CHEMICAL ENGINEERING TRANSACTIONS
VOL. 33, 2013
Guest Editors: Enrico Zio, Piero Baraldi

ISBN 978-88-95608-24-2; **ISSN** 1974-9791
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at: www.aidic.it/cet
DOI: 10.3303/CET1333012

Preprocessing of the Observed Data and the Recognition of the Hidden Signs of Fault

Vitalii Iakimkin, Aleksandr Kirillov, Sergey Kirillov

SmartSys Prognosis Center, Space Technology and Telecommunications Cluster, Skolkovskoye shosse
Moscow, 121353 Russian Federation
SmartTechAppl@ gmail.com

This paper describes methods and algorithms for preprocessing of the observed data to fault prognostics. The main focus is on identifying the hidden signs of fault (nonamplitude failure predictor) and the method, defined their evolution. The signal from the vibration sensor, installed on the engine case, is taken as the initial observed signal.
Further analysis is based on the representation of the signal in the form of coefficients of its wavelet decomposition. Each coefficient of wavelet decomposition is represented as a discrete sequence of data. Further pre-processing signal is a representation of the signal in the form of finite segments of wavelet coefficients. Thus, the transition is made to the high-dimensional vector processes with discrete time. The last stage of preprocessing is the decomposition of vector processes on "amplitude" and nonamplitude "phase" components (K-decompositions).
If the physics of process is known, methods of a nonlinear stochastic filtration are used for allocation of the hidden signs. In the opposite case, particularly interesting from a practical point of view, K-decompositions of the wavelet coefficients segments are analysed. Pseudo-dynamics of phase vector component is analysed to detect of non-amplitude failure predictors. Algorithm and software in automatic mode analyses the evolution of the phase component on the basis of continuous or periodic monitoring. Prognosis of evolution of the phase variable and life time estimate is based on the definition of evolution equations, or by monitoring the entropy characteristics of the pseudo-phase component of K-decomposition. Pseudo-phase component is physically interpreted as characteristics of process regularities. The described algorithm is implemented in PHM computing cluster. For full analysis of life time estimates the cloud computing service is used.

Introduction

Mechanics of internal combustion engines, hybrid engines and mechanical subsystems of ground transport, construction machinery and agricultural machines is very costly in terms of its maintenance and repair. On-board diagnostic systems are not able to diagnose the root cause of the mechanical failure of engine blocks and work equipment on transport. Appearing today PHM methods of analysis and prognosis of mechanics failures, identifying the causes of dysfunctions of subsystems of hydro-gas-dynamic mechanisms have the potential to answer many requests in operational service of transport mechanics and lead to a paradigm shift in the prevention, replacing the costly maintenance and repair on proactive prevention. However, there are certain difficulties in the development of PHM methods for mechanics. Undoubtedly mechanical vibrations of the body of blocks contain the most information the state of its mechanics. However, the ambiguity of data the vibration, the changing nature the vibration depending on the position of the sensor pose the prediction problem in the area of intractable problems. The situation is aggravated by the presence of multiple sources of vibration, the presence of vibration from the process of turbulent fuel combustion, and these vibrations passing through the body of mechanical block appear in the form of a complex nonlinear composition in the observed signal of sensor. Stochastic nonlinear filtering methods used to separation of the contributions from different sources, generally are ineffective because more often dynamic equations of all mechanical system with subsystems of fuel combustion and hydro-

Please cite this article as: Iakimkin V., Kirillov A., Kirillov S., 2013, Preprocessing of the observed data and the recognition of the hidden signs of fault, Chemical Engineering Transactions, 33, 67-72 DOI: 10.3303/CET1333012

gas-dynamics or are extremely complex and non-linear, or are unknown. This paper is devoted the development of the model of hidden signs on the basis of the K-decompositions and the description of their algorithms, described by Kirillov et al. (2011). Readings of vibration and high-precision rotation sensor are analyzed.

1. Model

CH & P model reflected in the article and report Kirillov et al (2012) is taken as the basis of computing algorithms to identify hidden signs of preventive prognosis that is, determining engine dysfunction, leading ultimately to appearance and development failures. The accelerometer signal and rotation sensor signal is taken as a source of information about the hidden signs, shown on Figure 1

The methods of a hidden signs of prediction prognosis are based on the methods of symbolic and topological dynamics. Hidden signs characterize the anomalous trajectory of the phase component of the K-decomposition of the observed signal by means of entropic and other characteristics. Staying on these trajectories, the system evolves with a maximum speed in the sets of dynamical regimes with high failure risks.

Are there any predictor of failures before boundaries of physical irreversibility? This is the most important issue. It adjoins number of perspective tasks related to the problems of self-treatment, self-maintenance, and therefore prolonging life of a mechanical system to theoretically possible time frame of the natural aging of the material.

If such signs exist, they should be degenerate relatively a stochastic vector process. That is, these signs are hidden. For a Gaussian vector process, a group of degeneracy of PDF in space $R^{N^*}$is $O(N^*)$. Factoring the group of degeneracy, on its subgroup of isotropy $O(N^*-1)$, finally it is obtained a homogeneous space

$$D = \frac{O(N^*)}{O(N^*-1)} = S^{N^*} \qquad (1)$$

This homogeneous space by analogy with similar spaces in the theory of defects of condensed matter is called the space of degeneracy. However, in the case of the analysis of stochastic processes, the situation is more complicated. Since the dimension N^* determined from the properties of the process, i.e. N **-** changeable number,$1 \ll N^* \leq \infty$.

Thus, in the general case there is the sequence of groups (1)

$$\subset \boldsymbol{O}(\boldsymbol{N}^* - \mathbf{1}) \subset \boldsymbol{O}(\boldsymbol{N}^*) \subset \boldsymbol{O}(\boldsymbol{N}^* + \mathbf{1}) \subset \boldsymbol{O}(\boldsymbol{N}^* + \mathbf{2}) \subset \cdots. \qquad (2)$$

$$\subset \frac{O(N^*)}{O(N^*-1)} \subset \frac{O(N^*+1)}{O(N^*)} \subset \frac{O(N^*+2)}{O(N^*+1)} \subset \cdots. \qquad (3)$$

and the sequence of degeneracy (Eq .2). At the transition to continuous time and when $\boldsymbol{N}^* \to \infty$, limit of sequence (Eq. 3) is a subgroup of rotation group of a nuclear space of a stochastic process.

Finally, by projecting the observed vector process of degeneration on space *D*, a set of hidden processes on homogeneous manifold of degeneration of multidimensional PDF are obtained. Hidden processes, i.e. processes on sequences of degeneracy spaces are related with the dynamics of the high complexity mechanisms. For example, in cases when the flows, generated by the degeneration groups, are countably-multiple Lebesgue spectrum, it may the definition of changing latent signs by calculating the Kolmogorov entropy or its estimates using the ε-entropy

Following the work of Martin and England (1981) entropy and the metric characteristics of finite segments of the wavelet coefficients of the observed signal are used as a hidden signs of preventive prognosis.

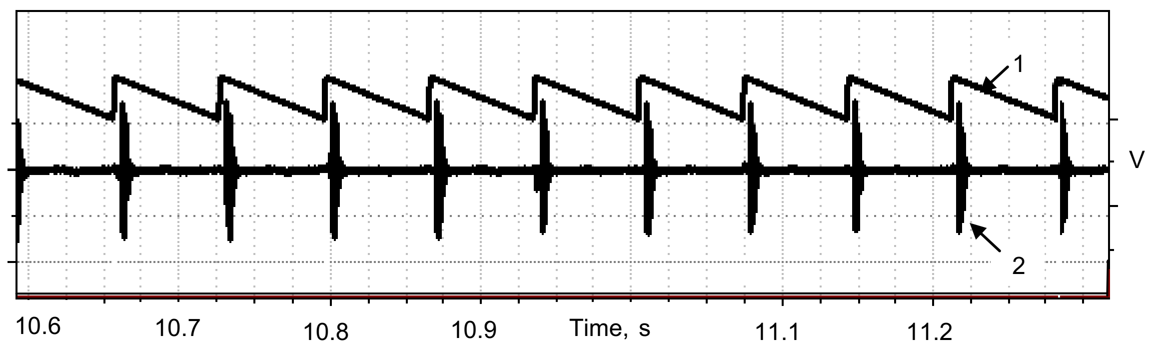

Figure 1: Signals are as a source of information about the hidden signs. 1- rotation sensor signal φ(t), 2 - accelerometer signal S(t)

After the procedure of secondary discretization, signal *s* represent as a vibration signal in angular variable $\boldsymbol{\varphi}$, $S'(\boldsymbol{\varphi})$, $\varphi \in [0,4\pi]$.
At the next step a signal is represented by a set of wavelet coefficients $\{{}_0^N w_{j,k}\}$. The basis for the statistical method is to construct empirical PDF, two-dimensional empirical PDF and conditional empirical PDF of the coefficients of wavelet decomposition of vibration signals at a fixed scale *j* and translation k, where (*j, k*) - the index of wavelet decomposition coefficients and *N*-number of cycle.
In the mechanisms of reciprocating or rotary action sensors signals are stochastic or chaotic quasi-periodic processes. This fact allows to represent each observed signal in quasi-period as an implementation of a stochastic process.
Finally, ${}_0^N w_{j,k}$ is a stochastic process with discrete time *N*.

$$\{{}_0^N w_{j,k}; N = 1,2,3 \dots N^*\} = \{\boldsymbol{R_i}: i = 1,2,3 \dots\}. \tag{4}$$

Next vector process is transformed as follows (K-decomposition):
$\boldsymbol{R_i} \to A_i \boldsymbol{e_i}$, where
A_i –Euclidean norm of finite segments, $\boldsymbol{e}_i \in S^{N^*}$, S^{N^*}- the unit sphere of dimension N^*
Input finite segment $\boldsymbol{e}_i$ comes in a block **C**1 of pre-processing, analysis and accumulation of statistics:
1. Calculation of ε-capacity of the set of finite segments$\{\boldsymbol{R}_i; \mathrm{i} = 1,2,3 \dots\}_i$;
2. Calculation of ε-entropy of the set of finite segments $\{\boldsymbol{R}_i; \mathrm{i} = 1,2,3 \dots\}$
3. Calculation of the correlation dimension $\{\boldsymbol{R}_i; \mathrm{i} = 1,2,3 \dots\}$
4. Determination of rate of change of these characteristics on the chronological database and the current value at constant or periodic monitoring.
Block scheme of a computing algorithm is presented in Figure 2
Then the calculated value of the capacity, entropy, correlation dimension and the set of segments $\{\boldsymbol{R}_i; \mathrm{i} = 1,2,3 \dots\}$ are transmitted to the block **C2** for further processing on the condition that value $\boldsymbol{R}_i$ far from the region of states of system $\boldsymbol{R^c}$ and evolve in some neighbourhood of the etalon state $\boldsymbol{R^e}$**.** Purpose of processing in block C of multidimensional components of the signal of basic sensor.
Signs of the causes of stage C (transition IV - II) are changes in dynamic trajectories of monitored mechanical system relatively the etalon. These changes are cause of the modes of mechanism operation, leading to the appearance and development of early signs of failures. These changes are characterized by changes in the structure of regularities in the final segments of the wavelet coefficients $\{{}_0^N w_{j,k}; N = 1,2,3 \dots\}$ of telemetry data signals and are characterized by a change of Kolmogorov complexity, conditional Kolmogorov complexity, Kolmogorov entropy, and also by various fractal dimensions and the family of entropy characteristics and entropy metrics.
The set of segments $\boldsymbol{R}_i$ goes in the block **C1** for the calculation of quantity and length periodic orbits of subsegments $\boldsymbol{R}_i$ of all dimensions smaller [N^*/2]
1. By enumerative technique the number of periodic vectors (segments) $\boldsymbol{R}_i$ (equivalently the return points) and the lengths of the periods expressed by number *n* are determined.
There are considered all segments with length *N= 1,2,3,…,N** .The number of periodic segments of dimension $\{N_m\ m = 1,2..\}$ and period n_m determine a set of values $\{M_{n_m}^{N_m}\}$.

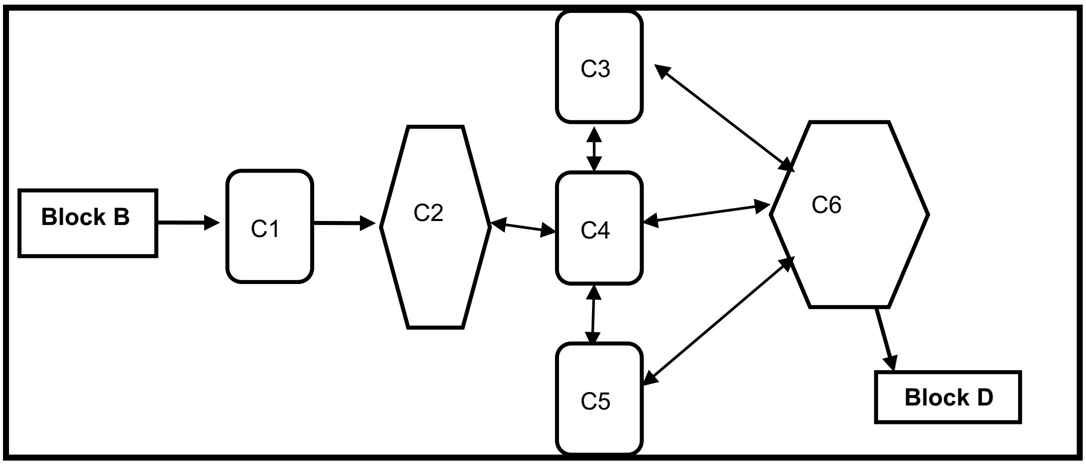

Figure 2: Block scheme of a computing algorithm

2. The estimate is made on basis of calculation п.(1) based on the methods of topological dynamics [ ]:
1. of topological entropy $h^{N_m}\left(M_{n_m}^{N_m}\right)$ on the basis of equality

$$h^{N_m}\left(M_{n_m}^{N_m}\right) = \lim_{n\to\infty}\sup\frac{1}{n}\log M_{n_m}^{N_m}) \quad (5)$$

2. of metric Kolmogorov entropy
3. of a generalized Shannon entropy calculated values together with the set $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ are transmitted to the block **C2** for further processing.
In the block **C2** it is established, to what type of possible following processes the stationary process $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ of the coefficients of wavelet decomposition of the signal $S(\varphi)$ belongs:
1. The sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ is an arbitrary stationary process;
2. The sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ is an arbitrary ergodic stationary random process;
3. The sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ is a sequence of independent test results.
Depending on the results of testing of stationarity, ergodicity and independence the sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ of the coefficients of wavelet decomposition of the signal $S(\varphi)$: together with the values of estimates 1-3 of the block **C1** is transmitted in blocks of estimates:
C3 if the sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ is an arbitrary stationary process;
C4 if the sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ is an arbitrary ergodic stationary random process;
C5 if the sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ is a sequence of independent test results.
In the blocks **C3-C5** estimation is performed:
1. K-complexity of finite segments $R_i = \left\{{}_0^N w_{j,k}; N = 1,2,3\ldots N^*\right\}$;
2. The distances between $R_i, R_{i+1)}$ in the Hamming metric, $Ham(R_i, R_{i+1})$;
3. The conditional K-complexity $(R_{i+1}: R_{i)})$
4. The complexity of individual trajectories $\mathcal{K}(W)$, $W = \left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$
5. The quantities of algorithmic information in the sequence of segments $R_i = \left\{{}_0^N w_{j,k}; N = 1,2,3\ldots N^*\right\}$
6. The quantities of conditional algorithmic information $I\left(R_i: R_{i+1)}\right) = K(R_{i+1}) - K(R_{i+1}: R_{i)})$
C3 the sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ is the arbitrary stationary process;
3.1. To estimate the K-complexity in addition to the set $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ from B1 value of the empirical cumulative distribution function $PDF(N)$ is transmitted;
3.2. Calculation of the coefficients of the polynomial, approximating $PDF(N)$;
3.3. Determination of the set of invariant finite trajectories {A} relatively $PDF(N)$
3.4. Calculation of Shannon entropy *h(A)* для {A};
3.5. Calculation of the Radon-Nikodym derivative for the function$PDF(A)h(A)$, denoted below as $\frac{d(Ph)}{dP}$;
3.6. Estimate of K-complexity on the basis of equality
$\sup_n \lim_{N\to\infty}\frac{K({}_0^N w_{j,k}(n))}{N} = \frac{d(Ph)}{dP}, n -$ is number of segments of length *N*
3.7 Estimate of conditional K-complexity $K(R_{i+1}: R_{i)})$ on the basis of the inequality

$$K(R_{i+1}: R_{i)}) \preccurlyeq l(d(R_i)) \quad (6)$$

l- is length of the word $d(R_i)$, $d(R_i)$ – is the number of elements R_i
C4 the sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ is the arbitrary ergodic stationary random process;
4.1 . To estimate the K-complexity in addition to the set $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ from C1 the value of the metric Kolmogorov entropy *H* is transmitted
4.2 estimate of K-complexity on the basis of equality

$$\sup_n \lim_{N\to\infty}\frac{K({}_0^N w_{j,k}(n))}{N} = H \quad (7)$$

$n -$ is number of segments of length *N*
4.3 estimate of conditional K-complexity $K(R_{i+1}: R_{i)})$ on the basis of the inequality

$$K(R_{i+1}: R_{i)}) \preccurlyeq l(d(R_i)) \quad (8)$$

l - is length of the word $d(R_i)$, $d(R_i)$ - is the number of elements R_i
C5 the sequence $\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ is the sequence of independent test results.
5.1 To estimate the K-complexity in addition to the set$\left\{{}_0^N w_{j,k}; N = 1,2,3\ldots\right\}$ from C1 the value of the a generalized Shannon entropy (q_k) is transmitted

5.2 estimate of K-complexity based on the Kolmogorov theorem

$K(w) \leqslant i(H(q_k) + \alpha(i)$

$$H(q_k) = -\sum_{k=1}^{2^r} q_k log_2 q_k \qquad (9)$$

$\alpha(i)) = C_r \frac{ln\, i}{i}$

5.3 3 estimate of conditional K-complexity $K(R_{i+1}: R_{i)})$ on the basis of the inequality

$$K(R_{i+1}: R_{i)}) \leqslant l(d(R_i)) \qquad (10)$$

l- is length of the word $d(R_i)$, $d(R_i)$ - is the number of elements R_i

Then the set of finite segments (state vectors) $\{\boldsymbol{R}_i; \mathrm{i} = 1,2,3 \ldots\}$ and all calculated or estimated characteristics of K-complexity, conditional K-complexity are transmitted to the block **C6** for further processing

In the block **C6** there are estimated:

6.1. complexities of individual trajectories $\mathcal{K}(W),\ W = \{{}_0^N w_{j,k}; N = 1,2,3 \ldots\}$ on the basis of equality

$$\mathcal{K}(W) = \sup_n \overline{\lim_{N^* \to \infty}} \frac{1}{N^*} \min_{W^{N^*}} K(\{{}_0^N w_{j,k}\}) \qquad (11)$$

$\mathcal{K}(W) = h_\mu$

6.2. The quantities of algorithmic information in the sequence of segments $R_i = \{{}_0^N w_{j,k}; N = 1,2,3 \ldots N^*\}$ denoted as $I(\{{}_0^N w_{j,k}\})$

6.3. The quantities of conditional algorithmic information $I(R_i: R_{i+1)}) = K(R_{i+1}) - K(R_{i+1}: R_{i)})$

6.4 The distances between $R_i, R_{i+1)}$ in the Hamming metric, $Ham(R_i, R_{i+1})$;

6.5. Determination of rate of change of characteristics, calculated or estimated, K-complexity and the information characteristics (hereinafter the characteristics) in the chronological database and the current value at constant or periodic monitoring.

6.6 Assessment of RUL (C) in terms of prognosis on the basis of polynomial approximation of the rate of change of the characteristics

I - is continuous analog *i*, в $P(\boldsymbol{R}, i)$; $\langle\xi\rangle$- is the average value, $\Delta\boldsymbol{R}_i$; $\boldsymbol{r_L}\ (\boldsymbol{s})$ **-** is parameterization of polygonal $\{\Delta\boldsymbol{R}_i; \mathrm{i} = 1,2,3 \ldots\}$.

2. Experiment

Signal of rotation sensor and vibration sensor signal are considered. The non-uniformity of stroke is expressed by time distance at turning the shaft on 360 °. General wear is determined by the second moment of the multidimensional distribution function of finite segments of high frequency coefficients of wavelet decomposition of vibration sensor signal.

Process consisting of pairwise differences, is constructed as analyzed process on the basis of a time series of data of non-uniformity of stroke. The normalized vector process, i.e. the process on the multidimensional sphere of normalized difference vector is considered. ε-entropy of the set of vectors for the observed period is determined.

Fact of change of ε-entropy is detected in the monitoring process. Figure 3. At the same time before change of ε-entropy the degree of wear changed in accordance with the formula of Brownian Motion type with a satisfactory prognosis, with an error of 12 %. In change of ε-entropy the process of general wear is more characteristic to process of random walk in a non-simply-connected domain. Prognosis is implemented by polynomial approximation of ε-entropy. Prognosis and real data indicate as dotted line and solid line, respectively. Cause for the change of ε-entropy has the change in combustion regimes, unrecorded by standard methods of diagnosis. Realized then preventive operations have led to the previous value of ε-entropy and have changed outlook to more favorable**.**

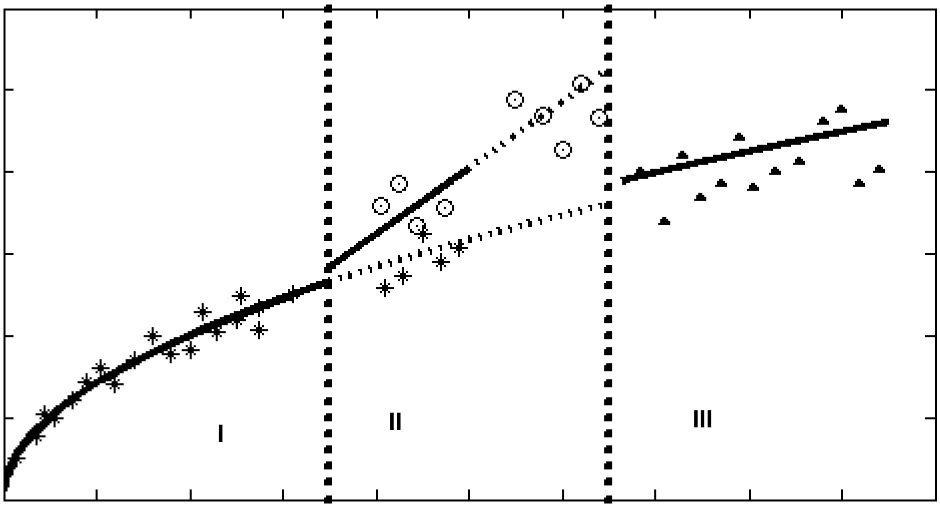

Figure 3: Change of ε-entropy: I - Fixed ε-entropy (normal situation); II - Increasing ε-entropy; III - Reducing ε-entropy to normal situation

Conclusion

Some concrete examples when self maintenance process is possible are demonstrated on the engines. In this case the there is a speech about early signs. For selection of predictive signs the sensor signals (processes) are exposed to the decomposition on the basis of group factorization of process degeneration into two components: conditionally stochastic component and conditionally dynamic component. Thus, the state of the mechanism is defined conditionally dynamic component of the signal decomposition. Analysis of the conditionally dynamical trajectories at engine work determines the earliest prognosis and the appearance of hidden signs of failure of mechanics. The analysis is based on estimates of multiple entropy characteristics.

There are described methods and algorithms of the root cause diagnostics and prognosis. Influence the evolution features of conditionally dynamical trajectories of mechanical unit or block of transport on the physical mechanism of failure or degradation of the material of the transport mechanism is revealed.

In particular, it found that at other being equal conditions, cause of mechanism breakage or material degradation is the mismatch of parameters of mechanism management with the characteristics of conditionally dynamical trajectories in transient regime.

Nucleation in the material the features of the elastic fields with complex geometry, which are the cause of the birth and development of micro-cracks in the material with its subsequent degradation is possible at such mismatches.

The conclusion is that a preventive diagnostics of hidden signs allows to eliminate hidden sign by the variation of control parameters and the change of dynamic modes. Thus, are signs hidden signs on the stage of physical reversibility or return conditionally dynamic trajectory to etalon or close to etalon.

In turn, the made statement means the possibility to create an intelligent self maintenance systems.

References

Kirillov A., Kirillova O., Kirillov S., 2011, Algorithmic method of analysis of time series data for definition of prognostic parameters of engine fault, 3rd International Conference on Advanced Computer Control (ICACC), 138 – 142, DOI: 10.1109/ICACC.2011.6016384

Kirillov, A.; Kirillov, S.; Pecht, M., 2012, The calculating PHM cluster: CH&P mathematical models and algorithms of early prognosis of failure, Conference on Prognostics and System Health Management (PHM), 1 – 11, DOI: 10.1109/PHM.2012.6228771

Martin N, England J., 1981, Mathematical Theory of Entropy, Addison-Wesley publishing, London, England

Zvonkin A.K., Levin L.A., 1970, The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms (In Russian), Uspekhi Mat. Nauk, XXV(6), 85-127.