Experimental modeling of radio information systems for solving the problem of failures forecasting

A Perlov¹, A Zyuzin², A Kokuev³, A Kochkarov⁴, K L’yov⁵ and A Tymoshenko⁶
¹Head of Sector, JSC “RTI”, Moscow, Russia
²Head of Department, Yaroslavl Higher Military School of Air Defense, Yaroslavl, Russia
³Professor, JSC “RTI”, Moscow, Russia
⁴Associate Professor, Financial University under the Government of the Russian Federation, Moscow, Russia
⁵Student, Lomonosov Moscow State University, Moscow, Russia
⁶Deputy General designer, JSC “RTI”, Moscow, Russia
E-mail: laperlov@yandex.ru

Abstract. We report a simulation model of technical condition data of radio information systems at solving the problem of failures forecasting. Time cost analysis on sampling for high-accuracy forecasting of radio information systems failures showed that the performance of samples accumulation by testing the equipment at the stands does not allows us to create radio information systems with automated operation management system on time. The existing level of product unification allows the use of both data from previous generation stations and current data from of the built-in control of new generation radar stations, due to their correlation connections. These features create the conditions for operative formation of a large amount of data on the technical states of the stations due to the development of a simulation data model of the built-in control. A feature of such a model concludes in its versatility, i.e. the formation of the data stream on both binary sensors (operational or not) and sensors of physical parameters. An analysis of the temperature recorded during the development process of the components was carried out to determine the minimum sample size required. The temperature prediction model described above was trained on samples of various volumes for determination of the influence of the training sample volume on forecast accuracy. The developed methodological apparatus for formation of data from the station’s built-in control provides a solution of the problem of high-precision failure prediction.

1. Introduction
Modern radio information systems (RIS) require equipping with automated operation management systems (AOMS), aimed to provide operating personnel with information and logistics support. The performance of such systems is determined by the quality of the product technical condition data, which allows the use of modern methods to solve the problem of forecasting critical product failures.

The problem of failures forecasting can be solved both by classical methods of statistical analysis that carry out the forecast by accumulating and processing statistics, and by machine learning methods that analyze time sequences and solve the classification problem [1, 2]. An analysis of publications of this subject area [3, 4] showed that the use of machine learning methods for organizing maintenance of complex multi-element products is increasingly used in leading military-industrial companies in the
USA and Europe. At the same time, the performance of these methods is largely defined by the formation of a necessary amount data for high-precision forecasting. Along with this circumstance, data accumulation is associated with significant time costs, since the so-called rare failures are most intrinsic for the new generation RIS.

Time cost analysis on sampling for high-accuracy forecasting of RIS failures showed that the performance of samples accumulation by testing the equipment at the stands does not allows us to create RIS with AOMS on time (Figure 1).

![Figure 1](image.png)

Figure 1. The time required to create RIS with AOMS in comparison with the sample size required for high-precision forecast

Thus, the need to create modern automated control systems by the operation of RIS in the directive period reveals the urgency to develop a methodological apparatus for generating data of integrated control RIS. It allows to solve the problem of high-precision failure prediction.

2. Research Method

To obtain the required amount of data we propose is to combine it from the integrated control system of the components of the RIS functional systems technical condition data from the simulation model. A functional system of RIS is a set of software and hardware components of RIS that ensure the performance of the functional characteristics of the product as intended.

The possibility of creating such a simulation model is due to the fact that modern RIS are unified series in which locators of a new generation, as a rule, consist of 35–65% of the previous generation RIS components (Figure 2) [5, 6].

This level of unification allows the use of both data from the previous generation of RIS and temporary one from the built-in control of the new generation of RIS, due to their correlations, which in turn creates conditions for the rapid formation of a large amount of technical condition data of RIS by developing a simulation model of the built-in control data. The feature of such a model concludes in its versatility, i.e. the formation of the data stream on both binary sensors (functional or not) and sensors of physical parameters (temperature, humidity, power).
For effective training of the forecast system this simulation model should have the following features of the functioning of modern RIS:

- the rare nature of failures;
- a large number of recorded variables of the product functional systems;
- the difference in the consequences of failure in the event of various states of RIS inoperability, depending on the level of disaggregation;
- implementation of a variety of backup methods and recovery strategies for product components;
- the possibility of several incompatible types of failures of the RIS elements, leading at a certain frequency and sequence of occurrence to various consequences at the level of the RIS functioning as a whole.

These features require a systematic approach to create a simulation model of data on the technical state of the RIS components [7].

We consider that:

- $X$ feature space is a $d$-dimensional Euclidean space $\mathbb{R}^d$ whose points correspond to the state of the blocks and the system at the current and previous points in time;
- $Y$ answer space is the $q$-dimensional Euclidean space $\mathbb{R}^q$ whose points correspond to the failure state of the system at future time instants;
- the space $F$ of models $f: X \rightarrow Y$;
- the space of distributions $P$ (probability measures) on $X \times Y$;
- error function $E: Y \times Y \rightarrow \mathbb{R}$, non-negative and equal to zero when the first (predicted answer) and second (true answer) parameters coincide. Usually, when the space $Y$ is Euclidean, a quadratic error is chosen $E(r, y) = \|r - y\|^2$;
- training sample $T=((x_1, y_1),..., (x_N, y_N))$, consisting of pairs (feature vector, response) $(x_i, y_i) \in X \times Y$, which are considered values of independent random variables with the same, but completely unknown, distribution $\pi \in P$.

It is required to build a prediction model $f \in F$ that minimizes the mathematical expectation of the error function on the distribution $\pi$ according to the data $X, Y, F, P, E$ and $T$

$$E(f) = \int_{(x,y) \in X \times Y} E(f(x), y)d\pi(x,y) \rightarrow \min_{f \in F}.$$
The next step is to replace the minimization of integral $E_n(f)$ with the empirical risk minimization function:

$$E(f, T) = \frac{1}{N} \sum_{i=1}^{N} E(f(x_i), y_i) \rightarrow \min_{f \in F}.$$ 

It is important to notice that the model obtained as a result of training by minimizing empirical risk depends on the training sample $T$, including its volume $N$.

Thus, the key point of forecasting with the required accuracy is to solve the problem of assessing the dependence of empirical risk $E(f, T(N))$ on the volume of the training sample $N$.

3. Results and Discussion

One can construct the simulation model of technical condition data by implementing the following algorithm:

- the parameters of the failure flow simulator (FS) are determined by priori data on the reliability characteristics of the unified components of the previous generation RIS; it takes into account the structural and technological features of the FS and others, information about which is reliably known at the current state of the development life cycle;

- dependences of FS characteristics from the technical state are determined by the results of the analysis of data on failures of FS and its components and are formed the requirements to the composition and structure of internal control to all key components of FS that allows to integrate data from sensors into a stream of denials from the reference simulator;

- it is determined the minimum necessary volume of the training sample, sufficient for high-precision forecasting of the technical condition of the FS by assessing the statistical stability of the required forecast accuracy by calculating confidence intervals for a different number of sample elements;

- the parameters of the simulation model are iteratively adjusted until the accuracy of the forecast of the real data is equal to the required one.

The scheme of the simulation model that provides the implementation of the above algorithm for high-precision prediction of RIS failures is shown in Figure 3.

The analysis of temperature values recorded during the processing of components (power amplification units) of the RIS transmission system was carried out to determine the minimum required sample size.

Figure 4 shows the results of estimates of the system failure probability at every moment in time.

It is used a metric – coefficient of determination ($R^2$-score) to analyze the performance of the model and the selection of its parameters; it allows us to determine how accurately the predicted temperature curve repeats the true curve:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2},$$

where $y_i$ is the true temperature value; $\hat{y}_i$ is the predicted temperature value; $\bar{y}_i = \frac{1}{n} \sum_{i=1}^{n} y_i$ is the average temperature value.
The value of the determination coefficient for predicted probabilities ($R^2$-score) for different durations of the forecast interval was calculated in order to determine how the forecast error changes with the duration of the forecast interval. This dependence has a downtrend reaching the level of 0.76.

Figure 3. The scheme of simulation model of technical condition data

Figure 4. The probability of failure one count ahead

Sharp jumps in the coefficient of determination occur due to the mismatch of predicted and true positions in which the system changes its state (Figure 5).

The temperature prediction model described above was trained on samples of different volumes to determine the influence of the training sample size on the accuracy of the forecast. After that, the model was tested, including the temperature forecast was repeatedly carried out, and the coefficient $R^2$ was calculated from the obtained values.

Figure 6 shows the dependence of the prediction accuracy of the individual component of RIS ($R^2$-score) on the volume of the training sample.
The analysis of changes in forecast accuracy showed that the achievement of statistical stability of the required forecast accuracy \( \sim 0.995 \) occurs with an increase in the training sample to \( 10^4 \) samples.

4. Conclusion

We have shown that it is fundamentally important to solve the problem of generation of technical condition data, the processing of which will provide the specified forecast accuracy for the timely and high-quality completion of the development of RIS equipped with an automated operation control system. The methodological apparatus developed for generation of data from the integrated control RIS provides a solution of high-precision failure prediction problem. The approach proposed in the article generally corresponds to the modern development directions of the leading military-industrial companies in the USA and Europe [6].

References

[1] Shmueli G and Koppius O R 2011 *Mis Quarterly* 553–572
[2] Wang Z et al 2017 *Optics Express* 25(16) 18553–18565
[3] Technology Today 2018 Retrieved from https://www.raytheon.com/sites/default/files/technology-today/2018/issue1/wp-content/uploads/2018/08/Raytheon_TechnologyToday_Issue1_2018.pdf
[4] Sun W G et al 2012 *Advanced Materials Research* 472–475 2650–2654
[5] Freund Y and Mason L 1999 *Proc of ICML '99* 124–133
[6] Boev S F et al 2018 *IOP Conf. Ser.: Mater. Sci. Eng.* 457 012010 doi: 10.1088/1757-899X/457/1/012010
[7] Boev S et al 2018 *IOP Conf. Ser.: Mater. Sci. Eng.* 457 012003 doi: 10.1088/1757-899X/457/1/0120035