Anomalies Detection in the Behavior of Processes Using the Sensor Validation Theory

Pablo H. Ibargüengoytia(✉), Uriel A. García, Alberto Reyes, and Mónica Borunda

Instituto de Investigaciones Eléctricas, Cuernavaca, Morelos, Mexico
{pibar,uriel.garcia,areyes,monica.borunda}@iie.org.mx

Abstract. Behavior can be defined as combination of variable's values according to external inputs or environmental changes. This definition can be applied to persons, equipment, social systems or industrial processes. This paper proposes a probabilistic mechanism to represent the behavior of industrial equipment and an algorithm to identify deviations to this behavior. The anomaly detection mechanisms, together with the sensor validation theory are combined to propose an efficient manner to diagnose industrial equipment. A case study is presented with the failure identification of a wind turbine. The diagnosis is conducted when detecting deviations to the turbine normal behavior.

Keywords: Anomaly detection · Model of behavior · Bayesian networks · Wind turbines

1 Introduction

> Anomaly detection refers to the problem of finding patterns in data that do not conform with the expected behavior [3].

In this sentence, several issues are identified and tackled in this paper. These are:

Behavior: a probabilistic model for the behavior of a process is proposed. Specifically, Bayesian networks are used to represent the relation that some variables have with others in a process. This is precisely the function of Bayesian networks.

Patterns in data: related variables maintain combinations of values according to the dynamic of the process. Some combinations respond to some contexts while different combinations or patterns corresponds to other contexts. Some patterns represent normal behavior while others represent failures.

Expected behavior: when a process changes due to external inputs or context changes, the process behavior can be recognized or expected using knowledge on the process dynamics.

This paper proposes a probabilistic model for representing the behavior of a process and a mechanism to detect patterns that reflects deviations of the

M. Montes-y-Gómez et al. (Eds.): IBERAMIA 2016, LNAI 10022, pp. 14–24, 2016.
DOI: 10.1007/978-3-319-47955-2_2

expected behavior. The idea is to collect historical data from a process when it is behaving properly according to experts. The historical data is formed by time series of several variables. The data forms a matrix where columns are the variables and the rows are the samples, measurements or instances. In this paper, the terms variable, sensor and node are used interchangeable.

This project develops the approach of anomaly detection in the behavior for the diagnosis of wind turbines. Historical data is collected and used to identify and model the normal behavior of the turbine. Related work has been revised for the diagnosis of wind turbines using advanced computational algorithms and artificial intelligence. Most of the consulted work is subscribed to the Condition Monitory (CM) community. The survey in [8] include the most common methods for fault detection. The main methods are vibration analysis, acoustic emission, ultrasonic testing techniques, oil analysis, strain measurement, thermography, shock pulse methods, radiographic inspections and others. However, those are traditional methods that usually require experts in the field and complex models difficult to construct and maintain. Computational methods include the work by [11], based on ontology and Failure Mode, Effects and Criticality Analysis (FEMCA). With that analysis, the method creates ontology and a knowledge base that is used on a expert system shell. However, uncertainty management is not considered. Other computational method is the work reported in [10] that also considers behavioral models obtained with SCADA historical data. However, their approach follows an adapted neuro-fuzzy interference system (ANFIS), but they make no distinction between the different operating modes of the turbine.

This paper models the behavior of the wind turbine and uses anomaly detection to find deviations to its normal behavior.

This paper is organized as follows. The next section introduces anomaly detection techniques and situates the proposal of this paper. Section 3 briefly explains the developed sensor validation theory. Next, Sect. 4 develops the behavior analysis of processes in order to detect deviations. Section 5 introduces the case study, namely the diagnosis of wind turbines with the detection of deviations in its normal behavior. Initial experiments are presented and discussed. Finally, Sect. 6 concludes the paper and indicates the future work in this project.

2 Anomaly Detection

Anomaly detection is an active area of research that is being used in several fields [3]. In finances, illegal transactions are identified in credit cards purchases. In a computer network, undesirable participants can be discovered. Anomaly detection is the identification of unexpected patterns when dealing with data collected from some process. The data recompilation can be conducted using sensors, measure instruments or even human sources like social networks or conversations.

Different types of anomalies can be detected in data representing the behavior of a process:

point anomalies- a single point is deviated in the observed process.

contextual anomalies- a data instance can be anomaly in some cases and normal in others. These cases are called *contexts*.
collective anomalies- a group of data instances is anomalous with respect to the entire data set.

In this project, the focus is in the contextual anomalies. They are data combinations that can be correct on one context, but invalid behavior in other contexts. For example, the behavior of the electric generator in a wind turbine is diverse when there is low wind (low speed) that when high winds blow in the park (high speed in the generator). For this detection, two kind of attributes are differentiated. First, **contextual attributes** are those that determine the neighborhood or context where the process is developed. Second, **behavioral attributes** are those that characterize the behavior in a specific context. For example, in a process that depends on weather, it is not the same a case in winter that a case in summer. The *season* variable can be considered the context variable, while ambient temperature can be considered a behavioral variable.

Table 1 shows the common techniques utilized in anomaly detection, and some examples of applications that can be found in the literature.

Table 1. Techniques used for anomaly detection and applications where they have been applied according to [3].

Techniques	Applications
Classification based	Cyber-intrusion detection
Clustering based	Fraud detection
Nearest neighbor based	Medical anomaly detection
Statistical	Industrial damage detection
Information theoretic image processing	Text anomaly detection
Spectral	Sensor networks

This paper presents a contribution for the anomaly detection community with the use of Bayesian networks. Also, an application in renewable energy is presented in this paper.

3 Sensor Validation Theory

The sensor validation theory was initially designed to find errors in the readings of sensors in industrial processes [7]. The basic idea is to calculate the probability of the values that a sensor provides, given the current values of the most related variables or sensors. Comparing the estimated value with the current variable reading, it is possible to detect failures between sensors readings.

The sensor validation theory follows the two phase approach that industry uses for diagnosis, namely fault detection and isolation (FDI) [5]. The first phase

detects that there is a failure between the variables, and the second phase isolates and discovers the faulty variables.

The first requirement of the sensor validation theory is the construction of a model that represents the probabilistic relations between variables in the application. This can be done with the learning algorithms for Bayesian networks available in the community. Using historical data corresponding to a normal behavior, the model is constructed. Figure 1 shows the network learned for the wind turbine. Once the model is defined, the sensor validation algorithm indicates that for all variables (or nodes in the model), instantiate all other nodes and propagate to calculate a posterior probabilistic distribution that indicates the probability of the real value. If there is a coincidence, then no failure is detected. Otherwise, an apparent failure is detected.

The sensor validation algorithm can be expressed as follows:

Algorithm 1. Detection algorithm

Require: A node n.
Ensure: Either correct or faulty.
1: assign a value (instantiate) to all nodes except n
2: propagate probabilities and obtain a posterior probability distribution of node n
3: read real value of variable represented by n
4: **if** $P(real\_value) \geq p\_value$ **then**
5: return(correct)
6: **else**
7: return(faulty)
8: **end if**

Where p_value is a threshold that can be adjusted to calibrate preferences in the failure detection. If the application requires the detection of all faults, even with the risk of false alarms, assign a p_value high. On the other hand, for the ability to catch the important failures even with the risk of unrecognized failures, assign p_value lower.

The detection algorithm can only specify a set of variables that contains an apparent failure. The failures are considered apparent since correct variables can be validated with faulty ones producing incorrect failure detections. In order to isolate the real faulty variable, a second stage is required. This is called the isolation algorithm. It is based in a property of Bayesian networks called the Markov blanquet (MB). The MB of a node in a BN is the set of nodes that when instantiated, isolates that node from changes in the nodes outside the MB. The MB of a node is formed by the set of parents, children and spouses of a node in a network [9]. Thus, when the set of apparent nodes coincides with the MB of a node, then this node has a real fault. Utilizing this property, if a fault exists in one of the variables, it will be exposed in all the sensors on its MB. On the contrary, if a fault exists outside a sensors' MB, it will not affect the estimation of that sensor. It can be said then, that the MB of a sensor acts as its protection against others faults, and also protects others from its inside failures. Thus, the

MB is utilized to create a *fault isolation* module that distinguishes the *real faults* from the apparent faults. This isolation stage utilizes a second Bayesian network to identify the real faulty variable.

Figure 2 shows the isolation network corresponding to the network of Fig. 1. The upper layer of nodes represents the vector of real failure. The lower layer represents the apparent failure. All nodes are binary representing {Fault, OK} values. The arcs in the network correspond to the MB of each node. Real faults cause apparent failures in all the variables MB, and the existence of an apparent fault indicate the existence of a real fault in one node in its MB. The apparent faulty nodes are instantiated with the detection cycle, and the propagation calculates the probability of real faults in the variables.

The isolation algorithm can be expressed as follows:

Algorithm 2. Isolation algorithm using the isolation network.

Require: A sensor n and the state of sensor n.

1: assign a value (instantiate) to the apparent fault node corresponding to n
2: propagate probabilities and obtain a posterior probability of all nodes *Real fault*
3: update vector $P_f(sensors)$

The sensor validation theory was successfully utilized in the validation of temperature sensors in a gas turbine of a power plant [6]. However, a fair question is still unanswered: what happen if a failure is detected but the sensor is working properly? The system identifies an improper behavior even if the sensors are working properly. The next section discusses the proposed model to detect deviations in a process behavior based on the sensors behaviors.

4 Proposed Model for Anomaly Detection in the Behavior

This paper utilizes the Sensor Validation Theory and Anomaly Detection mechanisms to diagnose wind turbines. The diagnosis is based on detecting anomalies in the behavior of turbines. Three steps are required in this diagnosis process:

1. Create a behavioral model, using Bayesian networks considering all the variables that may influence in the behavior.
2. Complete the isolation model and run the sensor validation algorithm to generate a pattern of faulty variables once that a fault is presented in the turbine.
3. Recognize a pattern of faulty variables that represents a failure in the turbine. This is learned with historical data and logbook of the wind turbine.

To create and use a model for anomaly detection in behavior, the next methodology is defined:

1. Select the participating variables among the complete historical SCADA data set. Notice that if many variables are used, more complex models can result and higher computational effort would be necessary for the diagnosis. Expert in the domain may advice in an adequate variables data set that represents the behavior.
2. Clean and discretize the data set. Most of the information is obtained through sensors that are prone to noise and failures. Discretization is required for using Bayesian networks. The number of intervals in the discretization should be chosen to balance computer power.
3. Identify a subset of variables that conform contexts in the process, and define the number of combinations that the context variables form. In this paper, the context variables are wind speed and power generation. The wind turbine behave different with high winds and hence high power, with respect low winds. Four contexts were defined in this work.
4. Separate a training data set for every context from the complete data set.
5. For every one of the contexts, utilize a learning algorithm and construct a Bayesian network that represents the probabilistic relations between the variables. In this paper, the *Greedy and search* algorithm [4] of the *Hugin* [1,2] package was used. Figure 1 shows the model obtained for medium speed winds. This is called the detection network.
6. For all the detection networks, identify the Markov blankets of all nodes and construct the isolation network as shown in Fig. 2 [7]. The isolation network produces a vector with the probability of fault in all variables considered in the model.

One important issue in this methodology is the amount of data recollected to learn the behavior model. The SCADA data set must be enough that most of the combinations of proper behavior are included in the learning data set. There could be a wide variety of normal behaviors between the variables. This is also the reason to separate the history of the system in contexts. Thus, enough data should be considered for every context in the diagnosis. If less data is available, some instances of normal operation of the turbine may be interpreted as abnormal. In this paper, 3 years of data are included in the learning process of the model. However, large periods of inactivity of the turbine produce useless data.

When the models have been defined for a specific application, the system is ready for diagnosis. The following procedure is followed:

1. For validation of the system, establish a data set for testing with the historical data of some known failure in the application.
2. From the testing data set, or from the on–line SCADA, read the current value of all variables.
3. Apply the sensor validation algorithm and identify the real faulty variable(s).
4. Register the pattern of real faulty variables for the diagnosed failure in the process.

The diagnosis is executed on-line, i.e., the response time should be enough for detecting insipient and unexpected failures in the wind turbine. The execution

time depends mainly on the interconnection of the behavior model. If some nodes contain many parents, the CPT tables grow exponential and the Markov blankets become also larger. If this happen, both inferences, on the detection and isolation networks, take longer time to execute. The recommendation is to use a learning algorithm that produces less interconnected possible model, as the Greedy algorithm used in this project. This algorithm allows limiting the maximum number of parents for nodes in the model. The time spent in one diagnosis cycle was 27 ms for the experiments of this paper. Nevertheless, the data set collected corresponds to periods of 5 min. If the diagnosis cycle corresponds to 5 min, and if the execution time of the system is below a second, then the response time is appropriate even with more complex behavior models.

The next section exemplifies the proposal in this paper for the diagnosis of a wind turbine.

5 Experiments and Results

The anomaly behavior detection was applied to a case study: the diagnosis of wind turbines.

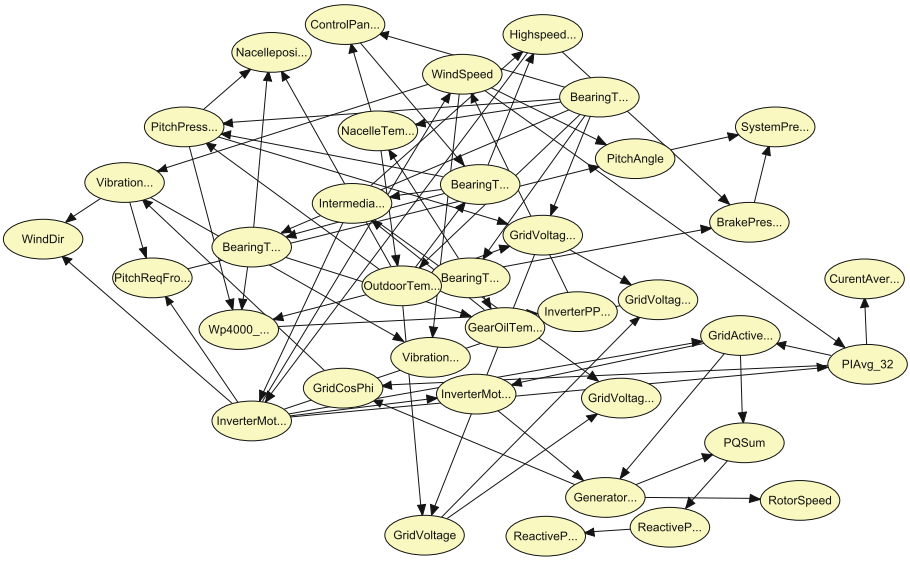

Fig. 1. Behavior model learned from historical data. It is also the detection network.

Wind turbines are devices that capture the kinetic energy of the wind and convert it first to mechanic energy, when the blades rotate, and then convert it in electric power. This power generation represents the higher percentage of renewable energy generation in the world. In Mexico, 1 % of the total generation

was produced by wind farms in 2013 and it is predicted that the produced power coming from clean energies, mainly wind energy, will reach 35 % by 2024.

The Electrical Research Institute (IIE in Spanish) possesses an experimental field with one wind turbine with the capacity to generate 300 kW. The wind turbine is controlled through a SCADA (supervisory control and data acquisition) system. The SCADA program has the function to store historical data of all variables values every 5 min. The total number of variables stored is 76. From those 76 variables, only 34 variables can be used to represent the turbine behavior.

In order to create the probabilistic behavior model, a specific context is chosen. The context variables are wind speed and power generation. For this experiment, data from normal behavior were filtered for the context of wind speed from 3 to 12 m/s, and a generation from 10 to 200 kW. Thus, historical data consisting of 3,300 registers from March 2013 to July of 2014 were selected to train the model. The network shown in Fig. 1 was learned using the *Greedy and search* algorithm [4] of the *Hugin* [1,2] software package. Using this network, the isolation network of Fig. 2 was obtained.

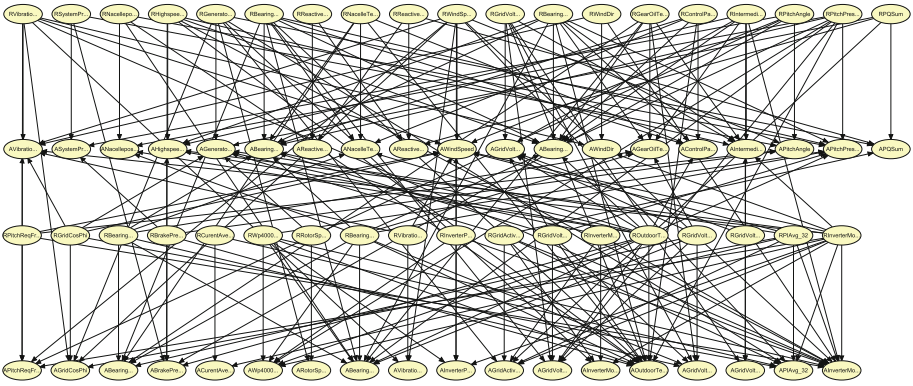

Fig. 2. Isolation network resulted from the model of Fig. 1.

To validate the system, historical data from August 2014 were selected for testing. Figure 3 shows the behavior of the wind speed and the power generation values. As can be noticed in this figure, the turbine was fired and protected to stop generation after a failure was detected by the operator. The turbine was generating around 60 kW when it drops to zero generation. The wind turbine operator informs that the failure detected at that time was the break of a screw causing the yaw mechanism unbalanced.

Figure 4 shows the results of the anomaly detection in the behavior of the turbine. The vertical axis represents the probability of failure of the variables. The horizontal axis represents the numbered instances of data selected for testing. According to Fig. 3, the turbine was behaving properly until something happen

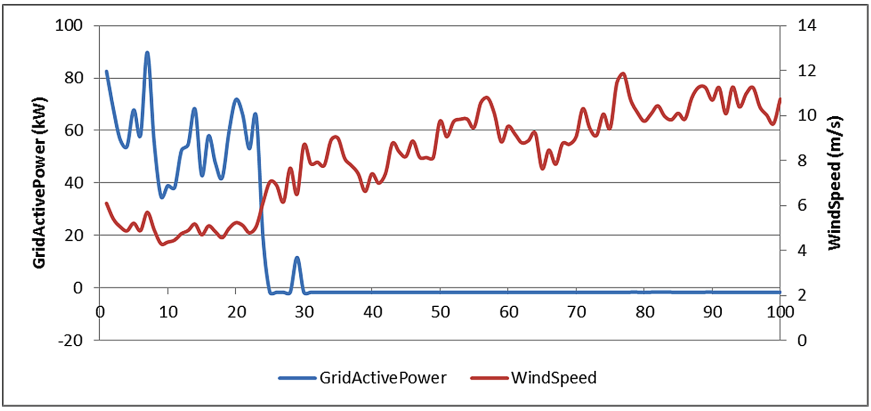

Fig. 3. Historic data from December 2014. Notice the difference in the behavior of the power generation variable and the wind speed.

and drop the generation to zero. This is shown around instance 25 in the graph of Fig. 3. At this same moment, some variables are found with an abnormal behavior. They are for example {BearingTempB, NacelleTemperature, NacellePosition, Vibration2WP4084_1}. According to the wind park operator and his logbook, the failure reported was a decalibration of the yaw break caused due to the loss of a screw. The yaw is the mechanism that faces the turbine to the direction of the wind in order to catch most of the wind energy. With a failure in this system, the nacelle could face a different angle, some extra vibration occurs and the bearings may increment the temperature. Therefore, this case study shows that most of the variables with failure detected as shown in Fig. 4 refer to the yaw system. Other faulty variables may not be completely related like the nacelle temperature. However, the behavior corresponds to a failure in the yaw position. Several tests are needed to complete the relations between changes in the normal behavior and actual failures in the wind turbine.

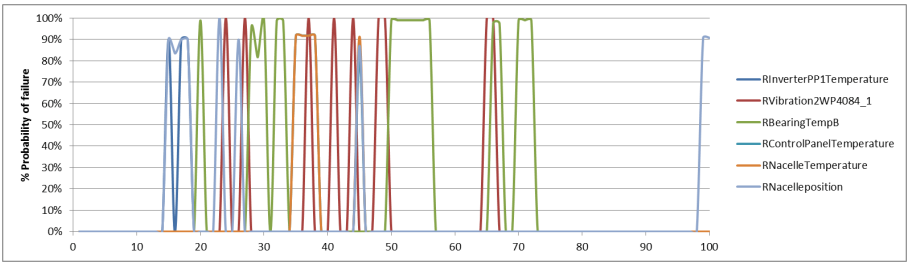

Fig. 4. Results of the anomaly detection. The vertical axis represents the probability of failure of the variables. The results are at the same time than the graph in Fig. 3

6 Conclusions and Future Work

This paper has described the use of sensor validation theory and anomaly detection techniques combined to build an online diagnosis system for wind turbines. The contribution of this project includes the generation of behavioral models based on Bayesian networks and an original form to detect anomalies in the behavior. The case study is the diagnosis of wind turbines where a model was constructed using SCADA historical data and filtering the different contexts proposed. Experiments show that it is possible to identify incipient deviations of normal behavior and the identification of the wind turbine failure.

Even with the promising results obtained in the experiments, several questions remain and require future work. Some of these are the following:

- Can a difference be identified between a failure in a sensor or a failure in an equipment?
- Which is the best way to identify the contexts in an application?
- Do different contexts require different models?
- Is it possible to guarantee a pattern of identified faulty variables for each fault in equipment?
- Is it possible to separate an application in sub–modules and apply this technique for each module?
- Is it worth to use dynamic Bayesian networks to consider time in the failure detection?

Acknowledgements. This work is a preliminary result of the P12 project of the Mexican Center of Innovation in Energy (CEMIE-Eólico), partially sponsored by Fund (FSE) CONACYT-SENER Energy Sustainability, and at the IIE, under the project 14629. Authors also thank the anonymous referees for their insightful comments.

References

1. Hugin expert, hugin expert A/S. Aalborg, Denmark (2000)
2. Andersen, S.K., Olesen, K.G., Jensen, F.V., Jensen, F.: Hugin: a shell for building bayesian belief universes for expert systems. In: Proceedings of the Eleventh Joint Conference on Artificial Intelligence, IJCAI, pp. 1080–1085, Detroit, Michigan, USA, 20–25 August 1989
3. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: a survey. Technical report TR 07-107, University of Minnesota, USA (2007)
4. Chickering, D.M.: Optimal structure identification with greedy search. J. Mach. Learn. Res. **3**, 507–554 (2002)
5. Frank, P.M.: Fault diagnosis in dynamic systems using analytical and knowledge based redundancy- a survey and some new results. Automatica **26**, 459–470 (1990)
6. Ibargüengoytia, P.H., Sucar, L.E., Vadera, S.: Real time intelligent sensor validation. IEEE Trans. Power Syst. **16**(4), 770–775 (2001)
7. Ibargüengoytia, P.H., Vadera, S., Sucar, L.E.: A probabilistic model for information and sensor validation. Comput. J. **49**(1), 113–126 (2006)
8. Márquez, F.P.G., Tobias, A.M., Pérez, J.M.P., Papaelias, M.: Condition monitoring of wind turbines: techniques and methods. Renew. Energy **46**, 169–178 (2012)

9. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Francisco (1988)
10. Schlechtingen, M., Santos, I.F., Achiche, S.: Wind turbine condition monitoring based on scada data using normal behavior models. Part 1: system description. Appl. Soft Comput. **13**, 259–270 (2013)
11. Zhou, A., Yu, D., Zhang, W.: A research on intelligent fault diagnosis of wind turbines based on ontology and FMECA. Adv. Eng. Inform. **32**, 255–270 (2014)

http://www.springer.com/978-3-319-47954-5

Advances in Artificial Intelligence - IBERAMIA 2016
15th Ibero-American Conference on AI, San José, Costa Rica, November 23-25, 2016, Proceedings
Montes y Gómez, M.; Escalante, H.J.; Segura, A.; Murillo, J. de D. (Eds.)
2016, XVI, 428 p. 113 illus., Softcover
ISBN: 978-3-319-47954-5