Prediction of residual life of complex electromechanical systems based on small training samples

P V Lekomtsev
Department of Mechatronic Systems, Kalashnikov Izhevsk State Technical University, Studencheskaya 42, Izhevsk, Russia
lekomtsev@istu.ru

Abstract. A model for predicting the residual life of complex electromechanical systems based on small training samples is presented. The developed algorithm involves the use of a probabilistic model of a complex electromechanical system. The statistical characteristics of the model parameters are determined by training samples, that is, by past experience of the functioning of this or similar system. The developed algorithm is based on recognition and forecasting algorithms based on a Wald’s sequential decision-making procedure, a modified V. S. Pugachev canonical decomposition, and an improved Parzen window density estimation algorithm.

1. Introduction
In the context of Industry 4.0 machines are becoming more complex. The main factor that limits the improvement of the quality and efficiency of production becomes a significant increase in the incidence of failures due to the complexity of the machines.

Thus, the development of a model for determining and predicting the actual resource of electromechanical systems is an urgent task. Currently, to calculate the residual life, a universal measure of which can be taken as a unit of mean time between failures [1], many different methods have been developed [2-5] Assessment of the reliability of electromechanical systems was investigated in the works [6,7].

This article presents an algorithm for predicting the residual life of complex electromechanical systems based on the synergetic unity of recognition and prediction algorithms. The developed algorithm is based on recognition and forecasting algorithms based on a Wald’s sequential decision-making procedure [8], a modified V. S. Pugachev canonical decomposition [9], and an improved Parzen window density estimation [10]. The developed algorithm assumes the use of a probabilistic model of a complex system, the statistical characteristics of the parameters of which are determined by training samples, i.e. from past experience of functioning of a similar system, or the same in a similar situation.

2. Recognition algorithm of random processes
An algorithm is proposed for making a decision on whether the analyzed sample belongs to one of the classes using the sequential Wald's algorithm for a known degree of connectivity λ (which can be approximately determined empirically). Let us write a sequence of density functions with increasing sample using the formula for conditional probabilities:
write the likelihood ratio in (1...4) as each subsequent value of the random process, the results of the previous calculation are used. Let us procedure is the calculation of the thresholds step we will obtain exactly the same likelihood ratios as for (to continue observation) is made if the assigned connection where \(\beta\) - errors of the first and second kind, respectively. From (1...4) it is easy to see the recurrence of calculating the likelihood ratio, i.e. when analyzing the direct calculation of the density function by the Parzen method at each step without taking into account the magnitude of the aftereffect \(\lambda\).

The form of notation (5), (6) due to recurrence has a significant computational advantage over the other methods. The decision \(\gamma_0\) (the hypothesis of the presence of the representation \(H_0\)) is made, if

\[
\frac{\prod_{i=1}^{k-\lambda+1} \omega(x_1, x_2, ..., x_{i-1}) \prod_{i=2}^{k-\lambda+1} \omega(x_1, x_2, ..., x_{i+\lambda-2}) | H_1)}{\prod_{i=1}^{k-\lambda+1} \omega(x_1, x_2, ..., x_{i+\lambda-1}) | H_0} \leq C_{ok}.
\]

(3)

and

\[
\frac{\prod_{i=1}^{k-\lambda+1} \omega(x_1, x_2, ..., x_{i-1}) \prod_{i=2}^{k-\lambda+1} \omega(x_1, x_2, ..., x_{i+\lambda-2}) | H_1)}{\prod_{i=1}^{k-\lambda+1} \omega(x_1, x_2, ..., x_{i+\lambda-1}) | H_0} > C_{1k}.
\]

(4)

From (1...4) it is easy to see that for each subsequent value of the random process, the results of the previous calculation are used. Let us write the likelihood ratio in (1...4) as

\[
l(x_1, ..., x_n) = l(x_1, ..., x_{n-1}) \frac{\omega(x_n | x_{n-\lambda}, ..., x_{n-1}) | H_1}{\omega(x_n | x_{n-\lambda}, ..., x_{n-1}) | H_0},
\]

(5)
or, expressing conditional functions through unconditional,

\[
l(x_n) = l(x_{n-1}) \frac{\omega(x_n | x_{n-\lambda}, ..., x_{n-1}) | H_1}{\omega(x_n | x_{n-\lambda}, ..., x_{n-1}) | H_0}.
\]

(6)
3. Modified V. S. Pugachev canonical decomposition of random processes algorithm

At the recognition stage, the vector $\overline{x}_n$ and the training sample $X_{MN}^H$ are transformed recursively into the vector $\overline{V}_n$ and the training sample $V_{MN}^H$ by the V. S. Pugachev canonical decomposition [9]. This simplifies the application of the Parzen estimation method, which are required later in the forecasting algorithm.

The canonical representation of the vector of coordinates $\overline{V}_n$ provided by the following recursive relations:

$$V_1 = x_1 - M_{x1},$$
$$V_i = x_i - M_{xi} - \sum_{v=1}^{i-1} V_v \cdot a_{iv}, \ i = 2, ..., n,$$

(7)

where

$$D_{V_1} = D_{x1},$$
$$D_{V_i} = r_{ii} - \sum_{v=1}^{i-1} |a_{iv}|^2 \cdot D_{V_v},$$
$$a_{ij} = \frac{1}{D_{V_j}} (r_{ij} - \sum_{v=1}^{j-1} a_{iv} \cdot a_{jv} \cdot D_{V_v}), \ j = 2, ..., i - 1.$$  

(8)

(9)

Here $r_{ij}$ are the elements of the covariance matrix of the random process $\overline{x}_n$; $D_{V_j}$ is the variance of the decorrelated component $V_i$; $M_{xi}$ is the mathematical expectation of the components of the observed vector $x_i$, $i = 1, ..., n$.

4. Prediction algorithm for the training sample

Let us write a sequential nonparametric prediction algorithm for the training sample $V_{MN}$ for the vector $\overline{V}_k$, the first $k$ components of which are determined, in the following form [12]:

$$Y_{k+1} = F_{k+1}(V_{k+1}|V_1, ..., V_k) = \int_{-\infty}^{V_{k+1}} \omega_{k+1}(V_1, ..., V_{k+1}) dV_{k+1}$$

$$Y_{k+2} = F_{k+1}(V_{k+2}|V_1, ..., V_{k+1}) = \int_{-\infty}^{V_{k+2}} \omega_{k+2}(V_1, ..., V_{k+2}) dV_{k+2}$$

$$Y_M = F_M(V_M|V_1, ..., V_{M-1}) = \int_{-\infty}^{V_M} \omega_M(V_1, ..., V_M) dV_M$$

(10)

where $\omega(\xi) = \begin{cases} 1, & 0 \leq \xi < 1 \\ 0, & \xi < 0, \ \xi \geq 1 \end{cases}$

Since the conditional distribution function $F_m(V_m|V_1, ..., V_{m-1})$ non-decreasing, then the search of extremum $|\xi - Y_m| = |\xi - F_m(V_m|V_1, ..., V_{m-1})|$ is easily done by dichotomy. In this case, the condition

$$|\xi - F_m(V_m|V_1, ..., V_{m-1})| < \delta_m$$

(11)

is checked, where $\delta_m$ is the a priori accuracy.

If (11) is satisfied, then the $V_m$ coordinate changes according to the formula:

$$V_{m}^{i+1} = V_m^{i} + \{\text{sign}(\xi - F_m(V_m|V_1, ..., V_{m-1})) \cdot V_{\text{max}}/2 \}.$$  

The superscript for $V_m$ indicates the number of the iteration step in the numerical solution (11). The estimate of the numerator and denominator of equations (10) will be sought by the Parzen method [10]. For the Gaussian kernel and decorrelated coordinates of the vector $V_m$, we have

$$\hat{\omega}_m(V_1, ..., V_m) = \frac{1}{N} 2\pi^{-\frac{m}{2}} h^{-m} \sum_{i=1}^{N} \exp \left[ -\frac{1}{2} h^{-2} \sum_{j=1}^{m} (V_j - V_{ij})^2 \right],$$

(12)
where $h$ is the smoothing parameter, the optimal value of which is calculated in accordance with [13].

Integrating (12) over the last coordinate, we get

$$Y_m = \sum_{i=1}^{N} \prod_{j=1}^{m-1} \eta'\left(\frac{V_{m-i} - V_{m-i-j}}{h}\right) \frac{V_{m-i} - V_{m}}{h} \cdot \sum_{j=1}^{N} \prod_{i=1}^{m-1} \eta'\left(\frac{V_{j} - V_{m-i-j}}{h}\right).$$

(13)

Expression (13) shows that the calculation of the conditional multivariate distribution function is only slightly exceeds the computational complexity of determining the multidimensional density function estimation of the same dimension Parzen method, due to the identity product in brackets for both the numerator and denominator. When solving equation (13) by the dichotomy method at each step of the forecast, only $\eta\left(\frac{V_{i} - V_{m-i}}{h}\right)$ will be the variable value, i.e.

$$\eta'\left(\frac{V_{j} - V_{m-i}}{h}\right) = (2\pi)^{-\frac{m+1}{2}} h^{-m+1} \exp \left\{ -\frac{1}{2} h^{-2} \sum_{j=1}^{m-1} (V_j - V_{m-i})^2 \right\}.$$

(14)

This means changing only the N-dimensional vector of each subsequent step. Consequently, the dichotomy algorithm, and hence forecasting, will not depend on the dimension of each implementation of the training sample.

5. Algorithm for predicting the residual life of complex electromechanical systems

Residual life will be determined by parameters such as the probability of failure-free operation, intensity of failures. Let us apply the reliability model of a complex system in the form:

$$f(t) = \lambda e^{-\lambda t},$$

(15)

where $f(t)$ is the density function of the exponential distribution law, $\lambda$ is the failure rate.

Mean time before failure (MTBF) [1]:

$$T_m = \frac{1}{\lambda}.$$

(16)

Then

$$P(T_m) = e^{-1}.$$

(17)

Let the probability of no-failure operation at the end of the operating time be given as $P_E$, and the MTBF as $t_E$, then

$$t_E = -\frac{1}{\lambda} \ln (P_E).$$

(18)

From (16) we find the failure rate

$$\lambda = \frac{1}{T_m}.$$

(19)

Thus, further software implementation of calculating the estimate of the MTBF of a complex electromechanical system involves the following steps of the algorithm:

1) Creating a training sample of conditionally fit products (space $S_0$) in the form of a matrix $X_0(i,j), i = 1, M_0, j = 1, N_0$.

2) Conversion of the training sample of conditionally fit products into a matrix $V_0(i,j), i = 1, M_0, j = 1, N_0$, consisting of decorrelated row vectors.

3) Creating a training sample of conditionally non-fit products (space $S_1$) in the form of a matrix $X_1(i,j), i = 1, M_1, j = 1, N_1$. 


4) Conversion of the training sample of conditionally non-fit products into a matrix $V_i(i,j), i = 1, M_1, j = 1, N_1$, consisting of decorrelated row vectors.

5) Input control data as a vector $x_k = [x_1, x_2, ... , x_k]$ for some products.

6) Recognition by the sequential Wald algorithm - assignment of the observed vector, if possible for step $k$, to the space $S_i, i = 0, 1$.

7) Decorrelation of the input recognized (i.e. referred to the space $S_0$ or $S_1$) vector by the V. S. Pugachev canonical decomposition method: $(v_k = [v_1, v_2, ... , v_k])$; here and below without designating the hypothesis number.

8) Run sequentially for the vector $(v_k = [v_1, v_2, ... , v_k])$ forecast for the components $[v_{k+1}, v_{k+2}, ... , v_N], N = N_0$ or $N_1$ (total $(N - k)$ times). Perform inverse transform: $(v_{N|k} = [v_{k+1}, v_{k+2}, ... , v_N]) \Rightarrow (x_{N|k} = [x_{k+1}, ... , x_N])$.

9) Repeat steps 7 and 8 $L$ times (a total of $L$ predicted vectors of length $(N - k)$ extensions of the vector $x_k = [x_1, x_2, ... , x_k]$).

10) Calculate the estimate of MTBF $\hat{T}_m$ using the formula $\hat{T}_m = \frac{1}{T} \sum_{i=1}^{L} T_i$, where $T_i = \lambda T$ - time interval of going out of tolerance $[(x_i(i) > (m(x_i) + \Delta x)) \text{ or } (x_i(i) < (m(x_i) - \Delta x))], l = 1, L, i = 1, N$ 1 is forecast function number, $\lambda$ is out-of-tolerance forecast number, $T$ is out-of-tolerance forecast number $x_i, i = 1,2, ... M$.

11) For the case $(m(x_i) + \Delta x) < x_i < (m(x_i) - \Delta x), l = 1, L, i = N$, write the MTBF as $T_i = NT$.

12) The mean time between failures corresponds to an index such that $\hat{T}_m = \frac{i}{\lambda T^n}$.

13) Calculation of the estimate of the failure rate according to the formula (19) $\hat{\lambda} = \frac{1}{\hat{T}_m}$, where $\hat{\lambda}$ is the estimate of the failure rate, $\hat{T}_m$ is the estimate of the mean time between failures.

14) Calculation of the MTBF estimate according to the formula (18): $\hat{t}_E = -\frac{1}{\hat{\lambda}} \ln (P_E)$.

6. Results of computational experiments

For modeling were developed programs of recognition, decorrelation, forecasting using MatLab software. The result of the program for predicting the boundaries of the assumed vibration velocity for the observed vector $x_k$, the first $k$ components of which are determined, for the training sample $X_{NM}$, is shown in Figure 1. Figure 1 shows only the upper threshold $C_1$, since a reduction in vibration is not a sign of bearing failure ($C_0 = 0$).

![Figure 1](image-url)  
Figure 1. Prediction of the development of bearing defects in the drive of a mechatronic device by vibration speed: a) upper bound of the forecast; b) lower bound of the forecast; c) theoretical upper limit of vibration speed at the worst development of defects [14]; d) theoretical lower limit of vibration velocity with favorable development of defects for higher quality bearings [14].
In model experiments the time interval for recording the level of vibration velocity is $T = 24$ hours. The decision to recognize the bearing as failed for each discrete forecast function is made at the stage of intensive development of defects in accordance with [14]. Therefore, the threshold is set to the level of vibration speed - 12 mm/s.

In this model experiment, the MTBF was $t_E = 100T$ with the probability of failure-free operation $P_E = 0.98$ for $\alpha = 0.05$ and $\beta = 0.1$.

7. Conclusion
The algorithm for predicting the residual life of complex electromechanical systems developed in this work allows:

- to form correlated and decorrelated training samples for conditionally fit and conditionally non-fit drives;
- assign the analyzed sample to one of the classes;
- calculate decorrelated input vector;
- calculate the predicted development of a random process;
- calculate the residual life of the drive.

The proposed algorithm makes it easy to change the statistical conditions for carrying out model experiments by changing the training samples, adding successful implementations or removing doubtful ones. The derivation of parameter estimates from small training samples is provided by an improved non-parametric Parzen approach to estimate a probability density function, and the reduction in the number of tests is achieved by the Wald’s sequential decision-making procedure.

Acknowledgements
The reported study was funded by RFBR, project number 18-08-00772 A.

References
[1] Kashtanov, V.A., Medvedev, A.I. (2010) Teoriya nadezhnosti slozhnykh system (Reliability theory of complex systems), 2nd ed., Rev., Moscow, FIZMATLIT. 608 p. (in Russian)
[2] Jardine, A.K.S., Lin, D. and Banjevic, D. (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20, 1483–1510.
[3] Stepanov, P., Lagutkin, S., Nikitin, Y. (2015) Development of the Algorithm for Definition of Residual Service Life by a Comprehensive Diagnosis of the Electromechanical Drive. Acta Mechanica Slovaca. 19(2). Pp. 52-56.
[4] ISO 13381–1. (2004) Condition Monitoring and Diagnostics of Machines Prognostics Part1: General Guidelines, International Organization for Standardization, Geneva.
[5] Zhukovskiy, Yu. L. et al (2017) IOP Conf. Ser.: Earth Environ. Sci. 87 032056
[6] Turygin, Yu., Bozek, P., Nikitin, Yu., Sosnovich, E., Abramov A. (2016) Enhancing the reliability of mobile robots control process via reverse validation. International Journal of Advanced Robotic Systems, November-December. Pp.1-8.
[7] Turygin, Yu., Bozek, P., Abramov, I., Nikitin, Yu. (2018) Reliability Determination and Diagnostics of a Mechatronic System. Advances in Science and Technology. Vol. 12, No. 2, June, pp. 274–290.
[8] Abraham, Wald, (1947) Sequential Analysis, New York, John Wiley & Sons.
[9] Pugachev, V.S. (1960) Teoriya sluchajnyh funkcij i ee primenenie k zadacham avtomaticheskogo upravleniya (Theory of random functions and its application to automatic control problems), 2nd ed., Rev., Moscow, FIZMATLIT. 883 p. (in Russian)
[10] Parzen, E. (1962) On Estimation of a Probability Density Function and Mode. The Annals of Mathematical Statistics. Vol. 33, No. 3, pp. 1065-1076.
[11] Levin B. R. (1975) Teoreticheskie osnovy statisticheskoj radiotekhniki (Theoretical foundations of statistical radio engineering), 2nd ed., Rev., Moscow, Sov. radio. Book two. 392 p. (in Russian)

[12] Trefilov, S.A. (1994) Razrabotka algoritmov dlya modelirovaniya bortovyh informacionno-izmeritel'nyh sistem na etape proektirovaniya: avtoreferat dissertacii na soiskanie stepeni kandidata tekhnicheskih nauk (Development of algorithms for modeling on-board information and measuring systems at the design stage: dissertation abstract for the degree of candidate of technical sciences). Izhevsk, 16 p. (in Russian)

[13] Botev, Z.I., Grotowski, J.F., Kroese, D.P. (2010) Kernel density estimation via diffusion. *Annals of Statistics*. Vol. 38, No. 5, pp. 2916–2957.

[14] Akutin M.V. (2009) Metod kontrolya sostoyaniya podshipnikov kacheniya na osnove sravneniya vejvlet skejlogramm: dissertaciya na soiskanie stepeni kandidata tekhnicheskih nauk (Method for monitoring the condition of rolling bearings based on the comparison of wavelet scaleograms: dissertation for the degree of candidate of technical sciences). Kazan', 132 p. (in Russian)