AN APPLICATION OF PATTERN ANOMALY DETECTION METHODS TO FLEET-WIDE ASSET LEVEL DIAGNOSTICS

Chance M. Kleineke, Nilimb Misal and Michael T. Santucci
Engineering Consultants Group, Inc.
3394 W. Market Street
Akron, OH 44333
330.869.9949
kleinekec@ecg-inc.com

Abstract: Centralized monitoring techniques have become more widely used as business demands and budgetary cuts for companies require streamlined operation and maintenance of a company's assets. These assets may be located at a single site where the monitoring is taking place, or they may be located all over a state, country or the world. Local data collection with consolidated servers allows a central maintenance center to pool big data for fleet-wide monitoring purposes. Advanced pattern recognition (APR) software solutions have been on the forefront of managing big data for dealing with a multitude of assets. APR techniques can provide evidence that a machine is not operating as expected, but the condition detected could indicate many possible underlying faults. The root cause may still be unknown.

Causal network analysis has been widely used in providing differential diagnosis in the medical field when a set of symptoms are known. This method is based on Bayesian probability which can handle uncertainty in the data, both input and output, and has a good theoretical foundation. This paper discusses methods to utilize pattern anomalies as symptoms for a causal network to diagnose asset conditions and to mitigate failures for predictive maintenance programs.

Key Words: Advanced Pattern Recognition; causal networks; asset condition diagnostics; health management; condition-based maintenance; predictive maintenance; centralized monitoring; fleet-wide monitoring; diagnostics; prognostics

Introduction: With ever increasing business demands in the energy utilities industry, many organizations look for methods to maximize departmental efficiency. Rather than depend solely on crews at a single location, power generation companies now look to consolidate maintenance and performance optimization roles to a single entity monitoring a fleet of plants or units. Breaking this down further, the monitoring center must be able to concentrate their efforts on optimizing performance and understanding the condition of each individual asset across the fleet.

At this level, a monitoring center may be inundated with tens of thousands of individual asset variables. A method is required to manage big data to best accomplish the monitoring objectives. There can be terabytes of new data generated daily that these centers must use to determine the health and performance of the organization's many assets. A single gas turbine, for example, may have ~2,000 variables being stored, with about 200 specific to the performance of the turbine, like gas temperatures and exhaust pressures. While there can be another 200 variables more specifically for mechanical integrity, i.e. vibration readings or bearing and oil temps. A single coal-fired generating unit may have 40-50 different assets to monitor, each with anywhere from a few recorded variables to hundreds.

Advanced Pattern Recognition: A simple approach to performance monitoring is setting hard-limit thresholds to alarm when data values extend outside of a predetermined "optimal" range. A drawback to this, however, is that the optimal range must be one that includes all operating loads and external input to avoid false alarms. For example, a bearing temp high limit would need to be set high enough to include the highest operating temperature on a hot day when ambient temps are affecting the bearing and the unit is operating at full capacity. But at different operating conditions, such as a cooler day when the ambient isn't lending to higher bearing temps and load is low, the bearing temp could be higher than it historically was, while remaining under the high threshold. This method would not alarm on faults occurring within the defined normal operating range.

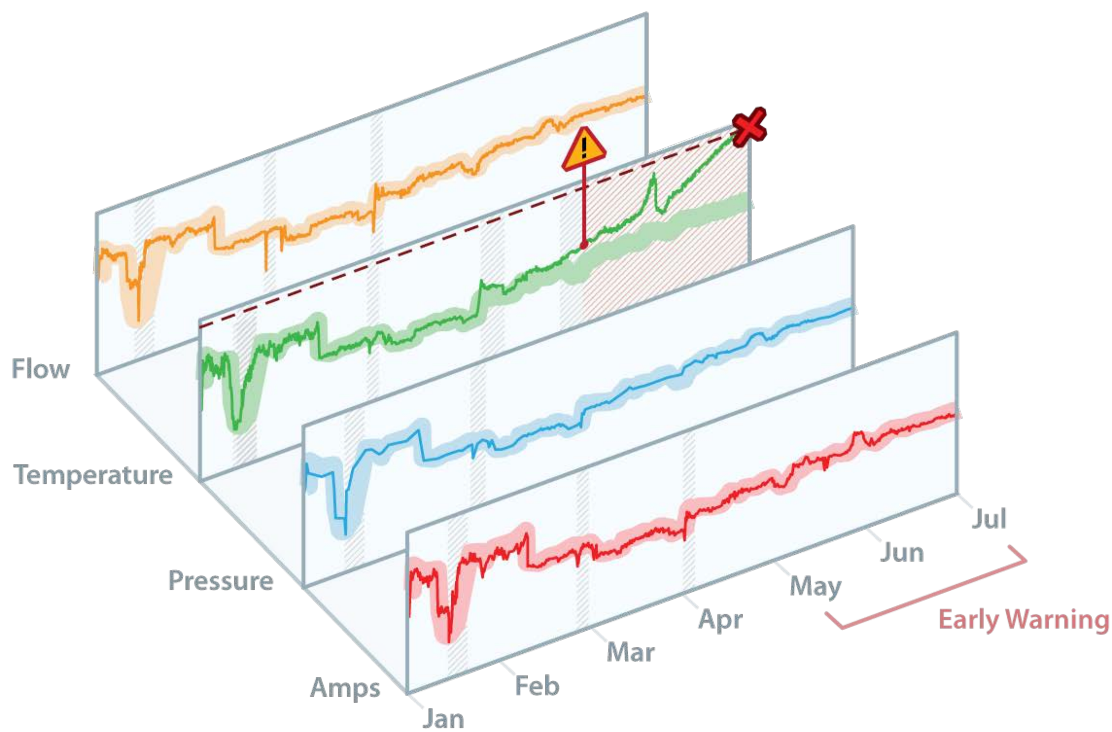


Figure 1: Early fault detection compared to hard-limit methods.

APR methods rely on the historical data of multiple variables of an asset to predict what the value of each variable should be at any given moment. The predictions are based on the activity of the other variables and relates to the performance of the variables in the past. APR methods are able to find faults much sooner than threshold monitoring techniques widely used in maintenance programs and can predict values dependent on the other variables, see Figure 1. The methods work best when there is at least some correlation among the variables being monitored, and so independent assets are modeled separately and the calculations utilize only variables relevant to that particular asset.

APR has become a common application for organizations that manage large amounts of performance and maintenance data on their assets. An engineer or data analyst building an APR model for the first time must determine boundary conditions for each asset, and subsections of those assets if necessary. Data variable selection can be completed using knowledge of the sensors fitted to each piece of equipment. Asset management and variable mapping can be completed easily by templating assets in a spreadsheet, Table 1. The models can then be imported into an APR software suite, such as Predict-It™ by ECG, Inc. [1].

Table 1: Building asset model templates.

Process Variable	1A Fan	1B Fan	2A Fan	2B Fan
Inlet Pressure	1APT1060	1BPT1060	2APT1060	2BPT1060
Outlet Pressure	1APT1065	1BPT1065	2APT1065	
Amps	1ACT2556	1BCT2556	2ACT2556	2BCT2556
Flow	1AFT1559	1BFT1559	2AFT1559	2BFT1559
Air Temperature	1ATT1205	1BTT1205	2ATT1205	2BTT1205
In Bearing Temp	1ATT1215	1BTT1215	2ATT1215	2BTT1215
Out Bearing Temp	1ATT1216	1BTT1216		2BTT1216
Damper Position	1AZT3324	1BZT3324	2AZT3324	2BZT3324
Vibration 1	1AVT0056	1BVT0056	2AVT0056	2BVT0056
Vibration 2	1AVT0057	1BVT0057	2AVT0057	2BVT0057
Vibration 3	1AVT0058		2AVT0058	2BVT0058

Once the boundary conditions are set and asset variables have been selected, the models that will be monitoring the assets can be created and trained on historical data. Relevant data should be selected that best represents optimal or normal operating conditions, Figure 2. Next, viewing the correlations among all the variables, the engineer can remove outliers using standard X-Y scatter plots, or other statistical charts that aid in narowing down the best data to be used for training, Figures 3 and 4.

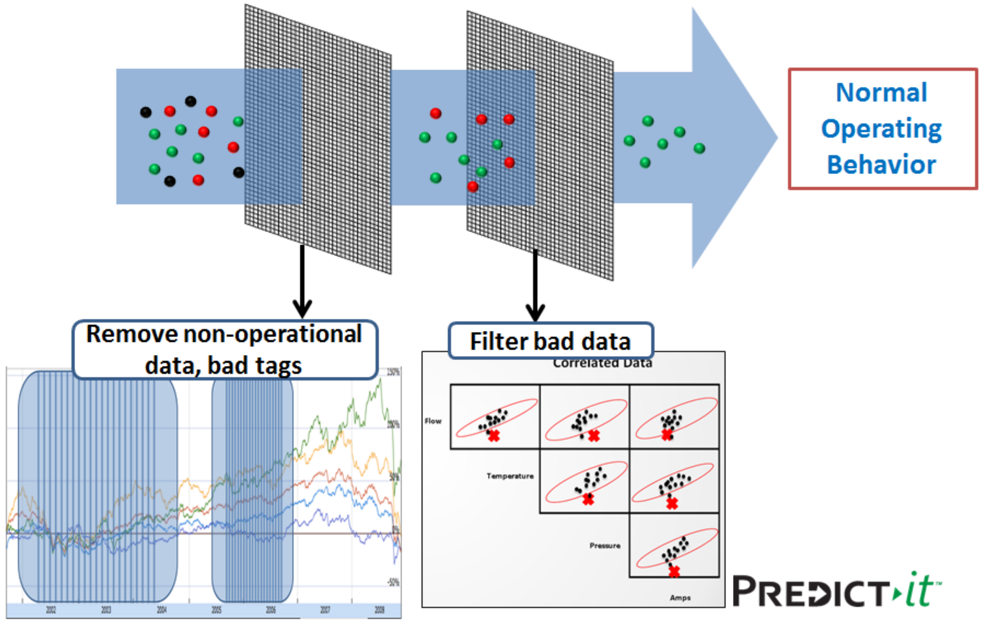


Figure 2: Selecting training time frames that best represent optimal conditions.

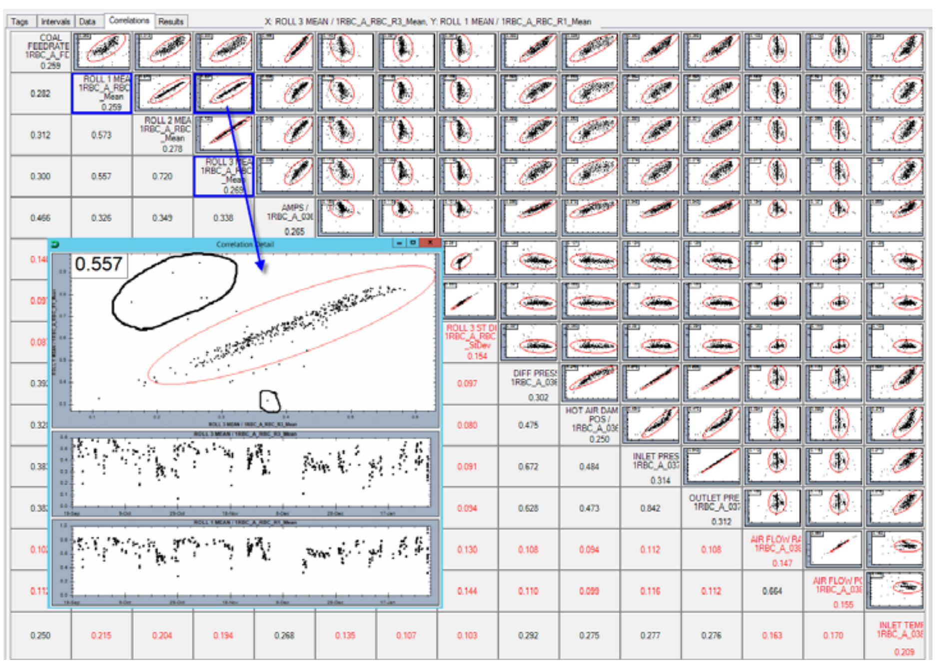


Figure 3: Correlation charts of all asset variables for the model. The model builder can manually remove outliers from individual plots (black circles).

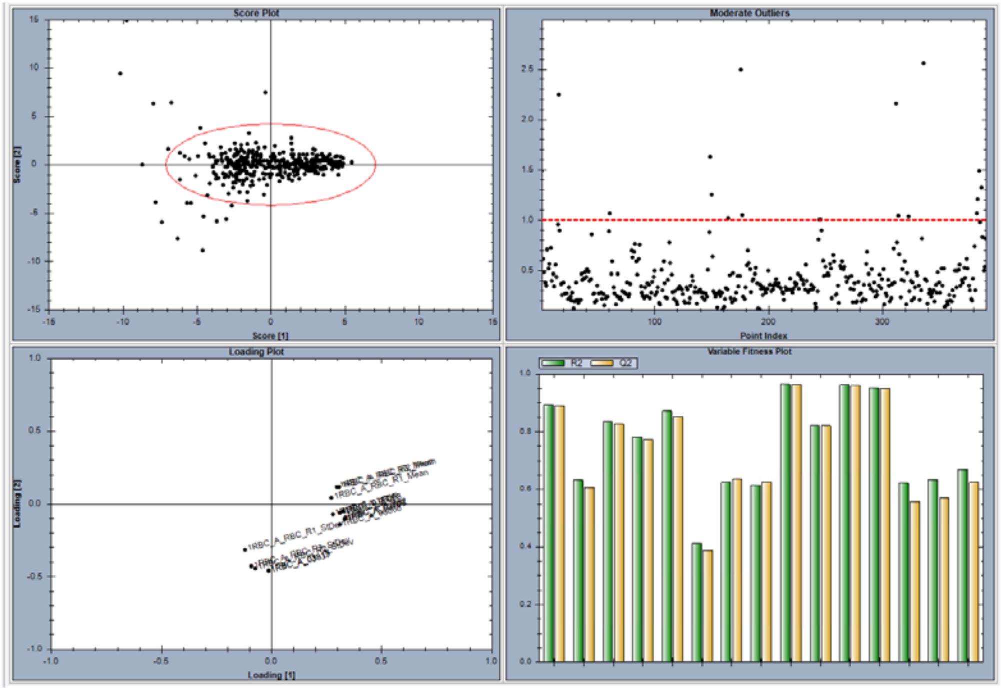


Figure 4: Alternate statistical plots for narrowing down training data.

Figure 5: Asset variable in alarm; actual value is lower than expected.

Once the data is cropped to an acceptable extent, the model can be programmatically trained. It can then be run against existing data as well as run against incoming current data. Figure 5 shows an example of the results of running a model against 2 weeks of past data. One of the variables is lower than expected and in alarm. Any APR solution typically produces a result similar to this, revealing an early warning when a particular variable, or variables, are not performing as expected. But the engineer or data analyst may be left wondering what this means. For the purpose of this paper, these variable deviations will be used as inputs to a Diagnostic Advisor that can be used as decision support for the end user. The supporting diagnostics are based on the theory of Bayesian networks.

Bayesian Networks: Bayesian networks are directed acyclic graphs (DAG) that map variables to each other with their associated dependencies. A DAG is a graph with directed arrows but which do not contain directed cycles. These dependencies are represented by conditional probability tables (CPT). The CPTs are a reformulation of the joint probability distribution between two variables. Joint probability is the probability of two or more events or variables occurring together. Conditional probability is the probability of a variable/event given that we know another state of another variable/event.

There is no distinction between independent and dependent variables and inference can be done in any direction. The method is non-parametric, nonlinear and can handle numerical and categorical variables. In statistics these models are also called directed graphical models.

The term "Bayesian networks" was coined by Judea Pearl in 1985 to emphasize three aspects [2, 3]:

1. The often subjective nature of the input information.
2. The reliance on Bayes' conditioning as the basis for updating information.
3. The distinction between causal and evidential modes of reasoning

Bayesian networks consist of nodes connected by arrows representing real causal relations. Bayesian networks have a big advantage in that they are direct representation of the world. Unlike neural networks, deep learning and other “black box” methods, Bayesian networks are transparent, intuitive to understand and are capable of providing clear explanations.

Bayesian network use the Bayes theorem to update the network. Mathematically Bayes’ theorem is stated as the following equation [4]:

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)},$$

Where A and B are events and $P(B) \neq 0$:

- $P(A|B)$ is conditional probability of event A occurring given B is true.
- $P(B|A)$ is the conditional probability of event B occurring given that A is true
- $P(A)$ and $P(B)$ are the probabilities of observing A and B independent of each other.

Humans naturally can think of problems by going from cause to effect. For example, it is known that the flu causes a cough. But a cough could be caused by the flu or the common cold. Bayes' theorem allows a reverse reasoning from effect to cause using sound probability theory. Since an engineer has to usually start with effects and then try to figure out the cause, the Bayes approach is very useful in this scenario. Starting from prior probabilities of faults, as soon as new evidence is available, Bayes theroem allows for these probabilities to be updated.

Bayesian networks can be converted to causal network when constrained in the sense that the parents of each node are its direct causes. Historically structural equations are commonly used to causally represent the different mechanisms active in a system. Bayesian representation has been shown to be equivalent to structural equations [5, 6].

Asset Management: Organization is crucial when dealing with such large amounts of data. Predict-It™ utilizes an asset hierarchy system that allows users the ability to order assets by unit or site. In addition, particularly large assets may be subdivided into smaller segments. Or one can monitor different modes of operations, such as mechanical soundness and process performance. Figure 6 represents an asset hierarchy of 5 coal mills on a single steam unit with the parent folder labeled as Mill Data. There would be other parent folders for other assets on this unit. Other units or plants could be similarly set up.

Defining an asset involves programming any possible symptoms the user believes could result from the pattern recognition calculations. The end user must then outline any possible faults for that particular asset. The faults must be defined with the corresponding directional symptoms, i.e. deviation high, deviation low, absolute high, absolute low, etc. User questions may also be set-up that allows input regarding data that may not be electronically recorded or tracked as an analog variable. Examples of this might be an asset's service time or observational items like noise level of the asset.

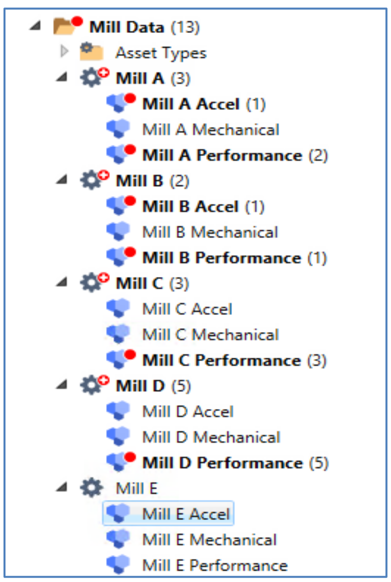


Figure 6: Asset hierarchy.

Actual instances that have been documented can be entered into the diagnostic calculations as well. When a fault occurs, the user should make note of any symptoms present to record this as a case. When these cases are entered, the probability of each fault is adjusted accordingly. Figure 7 shows what a fully configured coal mill might look like.

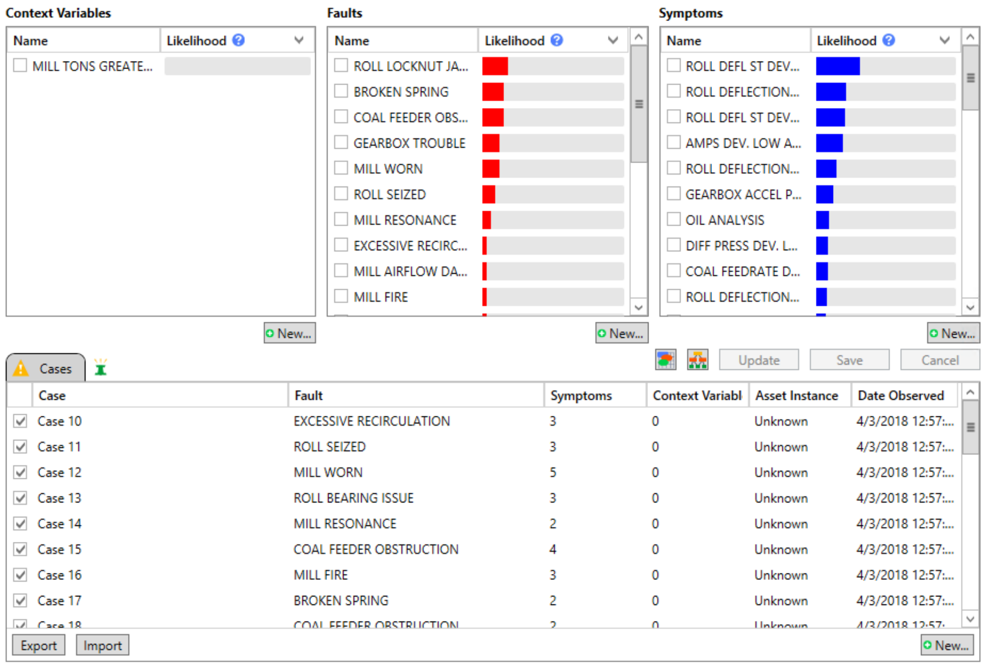


Figure 7: Defining an asset with context variables, possible faults and symptoms. The end user can submit actual cases to further solidify the diagnostic calculations.

Defining assets types allows the Diagnostic Advisor to create the Causal Asset Network (CAN). The CAN is represented as a network of Faults and Symptoms, with arrows linking associations from each column. These networks can become quite complex as the many variations of symptom combinations are set up. Figure 8 illustrates a simple CAN for an arbitrary asset. The red boxes are faults, blue boxes are symptoms and the green boxes represent context variables.

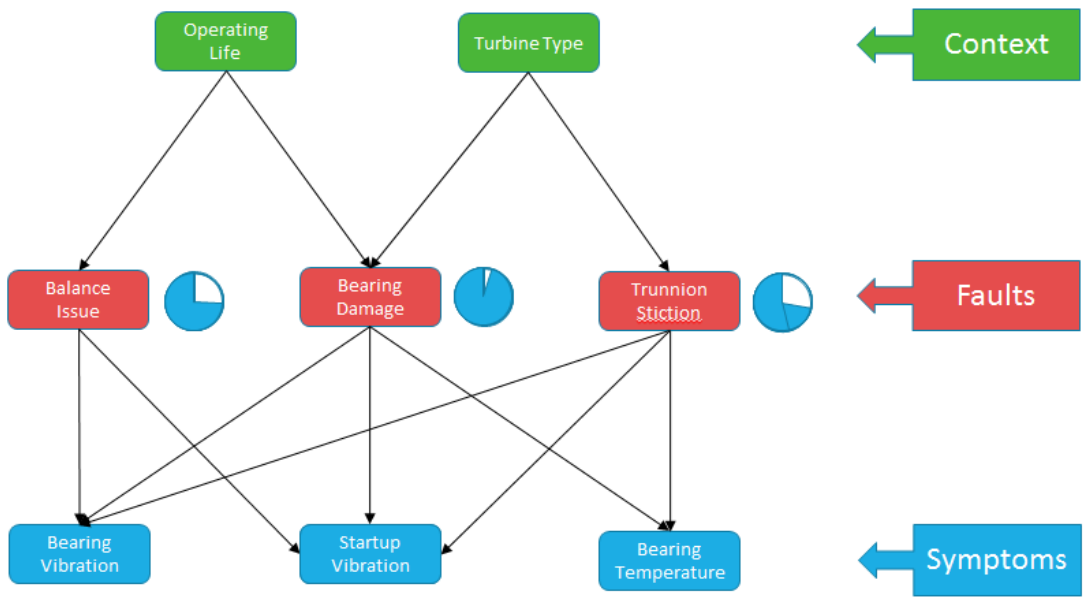
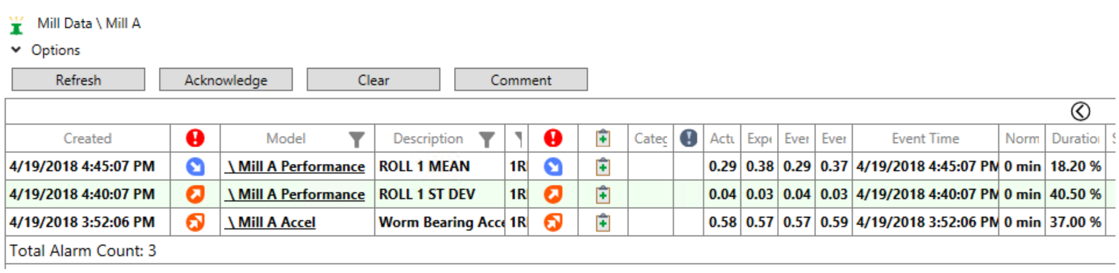


Figure 8: An example of a Causal Asset Network (CAN) created with an asset's faults and symptoms.

Diagnostic Reasoning: The Diagnostic Advisor will combine the functionality of the Bayesian networks with the results of pattern recognition fault detection. It will be used as a tool to visualize the diagnostics steps, experiment with diagnostic scenarios and answer questions related to asset conditions. Figure 9 shows an example list of pattern anomaly alarms that have been triggered for Mill A used in examples above. It shows that Roll 1 mean deflection is lower than expected and the standard deviation for that roll is higher than expected. It also shows that the accelerometer on the worm bearing is in alarm as an absolute high (hard-limit). These deviations from the expected are read into the diagnostic engine as symptoms.



Mill Data \ Mill A

Options

Refresh | Acknowledge | Clear | Comment

Created		Model	Description				Categ		Actu	Expe	Ever	Ever	Event Time	Norm	Duratio
4/19/2018 4:45:07 PM		\ Mill A Performance	ROLL 1 MEAN	1R					0.29	0.38	0.29	0.37	4/19/2018 4:45:07 PM	0 min	18.20 %
4/19/2018 4:40:07 PM		\ Mill A Performance	ROLL 1 ST DEV	1R					0.04	0.03	0.04	0.03	4/19/2018 4:40:07 PM	0 min	40.50 %
4/19/2018 3:52:06 PM		\ Mill A Accel	Worm Bearing Acce	1R					0.58	0.57	0.57	0.59	4/19/2018 3:52:06 PM	0 min	37.00 %

Total Alarm Count: 3

Figure 9: Current faults from the APR deviations.

The Diagnostic Advisor is divided into four main viewing blocks, see Figure 10. These blocks include Ranked Faults, Evidence, User Questions and Ranked observations. The Ranked Faults box will include any possible faults that have been added to the asset. If the asset has evidence from pattern recognition or user questions, the faults will be ranked by likelihood of presence. Any symptoms from pattern recognition or context variables that are currently present will be shown in the Evidence box, in order of influence to the faults.

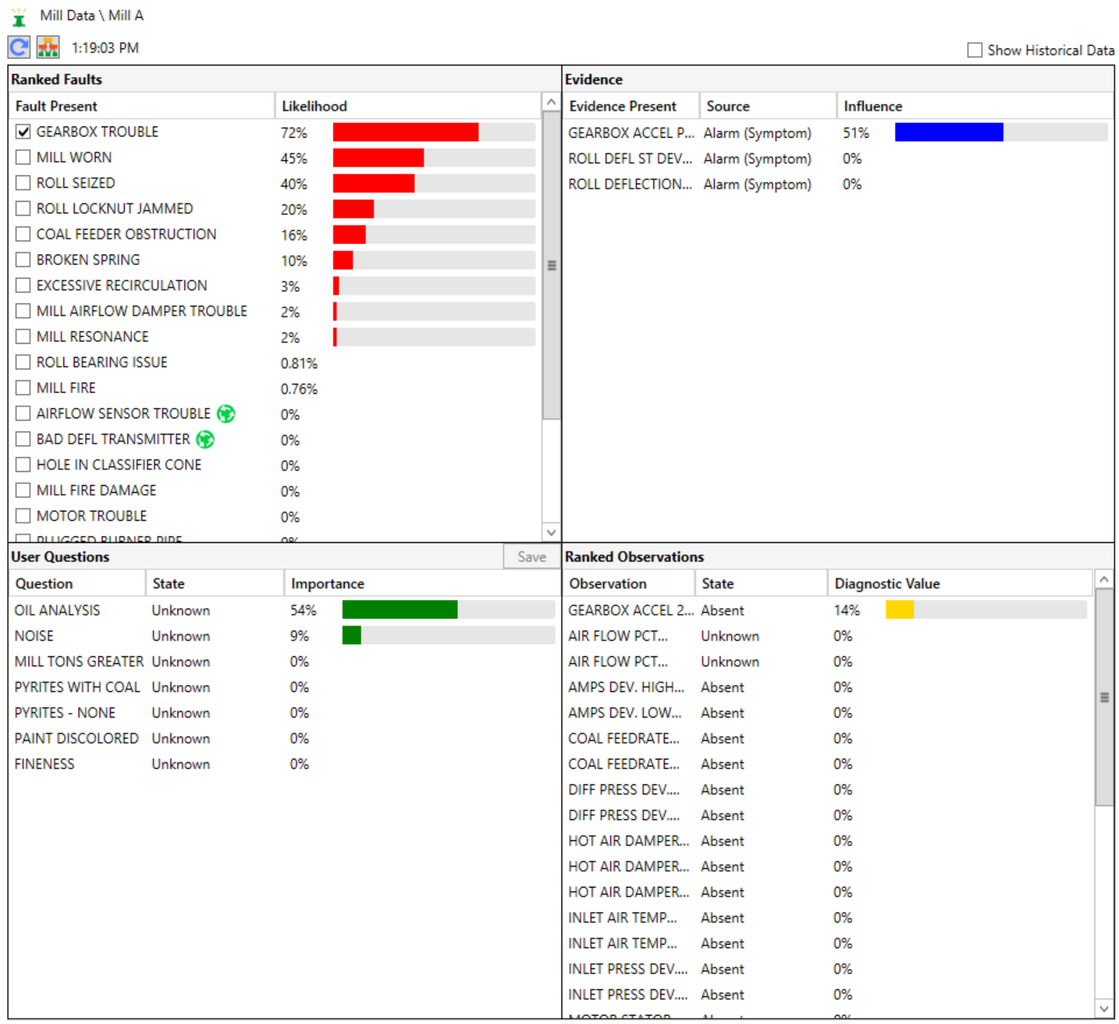


Figure 10: The four-block view of the Diagnostic Reasoner.

The User Questions block will contain any questions that are not present or are unknown; otherwise they will be included in the Evidence block. The importance of the remaining questions will also be shown in this box. The Ranked Observations contains any other possible symptoms that are absent or unknown.

The Ranked Faults box can guide decision making for the end user based on the likelihood of possible faults. When first opened, it will contain the current state of the diagnostic for that asset based on available information. The tool can be used for experimentation by tweaking the other inputs and answering user questions in the other boxes. The likelihood of the faults will adjust accordingly, without disrupting the live system.

Application to Fleet-Wide Monitoring: APR technologies have been employed by many organizations to aid in organizing and managing multiple assets with large amounts of sampled data. Utilizing the results of APR as observations in a causal network will allow these organizations not only the early detection of faults, but aid in diagnosing the root cause. This can assist in outage planning and preparation, as well as help mitigate and avoid catastrophic failures, preventing the possibility of hundreds of thousands, or even millions, of dollars of lost revenue.

References:

[1] Misal, N., Kleineke, C. and Santucci, M., 2010, Adapting Advanced Pattern Recognition Software to Predict and Optimize Pulverizer Performance, *MFPT Proceedings 2010*

[2] Pearl, J., 1985, Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning (UCLA Technical Report CSD-850017). *Proceedings of the 7th Conference of the Cognitive Science Society,* University of California, Irvine, CA., pp. 329–334.

[3] Pearl, J., 2002, *Causality: Models, Reasoning and Inference*, Cambridge University Press, New York, NY, ISBN 0-521-77362-8, 384 p.

[4] Wikipedia contributors. "Bayes' theorem." *Wikipedia, The Free Encyclopedia*. Wikipedia, The Free Encyclopedia, 18 Apr. 2018. Web. 18 Apr. 2018.

[5] Druzdzel, M. J. and Simon, H. A., 2013, Causality in Bayesian Belief Networks, *Proceedings of the Ninth Annual Conference on Uncertainty in Artificial Intelligence (UAI-93)*, Morgan Kaufmann Publishers, Inc., San Francisco, CA., pp. 3-11.

[6] Stuart, A. and Ord, K., 1994, *Kendall's Advanced Theory of Statistics, Volume 1, Distribution Theory,* Wiley, Hoboken, NJ, 676 p.