Optimizing thresholds of classification system for identifying status of transport communication systems

G I Linets, S N Sandryan, S V Melnikov

North-Caucasus Federal University, Institute of Information Technologies and Telecommunications, 2, Lenina Ave., Stavropol, 355029, Russia

E-mail: territorereer@yandex.ru

Abstract. This work proposes an approach to implementing the threshold optimization for classifying the states of complex technical systems based on the application of the normal law of random variables distribution.

1. Introduction

The problem of the efficient use of network resources and the requirement for implementing a distributed information development in telecommunication systems (TS) have brought up the issue of creating and implementing new methods of information and computing resources management in the modern communication networks allowing one to improve the efficiency of data exchange and to significantly reduce the amount of control information circulating through a network without reducing the operation quality of the communication network as a whole.

In terms of the implementation of control mechanisms, a communication network is viewed as a complex interacting system with dispersed adaptive control, based on the coordination of individual components under the influence of fluctuating processes, and no information on the system status is available [3, 6].

When researching the status of communication systems, one of the most pressing problems is the identification of their status. The correct solution to this problem significantly influences the quality of information exchange as well as the processing efficiency of the information transmitted. At the same time, the procedure for determining thresholds for classification of communication system states is essential. It is also noteworthy that this procedure is applicable for both wired and wireless transmission systems. For example, it is known [5] that to assess various conditions of satellite channels, the following laws are applied as change models for random ionospheric parameters: normal, Rayleigh, Rice, Nakagami. They characterize the parameters change in the transmitted radio signals for various ionosphere conditions. However, given there is the complexity of computational procedures for determining the optimal classification thresholds, the task of their optimization for the normal law has not been resolved [3, 6]. It is also noteworthy that that the normal law is a limiting law to which other laws of random variables distributions are approaching under very frequent typical conditions [4]. This work proposes an approach to implementing the threshold optimization for classifying the states of complex technical systems based on the application of the normal law of random variables distribution.
2. Current problem analysis
The urgency of developing a system for identifying the states of traffic info-communication networks is determined by the complexity of the tasks to assess their current status. The main difficulty arising at the design stage is to ensure sustainability and required quality of their condition assessment in all operation modes. This is due to the fact that modern transport information and communication systems are operating under conditions of great uncertainty [3, 6]. The information on the real values of some of its parameters may be inaccurate, and the laws of their possible changes are often known only at the qualitative level. In some cases, external influence and some parameters are to be described statistically. Given there is a large degree of uncertainty, the synthesis procedure for a control system should provide for the simultaneous solution of the problems of creating an effective control and a rational (in some sense, optimal) management algorithm which would use both a priori and a posteriori information obtained in the course of its operation.

A transport network management system consists of two main components: a management object and a control system. The control system performs the following functions [7]: the identification of the transport network status; compiling control inputs based on the management objectives, taking into account the transport network status and the environment; providing control action for the transport network. The status identification of is closely linked to the concept of pattern recognition. Therefore, the possibility of applying a pattern recognition system at the stage of its condition identification seems to be quite obvious and natural. [7]

In order to improve the quality operation and to achieve maximum performance, an automatic condition monitoring system should be adaptive. In this case, the term "adaptive system" implies the approach described in [1], when the lack of a priori information is compensated by a better (as compared to the non-adaptive system) use of the current information. That is, in order to reduce the uncertainty and to achieve the specified quality parameters of information exchange and management processes, purposeful forced change of the parameters and structure of the control device is performed based on the current (a posteriori) information. The process of purposeful change of the parameters, properties or structure of a transport network based on the information obtained in the course of its performance of core functions, in order to achieve its optimal status under the changing circumstances, given there is initial uncertainty, is herein described as “adaptation”.

Traditional network management systems are essentially parametric control systems, i.e. they do not control the transport network states, but manage its observable parameters. In [7] it is proven that the efficient management of complex systems requires control of the status and not the settings.

Given there is heterogeneity of modern information and communication networks, the task of determining the classification thresholds for adaptive management is crucial for wired and, particularly, wireless systems of information transmission. This is due to the fact that the basic characteristics of any radio communication systems and, in particular, radio systems are defined by the characteristics of their channels. Therefore, the characteristics of radio systems can be decreased artificially (e.g., ionospheric disturbances may lead to significant degradation of detection, measurement, and discrimination of signals). At the same time, various fading, signal amplitude fluctuations and similar signal distortions are described by the laws of random variables distribution.

The existing identification methods do not imply qualitative classification at the class intersection. If a system has multi-type attributes, not all methods are suitable for its classification [2]. The most commonly used solution is the artificial division of classes, for example, introduction of additional space for the classification using a support vector method in the case of the intersection of classes in a training sample. However, given there is a large number of characteristics, quite powerful computing capacities are required, and the introduction of additional characteristics exacerbates further this problem.

In [3, 6], the classification methods used in the design and development of advanced automatic control and adaptive control systems present solutions to establish the optimal classification thresholds only for certain laws of random variables distribution, exponential and Rayleigh law.
Hence, it is important to develop an approach ensuring the required accuracy for the optimization of the classification thresholds for random variables with the normal distribution law.

3. Problem statement
The authors applied the selection approach for optimal classification thresholds as stipulated in [3, 6]. It is essential to minimize the risk of Type I and Type II errors.

To achieve this, an object recognition task, in general, is formally reduced to the verification of numerous hypotheses

\[ B_1, B_2, \ldots, B_i, \ldots, B_k \]

where \( B_i \) is the hypothesis assuming the object belongs to class \( A_i \). Let us assume that the a priori probability distribution of these hypotheses is defined, i.e. probability \( P(B_i) \) of the object’s belonging to class \( A_i \) is known. Moreover, \( \sum_{i=1}^{k} P(B_i) = 1 \), as an object ought to belong to a specific class. Under this condition, the distribution density is as follows:

\[ p_i(x) = p(x|B_i) \]

The identification system used two hypotheses \( B_i = \mathcal{N} \) and \( B_i = \overline{\mathcal{N}} \) with their respective a priori probabilities of the occurrence of the normal \( p_1 = p(B_1) = p(\mathcal{N}) \) and the abnormal \( p_2 = p(B_2) = p(\overline{\mathcal{N}}) \) situations in the system. Here, \( p_1 + p_2 = 1 \).

As a decision rule ensuring the top accuracy of the identification system, let us apply the Neyman-Pearson criterion. By fixing false alarm probability \( P_{f,a} \) at constant level \( C \), minimum omission error \( P_{om}^{\text{min}} \) for disturbing the system operation is required. Then,

\[
P_{om}^{\text{min}} = \min p_2 \beta(x_0)
\]

with a limit of

\[
P_{f,a} = p_1 \alpha(x_0) = C = \text{const}
\]

where \( \alpha(x_0) \) is Type 1 errors; \( \beta(x_0) \) is Type 2 errors.

Here, the probability density of \( x \) and \( y \) parameter distribution as normal distribution laws are shown in fig. 1a and fig. 1b.

**Figure 1.** Density of \( x \) and \( y \) parameter distribution.

Let us determine Type 1 errors (“false alarm” errors, fig. 1a) and Type 2 errors (“abnormal situation” error, fig. 1b):

\[
\alpha(x_0) = \int_{x_0}^{\infty} f(x|\mathcal{N}) \, dx
\]

\[
\beta(x_0) = \int_{-\infty}^{x_0} f(x|\overline{\mathcal{N}}) \, dx
\]
The function of distribution density for parameter \( x \) at the proper functioning of the system is
\( f(x/N) = f(x) \) and at malfunctioning - \( f(x/N) = f_2(x) \). Then the Type I and Type II errors of the detector
(Stage 1) will be, respectively:
\[
\alpha_{det} = \int_{x_0}^{x} f_1(x)dx; \quad \beta_{det} = \int_{-\infty}^{x_0} f_2(x)dx
\]  
(5)

The same method is applied to determine Type I and Type II errors of the recognizer (Stage 2)
\[
\alpha_{rec} = \int_{y_0}^{y} f_1(y)dy; \quad \beta_{rec} = \int_{-\infty}^{y_0} f_2(y)dy.
\]  
(6)

In view of formulas (5) and (6), formulas (1) and (2) look as follows:
\[
P_{f,a} = p_1 \int_{x_0}^{x} f_1(x)dx \int_{y_0}^{y} f_1(y)dy = C = const; \quad (7)
\]  
(7)

As in formula (7), thresholds \( x_0 \) and \( y_0 \) are linked in functional relationship \( x_0 = \phi(y_0) \), by
differentiating (8) at \( y_0 \) and bringing it to zero, the following is received:
\[
\frac{dx_0}{dy_0} : f_2(x_0) \cdot \int_{y_0}^{y} f_1(y)dy + f_2(y_0) \cdot \int_{x_0}^{x} f_1(x)dx = 0 . \]  
(9)

The general formulation of the objective to optimize classification thresholds \((x_0, y_0)\) comes down to the
solution of simultaneous equations:
\[
\begin{align*}
\int_{x_0}^{x} f_1(x)dx \cdot \int_{y_0}^{y} f_1(y)dy &= C / p_1 \\
\frac{dx_0}{dy_0} : f_2(x_0) \cdot \int_{y_0}^{y} f_1(y)dy + f_2(y_0) \cdot \int_{x_0}^{x} f_1(x)dx &= 0
\end{align*}
\]  
(10)

with limitations: \( p_1 = const \) at \( p_1 + p_2 = 1 \). The task is to solve simultaneous equations (10) and to
establish optimal classification thresholds \( x_0 \) and \( y_0 \).

4. Problem solution
In order to solve simultaneous equations (10), the authors used the formula of normal distribution as presented
below [4,8]: \( \phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du \).

Its relator looks as follows: \( \phi(x) = \frac{1}{2} - \phi(0) \)  
(11)

where \( \phi(A) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{A} e^{-u^2/2} du \) \quad La Place formula.

For this formula, the authors determined the following relations:
\[
\int_{-a}^{a} f_2(u)du = \int_{-a}^{a} f_1(u)du = \phi(x-a); \quad \int_{-a}^{a} f_2(u)du = \phi(x+a);
\]  
(12)

\[
(\phi(x))' = -\frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}
\]  
(13)

For further solution, based on (12), let us put simultaneous equations (10) in a convenient format:
\[
\begin{align*}
\phi(x_0) \cdot \phi(y_0) &= C / p_1 \\
\phi(x_0-a) \cdot \phi(y_0-b) &\rightarrow \max
\end{align*}
\]  
(14)
By applying the known formula: 

\[
\frac{x+y}{2} = \sqrt{xy}, \text{ where } x > 0, y > 0 \text{ and } x + y = \text{const}, \text{ and } \max(x, y)
\]
is possible at \( x = y \). Then simultaneous equations (14) look as follows:

\[
\phi(x_0) \cdot \phi(y_0) = C^*; \tag{15}
\]
\[
x_0 - a = y_0 - b \tag{16}
\]
which gives us:

\[
\Delta = a - b; \quad x_0 = y_0 + \Delta; \tag{17}
\]
\[
\phi(y_0 + \Delta) \cdot \phi(y_0) = C^* = \text{const} \tag{18}
\]

Let us now consider the approximate solution of equations (10) applying Newton’s method for the systems of nonlinear equations.

By deriving equation (11) at \( x \), one receives:

\[
\phi'(x) = \left( \frac{1}{2} - \phi_0(x) \right) = -\frac{1}{\sqrt{2\pi}} \int_0^{\frac{\pi}{2}} e^{-\frac{u^2}{2}} \, du = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{u^2}{2}} \, u. \tag{19}
\]

According to formulas (11), (12) and (13), simultaneous equations (10) look as follows:

\[
\begin{cases}
\phi(x_0) \cdot \phi(y_0) = C / p_1, \\
\frac{dx_0}{dy_0} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x_0-a)^2}{2}} \cdot \phi(y_0 - b) + \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(y_0-b)^2}{2}} \cdot \phi(x_0 - a) = 0.
\end{cases}
\tag{19}
\]

By differentiating the first equation in system (14) at \( y_0 \), one receives:

\[
\frac{d\phi(x_0)}{dy_0} \bigg|_{y_0} = \frac{d\phi(y_0)}{dy_0} \bigg|_{y_0} \text{ OR } \left( \frac{d\phi(x_0)}{dy_0} \right) \bigg|_{y_0} \cdot \phi(y_0) + \phi(x_0) \cdot \phi'(y_0) \bigg|_{y_0} = 0. \tag{20}
\]

Given that \( x_0 = x_0 \bigg|_{y_0} \), let us receive \( \left( \frac{d\phi(x_0)}{dy_0} \right) \bigg|_{y_0} \cdot \phi(y_0) + \phi(x_0) \cdot \phi(y_0) \bigg|_{y_0} = 0. \tag{20}
\]

Based on equation (20), there is:

\[
\frac{dx_0}{dy_0} = \frac{\phi(x_0)}{\phi(y_0)}. \tag{21}
\]

By plugging (21) in simultaneous equations (19), let us transform the second equation as follows:

\[
e^{-\frac{x_0^2}{2}} \cdot \phi(y_0) \cdot e^{-\frac{y_0^2}{2}} \cdot \phi(x_0) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x_0-a)^2}{2}} \cdot \phi(y_0 - b) + \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(y_0-b)^2}{2}} \cdot \phi(x_0 - a) = 0,
\]
or

\[
e^{\frac{x_0^2}{2}} \cdot e^{-\frac{(x_0-a)^2}{2}} \cdot \phi(x_0) \cdot \phi(y_0 - b) + e^{\frac{y_0^2}{2}} \cdot e^{-\frac{(y_0-b)^2}{2}} \cdot \phi(y_0) \cdot \phi(x_0 - a) = 0.
\]

Given that \( \frac{x_0 - (x_0 - a)^2}{2} = x_0 \cdot a - \frac{a^2}{2} \text{ and } \frac{y_0 - (y_0 - b)^2}{2} = y_0 \cdot b - \frac{b^2}{2}, \)

the latter now looks as follows:

\[
-\frac{e^{-\frac{(x_0-a)^2}{2}} \cdot \phi(x_0) \cdot \phi(y_0 - b) + e^{\frac{y_0^2}{2}} \cdot e^{-\frac{(y_0-b)^2}{2}} \cdot \phi(y_0) \cdot \phi(x_0 - a)}{\phi(x_0) \cdot \phi(y_0 - b) + e^{\frac{y_0^2}{2}} \cdot e^{-\frac{(y_0-b)^2}{2}} \cdot \phi(y_0) \cdot \phi(x_0 - a)} = 0.
\]

Then simultaneous equations (19) is transformed to look as follows:

\[
\left\{ \begin{array}{l}
\phi(x_0) \cdot \phi(y_0) = C / p_1, \\
\phi(x_0) \cdot \phi(y_0 - b) + e^{\frac{y_0^2}{2}} \cdot e^{-\frac{(y_0-b)^2}{2}} \cdot \phi(y_0) \cdot \phi(x_0 - a) = 0.
\end{array} \right. \tag{22}
\]

By introducing the values:

\[
F_x(x_0, y_0) = \phi(x_0) \cdot \phi(y_0) - C / p_1 = 0,
\]
\[
F_y(x_0, y_0) = e^{\frac{x_0^2}{2}} \cdot e^{-\frac{(x_0-a)^2}{2}} \cdot \phi(x_0) \cdot \phi(y_0 - b) + e^{\frac{y_0^2}{2}} \cdot e^{-\frac{(y_0-b)^2}{2}} \cdot \phi(y_0) \cdot \phi(x_0 - a) = 0.
\]

It is necessary to seek a solution to equation [8]: \( \rightarrow X_{(n+1)} = X_{(n)} - W^{-1}(X_{(n)}) \cdot \tilde{F}(X_{(n)}) \), \( \tag{23.1} \)
where
\[
\bar{X}_{(n)} = \left( \frac{x_{n}^{0}}{y_{n}^{0}} \right), \quad \bar{F}(\bar{X}_{(n)}) = \left( \frac{F_{i}(x_{n}^{0}, y_{n}^{0})}{F_{j}(x_{n}^{0}, y_{n}^{0})} \right), \quad W(\bar{X}_{(n)}, \bar{Y}_{(n)}) = \begin{pmatrix} \frac{\partial F_{1}}{\partial x_{n}} & \frac{\partial F_{1}}{\partial y_{n}} \\ \frac{\partial F_{2}}{\partial x_{n}} & \frac{\partial F_{2}}{\partial y_{n}} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.
\]

Then, \( W^{-1}(x_{(n)}^{0}, y_{(n)}^{0}) = \frac{1}{W(\bar{X}_{(n)}, \bar{Y}_{(n)})} \begin{pmatrix} D & -B \\ -C & A \end{pmatrix} \),

where
\[
A = \frac{\partial F_{1}}{\partial x_{n}}, \quad B = \frac{\partial F_{1}}{\partial y_{n}}, \quad C = \frac{\partial F_{2}}{\partial x_{n}}, \quad D = \frac{\partial F_{2}}{\partial y_{n}}.
\]

The differentiation results of the Jacobian's determinant [8]:
\[
\begin{align*}
\frac{\partial F_{1}}{\partial x_{n}} &= -\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x_{n}^{2}}{2}} \cdot \left( \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \int_{0}^{y_{n}} e^{-\frac{u^{2}}{2}} du \right); \\
\frac{\partial F_{1}}{\partial y_{n}} &= -\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y_{n}^{2}}{2}} \cdot \left( \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \int_{0}^{x_{n}} e^{-\frac{v^{2}}{2}} dv \right); \\
\frac{\partial F_{2}}{\partial x_{n}} &= \phi(x_{n}) \cdot \phi(y_{n}) - \phi(x_{n}) - \phi(y_{n}); \\
\frac{\partial F_{2}}{\partial y_{n}} &= \phi(x_{n}) \cdot \phi(y_{n}) - \phi(x_{n}) - \phi(y_{n});
\end{align*}
\]

Therefore, the Jacobian's determinant in point \((x_{n}, y_{n})\) looks as follows:
\[
W(\bar{X}_{(n)}, \bar{Y}_{(n)}) = \begin{pmatrix} \frac{\partial F_{1}}{\partial x_{n}} & \frac{\partial F_{1}}{\partial y_{n}} \\ \frac{\partial F_{2}}{\partial x_{n}} & \frac{\partial F_{2}}{\partial y_{n}} \end{pmatrix}, \quad \text{where} \quad \phi(x) = \frac{1}{2} - \phi(x_{n}) = \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{t^{2}}{2}} dt.
\]

The matrix inverse to Jacobian's matrix is calculated as follows [8]:
\[
W^{-1}(x_{n}, y_{n}) = \frac{1}{W(\bar{X}_{(n)}, \bar{Y}_{(n)})} \begin{pmatrix} D & -B \\ -C & A \end{pmatrix}, \quad W(\bar{X}_{(n)}, \bar{Y}_{(n)}) = AD - BC \quad (24.3)
\]

5. Example
In order to establish the optimal classification thresholds \((x_{c}, y_{c})\) for the normal distribution of characteristics \(x\) and \(y\) for object detectors based on parameter \(x\) and for recognizers based on parameter \(y\), let us define the input parameter values as follows: \(a = 2\), \(b = 2.5\), \(c = C / \rho_{1} = 0.005\).

Based on (23, 24.1-24.3), the authors established the classification thresholds applying the Newton's method. The initial approach point is \(x = 1.4, y = 2.1\). Using formula (24.1-24.3), let us iterate until the accuracy of found equation (10) solution satisfies the given limitations. The calculation results are presented in Table 1. By applying the Newton's method, the solution was obtained after 66 iterations.

To validate the results, the authors calculated optimal classification thresholds \((x_{0}, y_{0})\) with the same input data applying the method of simple iterations. By applying the method of simple iterations, a solution was obtained after 37 iterations.

Hence, the desired equation solutions were found by applying Newton's method \(x_{0} = 1.57325\), \(y_{0} = 2.09999\) with the required accuracy. The verification was performed by applying the method of simple...
The solutions deviation amounted to about 0.01%, which is an acceptable value for engineering systems.

6. Conclusions
1. The problem of optimization of control errors and determining the optimal classification thresholds for general cases is resolved and the authors demonstrate an example of its application for a two-stage procedure of identifying abnormal situations under the normal distribution laws for parameter recognition.
2. The authors have obtained the analytical dependence of optimal thresholds of characteristics classification for detection $x_0$ and recognition $y_0$ using the two-stage procedure of identifying abnormal situations with the required accuracy. The value of these thresholds depends on the given probability of the “false alarm” and information capacity of the detection (a) and recognition (b) parameters, as well as the a priori occurrence probability of an abnormal situation in the TS.
3. The accuracy of the findings is confirmed by the results obtained by applying various solution methods (Newton’s and chords) for the optimization problem. In the future, it would be advisable to research the dependencies of the changes in Type 1 and Type 2 errors of the average value of the standard deviation of the normal law of random variables distribution, as well as to assess the benefit from the reduced volumes of the controlling information depending on the information content of the recognition features.

References
[1] Chulin H A 1981 About some sufficient stability conditions of multidimensional SoAC Works of MVTU 360 31-38
[2] Zaharakis I D, Pintelas P E, Kotsiantis S B 2006 Machine learning: a review of classification and combining techniques Artificial Intelligence Review, November 26, 3 159–190
[3] Linets G I 2014 Methods for structure-parametric synthesis, identification, and management of transport telecommunication networks in order to achieve maximum performance output (monograph). (Stavropol: Publishing-information centre «FABULA»).
[4] Taboga M 2012 Lectures on Probability Theory and Mathematical Statistics - 2nd Edition (CreateSpace Independent Publishing Platform)
[5] Marvin K Simon, Mohamed-Slim Alouini 2004 Digital Communication over Fading Channels 2nd Edition (Wiley-IEEE Press; 2 edition)
[6] Sandryan S N, Linets G I 2014 Optimization of control errors of the identification system for monitoring transport telecommunication network status. (Information and communication technologies in science, business and education (INFOCOM-6). Edited volume of the Sixth International Scientific and Technical Conference). pp. 88-91
[7] Simankov V, Lutsenko E 1999 Adaptive management of complex systems based on image recognition theory. (Krasnodar: Technical University of Kuban State Technological University).
[8] Werner C. Rheinboldt 1998 Methods for Solving Systems of Nonlinear Equations (USA, Society for Industrial and Applied Mathematics Philadelphia)