USE OF ARTIFICIAL INTELLIGENCE TECHNIQUES FOR PROGNOSTICS: NEW APPLICATION OF ROUGH SETS

Janusz Zalewski [1], Zbigniew Wojcik [2]

[1] Deptartment of Software Engineering, Florida Gulf Coast University
Fort Myers, FL 33965, USA
zalewski@fgcu.edu, http://www.fgcu.edu/zalewski/

[2] Smart Machines
13703 Morningbluff Drive, San Antonio, TX 78216, USA
wojcik.zbigniew@sbcglobal.net, http://smart-ma.com/

Abstract: *The objective of this paper is to set the context for the potential application of rough sets in prognostics. Prognostics is a field of engineering, which deals with predicting faults and failures in technical systems. Engineering solutions to respective problems embrace the use of multiple Artificial Intelligence (AI) techniques. The authors, first, review selected AI techniques used in prognostics and then propose the application of rough sets to build the system health prognostication model.*

Keywords: *Prognostics, Model-based Algorithms, Data Driven Algorithms, Rough Sets.*

1. INTRODUCTION

This paper builds on a previous limited survey of AI techniques in prognostics [1]. Since the first comprehensive survey of AI methods used for prognostics, completed in 2007 [2], a number of new algorithms based on various AI approaches have been either developed or applied to prognostics problems. Moreover, the entire discipline of Prognostics and Health Management (PHM) has been significantly developed during this time, and even though no specific new survey paper has been published, perhaps with one exception [3], several articles appeared, which summarize the accomplishments thus far, or set up the scene for the future [4-8].

For the purpose of this paper, we adopt the terminology and classification of AI methods for prognostics as outlined in [2]. Fault *diagnostics* is defined as fault isolation and fault identification, that is, making attempts to "determine the location of fault" (fault isolation) and actually "determining what is wrong" (fault identification). Fault *prognostics* is defined as "determining when a failure will occur based conditionally on anticipated future usage", which essentially means predicting how much time is left before a failure occurs. The fundamental term used in prognostics is the Remaining Useful Life (RUL) of a device or other equipment defined as "an estimation of a remaining life of a component prior to occurrence of a failure" [8].

This paper concerns only prognostics of technical systems, including electromechanical and electronic systems, leaving out other disciplines where prognostics is essential for proper operation, such as production systems and medical applications, for example. Respective survey papers exist, which review issues and techniques related to these other areas, for example [9] in medical prognostics.

Limited to the discussion of prognostics in technical systems, this paper is structured as follows. Section 2 discusses the taxonomy of prognostics algorithms, which is followed by a brief review of two basic categories of these algorithms in Section 3 (model based algorithms) and Section 4 (data-driven algorithms), including a discussion of the application of rough sets in prognostics. This is followed by a conclusion in Section 5.

2. PROGNOSTICS ALGORITHMS TAXONOMY

The essential categorization of prognostics algorithms, shown in Fig. 1, is based on the distinction between using analytical models (mathematical equations) for prediction, which leads to *model-based* approaches, and deriving

conclusions about system behavior from the analysis of measurement data, which leads to *data-driven* approaches [2]. This categorization is not disjoint, however, because analytical models can be enhanced and refined with the use of experimental data, and behavioral data can be analyzed using mathematical methods, which leads to *hybrid* approaches. Nevertheless, for the purpose of this paper such categorization seems logical, since it captures the essential difference between two views of approaching prognostics.
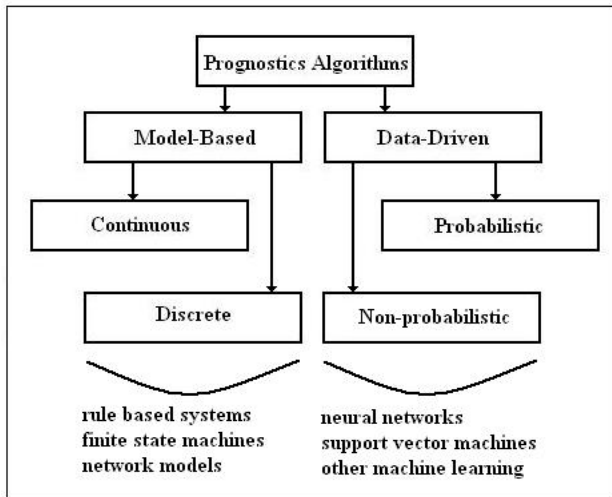

Fig. 1 – Taxonomy of prognostics algorithms

Model-based approaches can vary significantly, since there are a variety of mathematical theories one can use to describe the operation of technical systems, and they may include: differential equations, finite state machines, rule based systems, queuing theory, and other discrete models, such as various sorts of networks: Petri nets, Bayesian belief networks, and others. Many of these models are not purely analytical and use inference techniques typical to AI.

Data-driven models, on the other hand, rarely include analytical knowledge, therefore they all involve some sort of inference. When probability densities governing the distribution of data are available, then one can use well defined statistical methods, for example, linear regression. However, exact knowledge of probability distributions is rarely the case, so methods, which handle incomplete information are applied, including neural networks, support vector machines, and other machine learning approaches, in general.

Even though model-based systems are ideally based on mathematical equations and may not involve any AI reasoning techniques, this is rarely the case. So before venturing into the true AI methods, we first describe the principle of model-based approaches.

3. MODEL BASED ALGORITHMS

Due to a relatively large amount of various theories on which prognostics models can be based, model-based prognostics methods range, accordingly, from rather simple, deterministic, physics based models, to more complicated models, where the determination of critical model parameters may be very involved. Essentially, model-based prognostics can be derived from model-based techniques for fault detection, isolation and diagnosis, which have been used for a relatively long time [10-11]. The process illustrated in Fig. 2, developed for fault diagnosis, can be extended for prognostics by adding prediction estimates for respective model parameters. Most of the applications of model-based techniques do not necessarily use AI methods, unless they are hybrid, if analytical models are accurate enough not to require additional techniques to reason about data.
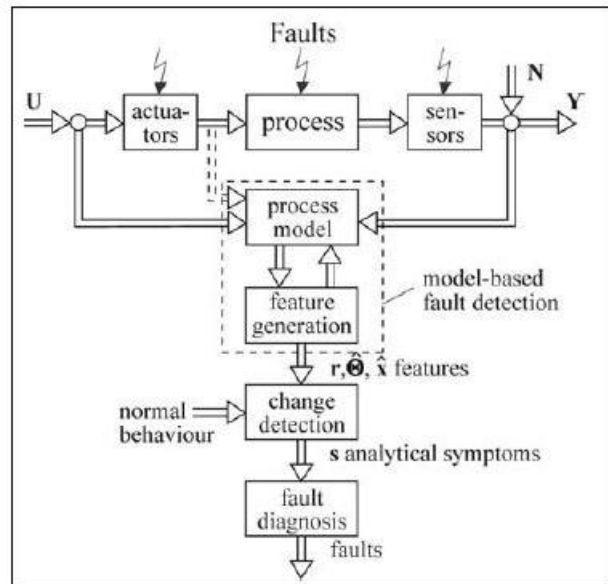

Fig. 2 – General scheme of model-based detection and diagnosis [10]

3.1. REVIEW OF SELECTED MODEL BASED TECHNIQUES

Luo et al. [12] describe such a model-based prognostic process, which starts with building a system model by using singular perturbation techniques of control theory. Based on the system model, Monte Carlo simulations are performed under different load conditions. Then, a prognostic model for component degradation, using Interacting Multiple Models (IMM) and parameter estimation method, is constructed based on simulated data. The degradation measure is tracked using and IMM estimator and the remaining life of the system is estimated under different future usage scenarios. The

process is illustrated in a suspension system of an automobile, with a failure mode involving a crack in the suspension spring caused by fatigue.

Thumati and Jagannathan [13] propose a more sophisticated scheme addressing both state and output faults by considering separate time profiles. The faults are modeled using input and output signals of the system. The fault detection comprises online approximator in discrete time with an adaptive term. An output residual, generated by comparing the fault estimator with that of measured system output, is used as a fault indicator when it exceeds a predefined threshold. A parameter update law is developed to estimate the time to failure, considered as a first step to prognostics. The effectiveness of this approach is demonstrated using a fourth-order multiple-input multiple-output satellite system.

There have been other multiple papers published on model-based prognostics. Hines and Garvey [14] developed a nonparametric prognostic model named PACE (path classification and estimation), which is applied to predict RUL of the steering system of a drill used for deep oil exploration. Sankavaram et al. [15] describe a hybrid approach that uses both a model-based prognosis applied to automotive suspension system and data-driven prognosis to Li-Ion batteries and electronic systems. Application of model-based prognostics methods extends beyond typical machinery and electronics, and is also used in production systems, for example [16].

3.2. PARTICLE FILTERS

Particle filters are the most recently used technique in prognostics and deserve special attention. They are best explained in the context of Kalman filters, used to estimate and predict internal state (over time) of a system described by a set of differential equations of a certain form, given a series of measurements corrupted by noise. The Kalman filter works best in the special case where the system of equations is linear and the system and process noise is Gaussian. Then the Kalman filter provides an optimal analytic solution.

There are several generalizations of the Kalman filter, such as the extended Kalman filter, that make simplifying assumptions about the system's non-linearity, so that the techniques of the linear Kalman filter may be applied to it. However this may cause problems in situations where the system is highly nonlinear or the distribution of the noise is multimodal or highly skewed. It turns out that Particle Filters introduced less than two decades ago [17] are suitable to deal with these sorts of problems.

Particle filters were introduced to solve the problem of state estimation in circumstances where

other methods, such as Kalman filters, are not particularly effective [18]. Particle filters use simulation to provide a numerical solution and are able to handle many more general situations than Kalman filter extensions. The only restriction is that the system must meet the Markov condition, which means that when the system makes an update, the new state depends only on the current state and is independent of any of the past states.

Particle filters are numerical methods for estimating the internal states of Bayesian models that use simulation. The system model describes the way the system updates over time and is given by

$$x_k = f_k(x_{k-1}, z_k)$$

where x_k has the distribution $p(x_k|x_{k-1})$ for $k > 0$ and x_0 has the distribution $p(x_0)$.

The vector x is called the state vector, z is a random vector called the process noise that has a known distribution, and f is called the state transition function. The value of the state at time k depends only on the value of the state at $k-1$ and is independent of the values at all previous times.

The measurement model is given by the following equation:

$$y_k = h_k(x_k, v_k)$$

where the vector y is called the measurement vector, v is a random vector called the measurement noise with a known distribution, and h is the observation function. The value of the measurement at time k depends only on the state at time k and is independent of previous measurements.
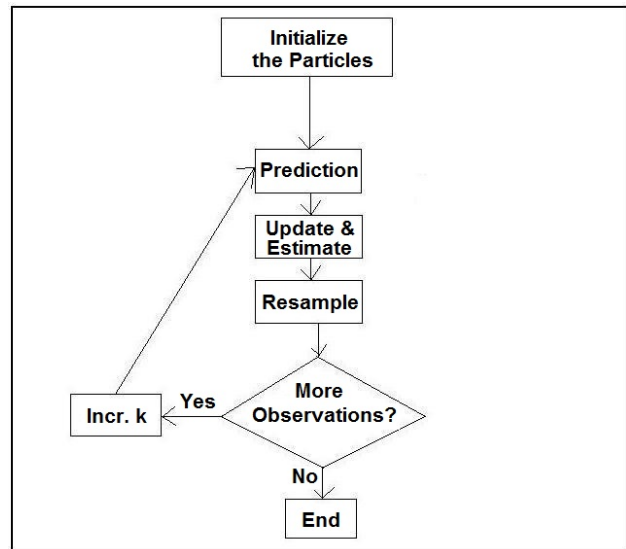

Fig. 3 – Basic steps in the particle filter algorithm

The computational algorithm uses simulation to approximate $p(x_k|y_{1:k-1})$, which is unknown. However it may not be possible to sample from this distribution, so another distribution $q(x_k|y_{1:k-1})$, called importance distribution (proposal) function, is sampled from instead. Although any distribution

with a support (set of values for which the distribution is not zero) that is a subset of the support of the prior may be used for q, the algorithm will perform better for a distribution with a shape close to that of $p(x_k|y_{1:k-1})$.

The algorithm is initialized by taking N samples, called *particles*, from the distribution of the initial state, $q(x_0)$. Let's denote the i-th particle as x_0^i for each i belonging to [1, N]. Each particle is assigned a weight of $w_0^i = p(x_0^i)/q(x_0^i)$. The weights are then normalized by dividing each weight by the sum of all the weights so that the normalized weights $\sim w_0^i$ sum to one. The algorithm then proceeds in three steps for each iteration, as described below and illustrated in Figure 3.

The first step, called *prediction* is to approximate the distribution $p(x_k|y_{1:k-1})$. Finding this exact distribution requires integrating $p(x_k|x_{k-1})p(x_k|y_{1:k-1})$ with respect to x_{k-1} and this integral is not tractable in general, so simulation is used to approximate it. The particles from the previous iteration are updated by taking a sample z_k^i from the process noise for each particle x_{k-1}^i and then evaluating the state transition function: $x_k^i = f(x_{k-1}^i, z_k^i)$.

The second step called *updating* is to estimate the posterior, $p(x_k|y_{1:k})$. Each particle is assigned an initial weight, $w_k^i = p(x_k^i|x_{k-1}^i)/q(x_k^i|x_{k-1}^i)$. The weights are then normalized by dividing each weight by the sum of all the weights so that the normalized weights $\sim w_k^i$ sum to one. Then $p(x_k|y_{1:k})$ may be approximated by the discrete distribution that is equal to $\sim w_k^i$ at each particle x_k^i. An estimate of the state at time k, x_k^*, may now be found by taking the weighted sum of all the particles.

The third step is called *resampling*. Here a sample of size N is taken from this discrete distribution to replace the current set of particles and each new particle is assigned a weight of *1/N*. It is necessary because the set of particles may consist of many particles from highly improbable regions that will result in all but one particle having negligible weights after several iterations, a situation called "degeneracy". When resampling is done, particles with low weights are unlikely to be selected and particles with high weights will likely be selected several times. This ensures that more particles are at regions where the true value of the state is likely to be and helps to avoid degeneracy.

The prediction step is now taken for time *k+1* and the algorithm continues until there are no more observations y_k to consider.

On-going research has shown great promise for application of particle filters to prognostics. They can be effectively used to track progression of system state in order to make estimations of remaining useful life (RUL), which is at the core of system prognostics and health management [19-25].

4. DATA DRIVEN ALGORITHMS

Data-driven algorithms are the ones which are truly based on AI techniques, since they require reasoning about data patters. In this paper, we discuss briefly one such technique, Support Vector Machines, and then pursue toward description of a new technique based on rough sets, which to our knowledge has not been used in prognostics before.

4.1. SUPPORT VECTOR MACHINES

One specific method used in prognostics, which has been mentioned in some surveys [2] but not extensively covered, is Support Vector Machines (SVM), also known as maximum margin classifier. SVM is a method considered as a nonlinear regressive model [26], in which the dependence of a scalar *d* on a vector *X* is described by the formula

$$d = f(X) + v$$

where both the function *f()* and the statistic properties of the additive noise *v* are unknown.

All the available information is a set of training data

$$\{(x_i, d_i)\}$$

where $i = 1...N$; x_i is the sample value of the input vector *X*, and d_i is a corresponding value of the output *d*. The problem is to give an estimate of the dependence of *d* on *X*. A nonlinear regression is performed by mapping the input vector *X* into a high-dimensional feature space, in which then a linear regression is performed.
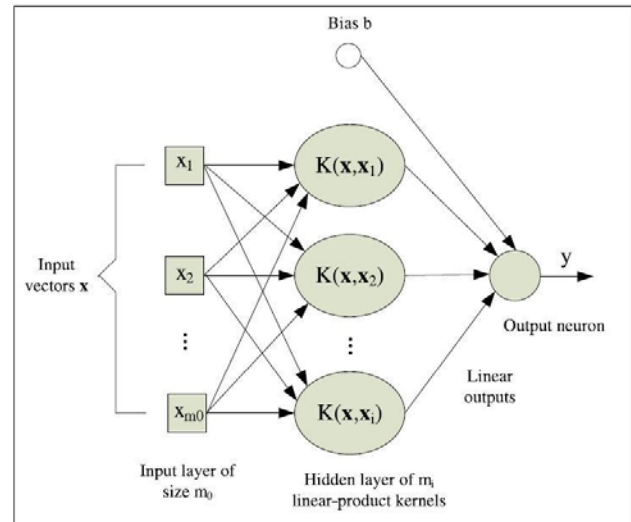

Fig. 4 – Architecture of an SVM [26].

In the architecture of the SVM shown in Fig. 4 [26], the kernel function $K(x, x_i) = \Phi^T(x)\Phi(x_i)$, and $i = 1, 2, ..., m_i$ is the dimension of the feature space. It is also assumed that $\Phi(X)$ is defined *a priori* for all *j*. Given such a set of nonlinear transformations, one may define a formula for *y*, an estimate of *d*, as follows

$$y = \Sigma\, w_j \Phi_j(X) + b, \text{ for } j=1...m_j$$

where w_j for $j=1...m_j$ denotes a linear weight connecting the feature space to the output, and b is a bias.

Caesarendra et al. [26] use SVM as a part of a more sophisticated approach, in a combination with a probability based approach, to predict a final time of a failure in a bearing. SVM is only a part of the entire procedure trained by kurtosis and the target vector to build the prediction model, which is then utilized to predict the final failure time of individual bearing.

While Caesarendra et al. [26] applied this method to the simulated bearing failure data as well as experimental bearing run-to-failure data, Kim et al. in their paper [27] verified the proposed model on bearing failure data of a high pressure LNG pump. Although the training error for classification of six classes using an SVM classifier was relatively high (18.75%), the authors claim that the estimated RUL follows closely the real remaining life of the machine, thus validates the proposed concept as potentially useful for application in industry.

Most recently, Widodo and Yang [28] used SVM in conjunction with survival analysis in machine health prognostics. In their work, survival analysis utilizes censored and uncensored data collected from condition monitoring (CM) and then estimates the survival probability of failure time of machine components. SVM is trained on data input from CM histories that correspond to target vectors of estimated survival probability. After validation, the SVM is employed to predict failure time of individual units of machine components. The method has been verified for simulation and experimental bearing data.

Multiple other applications of SVM's in prognostics exist, including specific technical systems, such as automotive [29], as well as non-technical systems, such as human patients in medicine [30].

4.2. USE OF ROUGH SETS FOR PROGNOSTICS

Rough sets, invented by Zdzisław Pawlak in 1981, are a mathematical method of dealing with incomplete information and uncertainty [31-32]. As such, they seem ideal to making predictions and prognostics. Nevertheless, there have not been any applications of rough sets to prognostics, except one we've found [33]. In this section, we present a method of using rough sets to build a system health prognostics model, and discuss how it fits into the overall taxonomy of AI methods in prognostics.

The rationale behind using a rough sets approach for health degradation estimation and prediction is three-fold:

a) It sets the focus on splitting the data into two parts: completely healthy and partly degraded (with different levels of degradation),

b) Offers tools to measure the level of health ambiguity and health degradation, and

c) Facilitates intelligent data mining, by considering relationships between unambiguous and ambiguous data, known from the psychology of human perception as intelligent.

A. Mathematical Background. The rough sets approach incorporates the Universe U of degradation signals (called test points) and a set R of equivalence relations. Examples of a test points are: (a) trf, the set of measurements of turbine fan vibration level over time, (b) tit, the set of Turbine Inlet Temperatures over time.

Each test point $u \quad U$ has its own corresponding equivalence relation $r \quad R$. Each first level equivalence relation r for each test point u describes the equivalence class of the measurements *"in healthy condition"* not looked for, because the focus is on health degradation. Health degradation of a part described by the test point u is represented by the complement of u/r in u.

Diagnostics of a health problem is frequently based on a specific non-empty intersection of the equivalence relations *"in degraded condition"*. For instance, a disintegration of an engine is likely to happen when

$$TRF/degraded \cap TIT/degraded \neq \emptyset$$

where $degraded \equiv$ *"'in degraded condition"*.

Engine *disintegration* is thus representable by non-empty indiscernibility relation [34] over two or more specific equivalence relations, e.g.

$$disintegration \leftarrow IND\{trf\_degraded; tit\_degraded\}$$
$$= trf/degraded \cap tit/degraded$$

executed on paired elements that occur at the same time.

Unhealthy situations are detected by rough sets variance. Rough sets variance prevents obscuration of degradation information by measuring distance to healthy interval. Statistical population variance of variable X:

$$Var(X) = 1/N\, \Sigma(x\text{-}x^{*})(x\text{-}x^{*})$$

where $(x\text{-}x^{*})$ is a measure of deviation of the values x from their mean x^{*}

The statistical variance obscures degradation information, because it collects into the sum also healthy measurements.

There may be no significant difference between: (a) a signal of a few unhealthy measurements accompanied with a large number of small deviations from the mean and (b) a signal of no unhealthy measurements accompanied with a large number of slightly higher (but still normal) deviations from the mean. Therefore, using rough

sets variance makes sense to measure degradations, because rough sets variance of variable X for N measurements is equal [34]

$$rsVar(X) = Median|x - L(X)|$$

where $L(X)$ is the value of the healthy range of X, nearest to x. The value of x must be outside of the healthy range, otherwise: $x - L(X) = 0$ and is rejected from considering to the median. Only non-zero degradations from the healthy range $L(X)$ count.

Rough sets variance is non-linear. For a large number of elements, the rough sets variance may also be expressed by [34]

$$rsVar(X) = 1/N \Sigma (x - L(X))(x - L(X))$$

or by

$$rsVm(X) = 1/N \Sigma |x - L(X)|$$

Statistical covariance between variables X and Y is equal:

$$Cov(X, Y) = 1/N \Sigma (x - x^*)(y - y^*))$$

Rough sets covariance between variables X and Y for N measurements is as follows [34]:

$$rsCov(X, Y) = Median(x - L(X))(y - L(Y))$$

Also, rough sets covariance provides very high signal-to-noise ratio, as opposed to the statistical covariance, which incorporates the mean. For a large number of elements the rough sets covariance may also be expressed by

$$rsCov(X) = 1/N \Sigma (x - L(X))(y - L(Y))$$

Bringing this to the world of health management, the equivalence class of measurements x taken at test point X of a part represented by the equivalence relation $r = in\ degraded\ condition$ equals

$$x/r = \cap\{x\ \varepsilon\ X: |x| > rsVar(X)\}$$

where $|x| > rsVar(X) \equiv r \equiv$ *"degraded"*. The sign $\cap$ means that $r \equiv$ "degraded" is the indiscernibility relation over all historical data of the test point X.

With the approximate knowledge denoted as $k=(X,"degraded")$, the degraded health at test point X can be approximated by two subsets [32]

$$LOxr = U\ \{\ x\ \varepsilon\ x|r:\ x \sqsubset X\}$$

$$UPxr = U\ \{\ x\ \varepsilon\ x|r:\ x\ crossSection\ X.neq.\ \emptyset\}$$

called, respectively, the lower and upper approximation of health degradation represented by the test point X.

B. Applicability of the Theory. The proposed rough sets mechanism analyzes data recorded from interrelated system components and makes predictions using the degradation levels of these components and health status measurements from historical data.

Statistical failure time does not predict failure of a particular item based on its specific level of degradation, but on statistics of a large number of items. Repairs can be made on a specific part that is less degraded than the statistical one, and an enormous risk is with any part that is degraded more than that statistical. Intelligence is of value here: if one knows, which part is less degraded than

expected, one can let it work for some time longer. But if it is used up too much than expected, then it is too dangerous not to make a repair or replacement somewhat earlier than expected.

Conditional maintenance not only costs less, but prevents unexpected crushes. Influence of one part on degradation of another part is also needed to make better predictions, because one cannot have a sensor of degradation at any part of the system. Influences can be learned and collected into a knowledge database. Without the deterministic knowledge of the degradation level and influences on degradations of other parts, no other health status can be determined except the probability of survival for the item being measured. This assumption is true, e.g., when material fatigue leads to a failure.

Application of the proposed model needs knowledge of events co-existing in time and properties of changes of observations over time. All this knowledge can be acquired in real time from historical databases. The proposed approach makes sense of knowledge.

The proposed data fusion system described in [34] consists of terms, which are the products of ratios of degradation and the ratios of influences on the degradations. The parts degradation data fusion system is deterministic for data sequenced by time. It exploits available degradation signals and gets optimized by removal of insignificant terms, e.g., when one part does not influence another. The influences are computed by rough sets correlations. The model is hierarchical and does not limit the object's complexity. Reasoning is deterministic and rule-like chaining is applicable.

Determinism of time sequences allows using simple inequalities with weights learned at the time of major events such as failure times, near-failures or major repairs. Patterns of degradation are created from historical data at times of known failures or major repairs. Predictions are made by: (i) matching current object conditions with the patterns of degradation learned through historical data, and (ii) computing and applying the increments of degradations and of the influences on the degradations learned from the patterns matched, minute-by-minute, to see when possible failures could happen.

A method is offered for computing elementary increments to extrapolate the actual and predicted patterns of degradation to the points of predicted failures and repairs. The system avoids obscuration of rare degradation measurements by healthy measurements by using rough sets: any healthy status is immediately considered irrelevant. Logistics operations can be optimized by accurately preparing to predicted repairs. Influences of environmental degradations that start to affect

system at the earliest stages can be traced back.

The existing structural modeling and regression models are equalities, not inequalities. The new proposed reasoning is based on testing the inequality in each modeling record, making a rule that fires an alarm if the measurement data implies a significant level of degradation. Records of lower-level components are listed and executed first, and whenever inequality is true, using its right side as an operand of one of the inequalities that is listed later as a higher-level component. Coefficients are not computed as in structural modeling by minimizing an error, but by a new machine learning from time-sequenced events.

By incorporating rough sets degradation covariances, that is, influences of degradations of parts on each other, the proposed method directly rejects obscuration by all measurements indiscernible by healthy status, automatically sets to 0 all influences of no impact on the system degradation, and retains only critical influences degrading the system. By this the proposed system correctly assesses the critical factors for prediction. Relating health status measurement to the equivalence class of the healthy ranges and not to the mean or to median is the key difference of the proposed approach with respect to statistics. Ratios directly representing health problem level are incorporated, and not the values of the variables as commonly used in statistics.

A numerical example of the application of this method to health prognostics of a jet engine can be found in [34].

5. CONCLUSION AND FUTURE WORK

The objective of this paper was to place rough sets in the context of AI techniques used for equipment health prognostics and outline the applicability of rough set theory for the evaluation of health degradation. In this view, it seems that rough sets allow for accurate data mining and prediction, with such new tools as rough sets variance, rough sets covariance, time-sequenced data fusion model, and deterministic machine learning. The rough sets compound covariance proves successful to mine and detect engine disintegration, and new time-sequenced data fusion model to predict it in real time. Determinism of these new tools assures higher accuracy of predictions and increased precision when mining of historical data.

The proposed rough sets prognostic model considers only data sequenced by time [34]. The time makes data mining deterministic in the domain of health degradations. Predictions are more accurate when data are deterministic, or much less amount of data is needed for the same quality of predictions.

Most data analysis tools are concerned with implicit static-like data (i.e., not necessarily sequenced by time) and rather probabilistic type of events, where time plays no significant role. Instead, randomness is important to be statistical. Prediction is inherently time-dependent and it is not easy to apply a statistical inference to it because statistics is basically static. Needles to say, also the basic rough sets theory is concerned with implicit static-like data. The if-then rules generated from this type of data are probabilistic in nature. In contrast to that, the new rough sets model involves only data sequenced by time.

The state-of-the-art machine learning methods do not adapt to the time of an event but to the event itself: such prediction of time of fault or time of repair is not direct and therefore incurs inaccuracies. A near-deterministic approach proposed allows making predictions based on degradation patterns gathered automatically for different applications and at the time moments of failures. Prediction of failures and the health status takes place by deterministic application of these increments to the actual patterns of degradation matched. Statistical (and non-deterministic) predictions would require more cases, which are not always available.

Machine learning is commonly understood as a slow adjustment of weights, and not just by setting the weights using knowledge. Rough sets variance and covariance set values and weights of the new deterministic data fusion model, which constitute knowledge model for time-sequenced prognostic, faster and more accurate compared to statistical or stochastic adjustment of weights.

The advantages of the proposed machine learning that computes each weight as unknown from the equations are: (i) high precision of classification; (ii) high speed of accurate machine learning even from one example to classify one class of degradation cases; and (iii) simplicity, by adjusting weights directly on the level of dichotomization concepts and by this easiness to interpret the machine learning process.

This new approach uncovers large areas of applications for rough sets, including predictions of health problems, diagnostics, mission readiness evaluation, equipment maintenance, optimization of logistics operations, avoidance of unnecessary repairs and high risks of failures with crushes, and data fusion for decision making.

6. ACKNOWLEDGEMENT

prognostics.

7. REFERENCES

[1] J. Zalewski, Latest development in artificial intelligence techniques for prognostics – a brief survey, *Proc. PRIP 2011, 11th International Conference on Pattern Recognition and Information Processing*, Minsk, Belarus, May 18-20, 2011, pp. 21-24.

[2] M. Schwabacher, K. Goebel, A survey of artificial intelligence for prognostics, *Proc. 2007 AAAI Fall Symposium.* Arlington, Virginia, November 8-11, 2007, AAAI Press, Menlo Park, Calif., pp. 108-115.

[3] J.W. Hines, A. Usynin, Current computational trends in equipment prognostics, *Intern. Journal of Computational Intelligence Systems*, (1) 1 (2008), pp. 94-102.

[4] Y. Gao et al., Research status and perspectives of fault prediction technologies in prognostics and health management system, *Proc. ISSCAA 2008, 2nd Int'l Symposium on Systems and Control in Aerospace and Astronautics,* Shenzhen, China, December 10-12, 2008.

[5] Y.G. Bagul, I. Zeid, S.V. Kamarthi, A framework for prognostics and health management, *Proc. 2008 IEEE Aerospace Conference*, Big Sky, Montana, March 1-8 2008.

[6] M. Pecht, R. Jaai, A prognostics and health management roadmap for information and electronics-rich systems, *Microelectronics Reliability*, (50) (2010), pp. 317-323.

[7] E. Zio, The challenges of system health management and failure prognostics, *IEEE Trans. on Reliability*, (58) 2 (2009).

[8] S. Uckun, K. Goebel, P.J.F. Lucas, Standardizing research methods for prognostics, *Proc. PHM 2008, Int'l Conference on Prognostics and Health Management.* Denver, Colo., October 6-9, 2008.

[9] P.J.F. Lucas, A. Abu-Hanna, Prognostics methods in medicine, *Artificial Intelligence in Medicine*, Vol. 15, 1999, pp. 105-119.

[10] J. Gertler, Survey of model-based failure detection and isolation in complex plants, *IEEE Control Systems*, (8) 3 (1988), pp. 3-11.

[11] R. Isermann, Model-based fault detection and diagnosis – status and applications, *Annual Review in Control*, (29) 1 (2005), pp. 71-85.

[12] J. Luo et al., Model-based prognostic techniques applied to a suspension system, *IEEE Trans. on Systems, Man and Cybernetics – Part A: Systems and Humans*, (38) 5 (2008), pp. 1156-1168.

[13] B.T. Thumati, S. Jagannathan, A model-based fault-detection and prediction scheme for nonlinear multivariable discrete-time systems with asymptotic stability guarantees, *IEEE Trans. on Neural Networks*, (21) 3 (2010), pp. 404-423.

[14] J.W. Hines, D.R. Garvey, Non-parametric model-based prognostics, *Proc. 2008 Annual IEEE 2008 Annual Reliability and Maintainability Symposium*, Las Vegas, NV, January 28-31, 2008, pp. 469-474.

[15] C. Sankavaram et al., Model-based and Data-driven Prognosis of Automotive and Electronic Systems, *Proc. 5th Annual IEEE Conference on Automation Science and Engineering*, Bangalore, India, August 22-25, 2009, pp. 96-101.

[16] P. Maier et al., Integrating model-based diagnosis and prognosis in autonomous production, *Proc. PHM'09, First International Conference on Prognostics and Health Management.* San Diego, Calif., September 27 – October 1, 2009.

[17] N.J. Gordon, D.J. Salmond, A.F.M. Smith, Novel approach to Nonlinear/Non-Gaussian Bayesian state estimation, *IEE Proceedings-F*, Vol. 140, 1993, pp. 107-113.

[18] B. Ristic, S. Arulampalam, N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications,* Artech House, 2004.

[19] M. Orchard et al., Advances in uncertainty representation and management for particle filtering applied to prognostics, *Proc. PHM 2008, 1st Intern. Conference on Prognostics and Health Management*, October 6-9, 2008.

[20] M. Daigle, K. Goebel, Model-based prognostics with fixed-lag particle filters, *Proc. PHM 2009, 2nd Annual Conference on Prognostics and Health Management*, San Diego, Calif., September 27 – October 1, 2009.

[21] M. Orchard, G. Vachtsevanos, A particle filtering approach for on-line fault diagnosis and failure prognosis, *Trans. of the Institute of Measurement and Control*, (31) 3-4 (2009), pp. 221-246.

[22] S. Butler, J. Ringwood, Particle filters for remaining useful life estimation of abatement equipment used in semiconductor manufacturing, *Proc. SysTol 2010, Conference on Control and Fault-Tolerant Systems,* Nice, France, October 6-8, 2010, pp. 436-441.

[23] M. Orchard et al., Risk-sensitive particle-filtering-based prognosis framework for estimation of remaining useful life in energy storage devices, *Studies in Informatics and Control*, (19) 3 (2010), pp. 209-218.

[24] E. Bechhoefer, S. Clark, D. He, A state space model for vibration based prognostics, *Proc. PHM 2010, 3rd Annual Conference on Prognostics and Health Management*, Portland, Ore., October 10-14, 2010.

[25] M. Daigle, K. Goebel, Multiple damage progression paths in model-based prognostics, *Proc. 2011 IEEE Aerospace Conference*, Big Sky, Montana, March 5-12, 2011.

[26] W. Caesarendra, A. Widodo, B.-S. Yang, Combination of probability approach and support vector machine towards machine health prognostics, *Probabilistic Engineering Mechanics*, Vol. 26, 2011, pp. 165-173.

[27] H.-E. Kim et al., Machine prognostics based on health state estimation using SVM, *Proc. WCEAM-IMS 2008, Third World Congress on Engineering Asset Management and Intelligent Maintenance Systems Conference*, Beijing, China, October 27-30, 2008, pp. 834-845.

[28] A. Widodo, B.-S. Yang, Machine Health Prognostics Using Survival Probability and Support Vector Machine. *Expert Systems and Applications*, Vol. 38, 2011, pp. 8430-8437.

[29] D.T. Vollmer, M. Manic, CI-PASM – Computational intelligence based prognostic automotive system model, *Proc. ICIEA 2009, 4th IEEE Conference on Industrial Electronics and Applications*, Xi'an, China, May 25-27, 2009, pp. 3714-3719.

[30] V. Van Belle et al., Survival VSM: a practical scalable algorithm, *Proc. ESANN 2008, European Symposium on Artificial Neural Networks – Advances in Computational Intelligence and Learning*, Bruges, Belgium, April 23-25, 2008.

[31] Z. Pawlak, Rough Sets, *International Journal of Information and Computer Sciences*, Vol. 11, No. 5, pp. 341-356, 1982. Available at: http://chc60.fgcu.edu/Images/articles/PawlakOriginal.pdf

[32] Z. Pawlak, *Rough Sets – Theoretical Aspects of Reasoning about Data,* Kluwer Academic Publishers, 1991.

[33] S. Lee, *An Architecture for a Diagnostic/Prognostic System with Rough Set Feature Selection and Diagnostic Decision Fusion Capabilities*. PhD Dissertation, Georgia Tech, December 2002.

[34] Z. Wojcik, System Health Prognostic Model Using Rough Sets. *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing. Lecture Notes in Computer Science*, Vol. 3641, pp. 522-531, 2005.
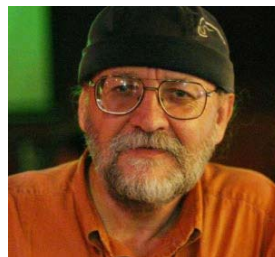
***Janusz Zalewski** is a professor of Computer Science and Software Engineering at Florida Gulf Coast University, in Ft. Myers, Florida, USA. He previously worked at nuclear research labs in* Europe and the U.S. and consulted for the government and industry. He also had fellowships at NASA and Air Force Research Labs. His research interests include real-time embedded systems, prognostics of complex systems, and software engineering education.

***Zbigniew Wojcik** obtained Ph.D. from the Polish Academy of Sciences in the area of pattern recognition. He is the autor of several models, algorithms and methods for prognostics and forecasting, speech signal recognition,* computer vision and image processing, non-blurring image filtering and enhancement, methods of detection of complete image features independent of positions and rotations of the objects, and parallel shape recognition algorithms. He is a Principal at Smart Machines.