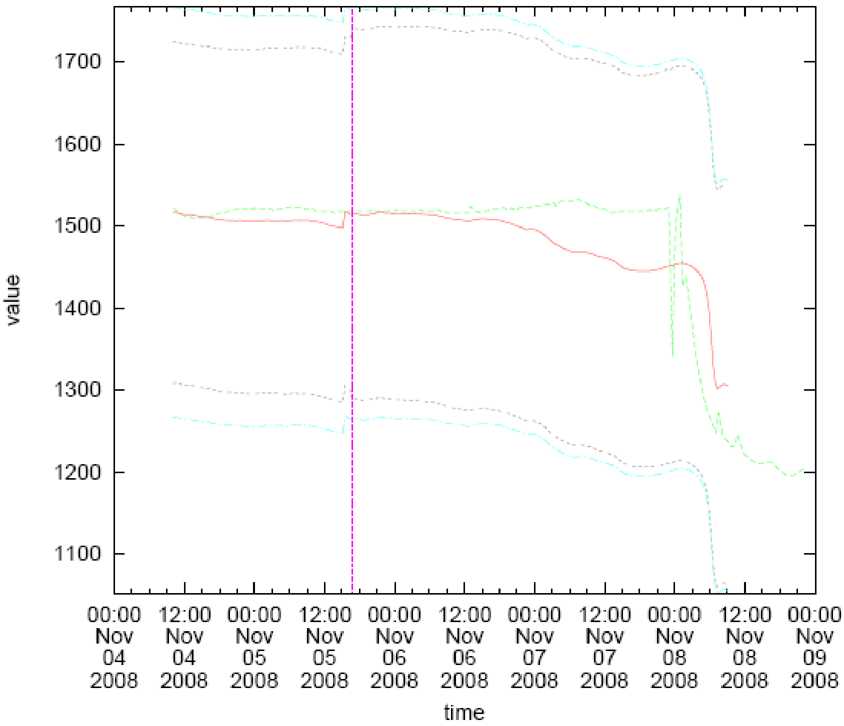

Fig. 2: Here we see the actual measurement (green) with its measurement uncertainty (the two blue lines) versus the model output (red) with its uncertainties (the two black lines) over a little history (left of the vertical pink line) and for the future two days (right of the vertical pink line). We observed a close correspondence between the measurement and the model. Particularly the event, the sharp drop, is correctly predicted two days in advance.

Thus, we can predict accurately that something will take place in two days from now with an accuracy of a few hours. Indeed it is apparent from the data that it would have been impossible to predict this particular event more than two days ahead of time due to the qualitative change in the system (the failure mode) occurring a few days before the event. The model must be able to see some qualitative change for some period of time before it is capable of extrapolating a failure and so the model has a reaction time. Events that are caused quickly are thus predicted relatively close to the deadline but two days warning is enough to prevent the major damages in this case. In general, failure modes that are slower can be predicted longer in advance.

It must be emphasized here that the model can only predict an event, such as the drop of a measurement. It cannot label this event with the words “blade tear.” The identification of an event as a certain type of event is altogether another matter. It is possible via the same sort of methods but would require examples of blade tears and this is a practical difficulty. Thus, the model is capable of giving a specific time when the turbine will suffer a major defect; the nature of the defect must be discovered by manual search on the physical turbine.

This is interesting but to be truly helpful, we must be able to locate the damage within the large structure of the turbine, so that maintenance personnel will not spend days looking for the proverbial needle in the haystack.

Fault detection and localization is now done by performing an advanced data-mining methodology that tracks frequency distributions of signals over the history and can deduce qualitative changes. Over the 80 measurements points, we are able to isolate that four of them contain a qualitative shift in their history and that two of these four go through such a shift several days before the other two. Thus, we are able to determine which two out of 80 locations in the turbine are the root cause for the event that is to occur in two days. See figure 3 for an illustration.

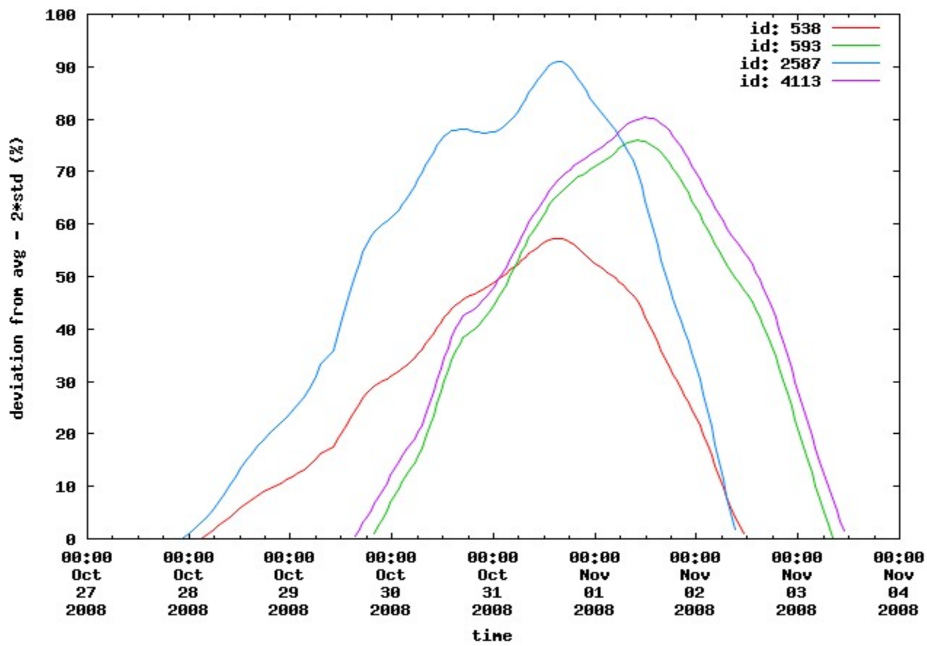

Fig. 3: We compute a deviation from normal being tracked over a window of about four days length. So we observe that two points start behaving abnormally and two days later, two other points behave abnormally. About 3.5 days after the start of the abnormal behavior, this new behavior has become normal and so the deviation from normal is seen to reduce again. Therefore, we observe a qualitative change in the performance of these four points.

The localization that is possible here is to identify the sensor that measures an abnormal signal and that will be the first to show the anomaly that will develop into the event. It is, of course, not possible to compute a physical location on the actual turbine more accurately than the data provided. However, a physical search of the turbine, after the actual blade tear, found out that the cause was indeed at the location determined by the data-mining approach.

5. Prospects

Even more than this is possible. If the damage mechanism is slower (aging, dirt accumulation, fatigue, etc.), then much longer prediction intervals are possible. For example, in the context of catalyst aging in (petro-) chemical processes, prediction windows of several months are realistic.

Also complex failures can be predicted in which the cause of the failure has several effects over time before the equipment actually shuts down. For instance, an abnormally small amount of lubricant in a machine can cause increased friction that can cause blockages that can cause moving parts to stop that can cause spindle breakages. We have many links before an effect visible from the outside occurs and yet we may trace abnormalities to their beginning and predict the end.

6. Conclusion

It is possible to reliably and accurately predict a failure on a steam turbine two days in advance. Furthermore, it is possible to locate the cause of this within the turbine so that the location covered by the sensor that measures the anomaly can be focused on by the maintenance personnel. The combination of these two results, allows preventative maintenance on a turbine to be performed in a real industrial setting saving the operator a great expense.

Not only is this possible, it has been done with automated mathematical means that do not require any human knowledge to be inserted into the model. The model thus requires no time investment by the plant’s engineers and requires no detailed knowledge of the plant itself – it requires only historical data. Thus this effort is very economical.

The approach has been successfully tested on a real turbine blade tear at a coal power plant in Germany.

[1] Bishop, C.M.: Pattern Recognition and Machine Learning. Heidelberg: Springer 2006

[2] Rosenblatt, F.: The perceptron: a probabilistic model for information storage and organization in the brain. Psychological Reviews 65, 1958, 386-408

[3] Mandic, D., Chambers, J.: Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability. Hoboken: Wiley 2001