RUSSIAN SPACE SYSTEMS

Scientific and technical journal

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS

Volume 6. Issue 2. 2019

Scientific and technical journal

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS

Vol. 6. No. 2. 2019

Founder:
Joint Stock Company “Russian Space Systems”

The publication frequency is four issues per year.
The journal is included into the Russian Science Citation Index.
The journal is included into the List of peer-reviewed scientific publications approved by the Higher Attestation Commission (VAK RF).
The opinions expressed by authors of the papers do not necessarily those of the editors.
ISSN 2409-0239
DOI 10.30894/issn2409-0239.2019.6.2
The subscription number of the journal in the united catalogue “The Russian Press” is 94086.

Joint Stock Company “Russian Space Systems”,
ul. Aviamotornaya 53, Moscow, 111250 Russia
Tel. +7 (495) 673-96-29
www.russianspacesystems.ru, www.spacedevice.ru
e-mail: journal@spacecorp.ru

Moscow
FIZMATLIT ®
2019

Contents

Vol. 6, Iss. 2, 2019

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 2–13

The Determination of Satellite Clock Corrections for Precise Point Positioning with CDMA GNSS Signals

A.A. Povalyaev, *Dr. Sci. (Engineering), Prof., povalyaev_aa@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
Moscow Aviation Institute (National Research University), Moscow, Russian Federation
A.N. Podkorytov, *Cand. Sci. (Engineering), povalyaev_aa@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
Moscow Aviation Institute (National Research University), Moscow, Russian Federation
S.A. Nikitin, *povalyaev_aa@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
D.V. Filimonova, *filimonova.dv@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation

Abstract. This paper deals with the algebraic principles of Precise Point Positioning with CDMA GNSS signals. It covers specifically the case of a network solution that is the determination of precise satellite corrections by joint processing of measurements generated by a network of ground stations. These satellite corrections are further delivered to user receivers via communication channels and are used to compute user coordinates with errors, usually not exceeding 1 cm. The determination of these corrections requires high precision and is carried out by processing phase measurements with resolved integer ambiguities. Resolution of phase ambiguities considerably improves the accuracy of estimated corrections and, at the same time, sharply reduces the time required to achieve centimeter-level positioning accuracy in the user solution. Algebraic principles of the user solution with integer ambiguity resolution of phase measurements with GNSS CDMA signals were studied in the previously published paper of the authors [1].

Keywords: precise point positioning (PPP), ambiguity resolution (AR), Float PPP, Integer PPP

Introduction

For the methodology and notation as well as key concepts and terminology used in this paper, the readers are referred to the previous publication of the authors [1]. In fact, in both cases of user and network solutions, the main challenge of measurement processing is in dealing with the rank deficiency of linear systems obtained by linearizing nonlinear mathematical models of code and phase measurements preserving integer nature of phase ambiguities.

The integer nature of phase ambiguities allows one to apply ambiguity resolution. This entails considerable increase in the accuracy of estimated corrections and, therefore, dramatic reduction in time required to achieve centimeter-level accuracy of user's positional solution.

Mathematical models of code and phase measurements in Integer Precise Point Positioning (PPP) at ionosphere-free GNSS frequencies with CDMA

Mathematical models of the code $\rho^j_{ifr.i}$ and the phase $L^j_{ifr.i}$ measurements and of the Melbourne–Wübbena combination mw^j_i in Integer PPP algorithms at ionosphere-free GNSS frequencies result in a similar efficiency of ambiguity resolution as with the models based on initial frequencies. Such models were already introduced in the previous paper of the authors [1] and read as follows:

$$\begin{aligned}
&\rho^j_{ifr.i} = R^j_i + w^j_i \Delta D_i + dT_{\rho.ifr.i} - dt^j_{\rho.ifr.i} + \xi^j_{\rho.ifr.i}\\
&L^j_{ifr.i} = R^j_i + w^j_i \Delta D_i + dT_{L.ifr.i} - dt^j_{L.ifr.i} -\\
&-\lambda_{\Delta nifr} N1^j - \lambda_{n_2 ifr} N^j_{mw} + \xi^j_{L.ifr.i}\\
&mw^j_i = b_{mw} - b^j_{mw} - \lambda_{mw} N^j_{mw} + \xi^j_{mw.i}\\
&j = \overline{1, J_i}
\end{aligned} \qquad (1)$$

The notations are identical to [1].

The network solution

In the description of a network solution the following notations are used: J_i is the number of space vehicles (SV) in CDMA GNSS which are simultaneously visible by the entire network of ground stations at the i-th epoch; M is the number of ground network stations; $J_{m.i}$ is the number of SVs visible by the m-th ground station $m = \overline{1, M}$ at the i-th epoch. Table 1 is taken from [2] and is somewhat reduced for brevity. This table gives an example of satellite visibility by a network of GPS ground stations; further, it is referred to as a 'scenario matrix'.

In Table 1 "1" corresponds to visible SVs, while "0" – to invisible SVs. The meaning of indices beside the units and shadowing of several units will become clear further. For the scenario given in Table 1, at the i-th epoch M=7, J_i=12, $J_{m.i}$ equals the number of units in the scenario matrix lines, i.e., to the number of SVs visible by each m-th ground station at the i-th epoch. $J_{\Sigma.k} = \sum_{m=1}^{M} J_{m.k}$ (note that for the scenario of Table 1, $J_{\Sigma.i}$ = 32).

The purpose of a network solution is to estimate the ionosphere-free code clock offsets $dt^j_{\rho.ifr.i}$, phase clock offsets $dt^j_{L.ifr.i}$ and hardware biases b^j_{mw} of the Melbourne–Wübbena combination ($j = \overline{1, J_{m.i}}$) for the whole set J_i of SVs visible by all ground stations in the network using the measurements of all the M ground stations comprising the network. In this case, the coordinates of stations and J_i observed satellites are assumed to be known with high accuracy. Subject to this, the system of nonlinear equations (1) of ionosphere-free measurements of the m-th station of the network ($m = \overline{1, M}$) can be represented in the network solution in the following linearized form (instead of the lower ifr index, the index m indicating the station number is applied):

$$\begin{aligned}
&\Delta\rho^j_{m.i} = w^j_{m.i} \Delta D_{m.i} + dT_{\rho.m.i} - dt^j_{\rho.i} + \xi^j_{\rho.m.i}\\
&\Delta L^j_{m.i} = w^j_{m.i} \Delta D_{m.i} + dT_{L.m.i} - dt^j_{L.i} -\\
&-\lambda_{\Delta nifr} N1^j_m - \lambda_{n_2 ifr} N^j_{mw.m} + \xi^j_{L.m.i}\\
&\Delta mw^j_{m.i} = b_{mw.m} - b^j_{mw} - \lambda_{mw} N^j_{mw.m} + \xi^j_{mw.m.i}\\
&m = \overline{1, M}, \; j = \overline{1, J_{m.i}}
\end{aligned} \qquad (2)$$

where $\Delta\rho^j_{m.i} = \rho^j_{m.i} - R^j_{m.i}$, $\Delta L^j_{m.i} = L^j_{m.i} - R^j_{m.i}$, $\Delta mw^j_{m.i} = mw^j_{m.i}$ are the residuals of ionosphere-free combinations of the code $\rho^j_{m.i}$ and the phase $L^j_{m.i}$ as well as the Melbourne-Wübbena combinations $mw^j_{m.i}$ for measurements of the m-th station; $R^j_{m.i}$ is the geometric distance between the m-th station $m = \overline{1, M}$ and j-th SV $j = \overline{1, J_{m.i}}$.

Table 1. SV visibility by seven network stations: 1 – SV is visible, 0 – SV is invisible

	m	PRN numbers of GPS SVs visible by the entire network of ground stations												$J_{m.i}$
		1	2	3	10	16	17	21	22	23	26	27	31	
		Numbers of SVs in their positions in the united array												
		1	2	3	4	5	6	7	8	9	10	11	12	
Ground network stations	1	0	0	$1_{1,1}$	0	0	1_{2}	1_{3}	0	1_{4}	$1_{5,2}$	0	0	5
	2	$1_{6,3}$	0	1_{7}	0	0	$1_{8,4}$	$1_{9,5}$	0	1_{10}	0	0	0	5
	3	1_{11}	0	1_{12}	0	0	1_{13}	$1_{14,6}$	0	$1_{15,7}$	0	0	0	5
	4	0	0	0	0	$1_{16,8}$	0	0	$1_{17,9}$	0	0	0	$1_{18,10}$	3
	5	0	$1_{19,11}$	0	$1_{20,12}$	0	0	0	0	0	$1_{21,13}$	1_{22}	0	4
	6	0	0	1_{23}	0	0	1_{24}	$1_{25,14}$	0	1_{26}	$1_{27,15}$	$1_{28,16}$	$1_{29,17}$	7
	7	0	1_{30}	0	1_{31}	$1_{32,18}$	0	0	0	0	0	0	0	3

The system of linearized equations (2) for the network solution can be rewritten with the use of the matrix notations:

$$\underset{3J_{\Sigma.i}\times 1}{\boldsymbol{Y}_{net.i}} = \underset{3J_{\Sigma.i}\times nx}{\boldsymbol{H}_{net.i}} \cdot \underset{nx_i\times 1}{\boldsymbol{x}_{net.i}} + \underset{3J_{\Sigma.i}\times 1}{\boldsymbol{\Xi}_{net.i}} \tag{3}$$

where $\underset{3J_{\Sigma.i}\times 1}{\boldsymbol{Y}_{net.i}} = \begin{bmatrix} \underset{1\times J_{\Sigma.i}}{(\boldsymbol{\Delta\rho}_i)^T} & \underset{1\times J_{\Sigma.i}}{(\boldsymbol{\Delta L}_i)^T} & \underset{1\times J_{\Sigma.i}}{(\boldsymbol{\Delta mw}_i)^T} \end{bmatrix}^T$ is the vector of observations, in which $\underset{J_{\Sigma.i}\times 1}{\boldsymbol{\Delta\rho}_i} = \begin{bmatrix} \underset{1\times J_{1.i}}{\boldsymbol{\Delta\rho}_{1.i}^T} & \underset{1\times J_{2.i}}{\boldsymbol{\Delta\rho}_{2.i}^T} & \cdots & \underset{1\times J_{M.i}}{\boldsymbol{\Delta\rho}_{M.i}^T} \end{bmatrix}^T$ is the vector of ionosphere-free code residuals arranged in the order of stations and in order of SVs for each station;

$$\underset{J_{m.i}\times 1}{\boldsymbol{\Delta\rho}_{m.i}} = \begin{bmatrix} \Delta\rho_{m.i}^1 & \Delta\rho_{m.i}^2 & \cdots & \Delta\rho_{m.i}^{J_{m.i}} \end{bmatrix}^T, \qquad m = \overline{1, M}$$

is the vector of ionosphere-free code residuals for the m-th station ordered by SVs; $\Delta\rho_{m.i}^j$, $m = \overline{1, M}$, $j = \overline{1, J_{m.i}}$ is the residual of ionosphere-free code measurement of the j-th SV at the m-th station; $\underset{J_{\Sigma.i}\times 1}{\boldsymbol{\Delta L}_i} = \begin{bmatrix} \underset{1\times J_{1.i}}{\boldsymbol{\Delta L}_{1.i}^T} & \underset{1\times J_{2.i}}{\boldsymbol{\Delta L}_{2.i}^T} & \cdots & \underset{1\times J_{M.i}}{\boldsymbol{\Delta L}_{M.i}^T} \end{bmatrix}^T$ is the vector of ionosphere-free carrier phase residuals arranged in the order of stations and in the order of SVs for each station; $\underset{J_{m.i}\times 1}{\boldsymbol{\Delta L}_{m.i}} = \begin{bmatrix} \Delta L_{m.i}^1 & \Delta L_{m.i}^2 & \cdots & \Delta L_{m.i}^{J_{m.i}} \end{bmatrix}^T$, $m = \overline{1, M}$ is the vector of ionosphere-free carrier phase residuals of the m-th station arranged in the order of SVs; $\Delta L_{m.i}^j$, $m = \overline{1, M}$, $j = \overline{1, J_{m.i}}$ is the residual of ionosphere-free carrier phase measurement of the j-th SV at the m-th station; $\underset{J_{\Sigma.i}\times 1}{\boldsymbol{\Delta mw}_i} = \begin{bmatrix} \underset{1\times J_{1.i}}{\boldsymbol{\Delta mw}_{1.i}^T} & \underset{2\times J_{1.i}}{\boldsymbol{\Delta mw}_{2.i}^T} & \cdots & \underset{1\times J_{M.i}}{\boldsymbol{\Delta mw}_{M.i}^T} \end{bmatrix}^T$ is the vector of residuals of Melbourne-Wübbena combinations arranged in the order of the stations and in the order of SVs for each station. $\underset{J_{m.i}\times 1}{\boldsymbol{\Delta mw}_{m.i}} = \begin{bmatrix} \Delta\, mw_{m.i}^1 & \Delta\, mw_{m.i}^2 & \cdots & \Delta\, mw_{m.i}^{J_{m.i}} \end{bmatrix}^T$, $m = \overline{1, M}$ is the vector of residuals of Melbourne-Wübbena combinations at the m-th station arranged in the order of SVs; $\Delta\, mw_{m.i}^j$, $m = \overline{1, M}$, $j = \overline{1, J_{m.i}}$ is the residual of the Melbourne-Wübbena combination of the j-th SV at the m-th station.

$$\underset{nx_i\times 1}{\boldsymbol{x}_{net.i}} = \Big[\underset{1\times M}{\boldsymbol{\Delta D}_i^T} \quad \underset{1\times M}{\boldsymbol{dT}_{\rho.i}^T} \quad \underset{1\times M}{\boldsymbol{dT}_{L.i}^T} \quad \underset{1\times M}{\boldsymbol{B}_{mw.i}^T}$$
$$\underset{1\times J_i}{\boldsymbol{dt}_{\rho.i}^T} \quad \underset{1\times J_i}{\boldsymbol{dt}_{L.i}^T} \quad \underset{1\times J_i}{\boldsymbol{b}_{mw.i}^T} \quad \underset{1\times J_{\Sigma.i}}{\boldsymbol{N1}_i^T} \quad \underset{1\times J_{\Sigma.i}}{\boldsymbol{N}_{mw.i}^T} \Big]^T \tag{4}$$

is the vector of the estimated variables in the network solution with the dimension nx_i=4M+3J_i+2$J_{\Sigma.i}$, where $\underset{1\times M}{\boldsymbol{\Delta D}_i^T} = \begin{bmatrix} \Delta D_{1.i}^T & \Delta D_{2.i}^T & \cdots & \Delta D_{M.i}^T \end{bmatrix}^T$ is the M-vector of the uncompensated wet component of zenith tropospheric delays (m) at the locations of ground stations; $\underset{1\times M}{\boldsymbol{dT}_{\rho.i}^T}$ is the M-vector of receiver ionosphere-free code clock offsets for all the M stations; $\underset{1\times M}{\boldsymbol{dT}_{L.i}^T}$ is the M-vector of receiver ionosphere-free phase clock offsets for all the M stations of the ground network; $\underset{1\times M}{\boldsymbol{B}_{mw.i}^T}$ is the M-vector of receiver hardware biases of the Melbourne–Wübbena combinations for the M stations of the ground network; $\underset{1\times J_i}{\mathbf{dt}_{\rho.i}^T}$ is the J_i-vector of ionosphere-free code clock offsets for J_i SVs visible by the entire network of ground stations at the i-th epoch; $\underset{1\times J_i}{\mathbf{dt}^T_{L.i}}$ is the J_i-vector of ionosphere-free code clock offsets for J_i SVs visible by the entire network of ground stations at the i-th epoch; $\underset{1\times J_i}{\boldsymbol{b}^T_{mw.i}}$ is the J_i-vector of satellite hardware biases of Melbourne-Wübbena combinations for J_i SVs visible by the entire network of ground stations at the i-th epoch; $\underset{1\times J_{\Sigma.i}}{\boldsymbol{N}1^T_i}$ is the $J_{\Sigma.i}$-vector of integer ambiguities $N1$ included in the system (2) arranged in the order of stations and in the order of SVs for each station at the i-th epoch; $\underset{1\times J_{\Sigma.i}}{\boldsymbol{N}^T_{mw.i}}$ is the $J_{\Sigma.i}$-vector of integer ambiguities N_{mw} included in the system (2) arranged in the order of stations and in the order of SVs for each station at the i-th epoch; $\underset{3J_{\Sigma.i}\times 1}{\boldsymbol{\Xi}_{net.i}} = \begin{bmatrix} \underset{1\times J_{\Sigma.i}}{\boldsymbol{\Xi\rho}^T_i} & \underset{1\times J_{\Sigma.i}}{\boldsymbol{\Xi L}^T_i} & \underset{1\times J_{\Sigma.i}}{\boldsymbol{\Xi mw}^T_i} \end{bmatrix}^T$ is the vector of measurement errors of the ionosphere-free code $\underset{1\times J_{\Sigma.i}}{\boldsymbol{\Xi\rho}^T_i}$, ionosphere-free phase $\underset{1\times J_{\Sigma.i}}{\boldsymbol{\Xi L}^T_i}$ and of Melbourne-Wübbena $\underset{1\times J_{\Sigma.i}}{\boldsymbol{\Xi mw}^T_i}$ combination, which are formed according to the same principle as the vector $\underset{3J_{\Sigma.i}\times 1}{\boldsymbol{Y}_{net.i}}$.

$$\underset{3J_{\Sigma.i}\times nx_i}{\boldsymbol{H}_{net.i}} = \begin{bmatrix} \underset{3J_{\Sigma.i}\times M}{\boldsymbol{HW}_i} & \underset{3J_{\Sigma.i}\times 3(M+J_i)}{\boldsymbol{d}_i} & \underset{3J_{\Sigma.i}\times 2J_{\Sigma.i}}{\boldsymbol{\Lambda}_i} \end{bmatrix} \tag{5}$$

is the design matrix for the observation vector $\underset{3J_{\Sigma.i}\times 1}{\boldsymbol{Y}_{net.i}}$ and the vector of the estimated variables $\underset{nx_i\times 1}{\boldsymbol{x}_{net.i}}$ (4), where $\underset{3J_{\Sigma.i}\times M}{\boldsymbol{HW}_i} = \begin{bmatrix} \underset{M\times J_{\Sigma.i}}{\boldsymbol{Hw}^T_i} & \underset{M\times J_{\Sigma.i}}{\boldsymbol{Hw}^T_i} & \underset{M\times J_{\Sigma.i}}{\boldsymbol{0}^T_i} \end{bmatrix}^T$ is the $(3J_{\Sigma.i}\times M)$-matrix of mapping functions $w^j_{m.i}$ ($m=\overline{1,M}$, $j=\overline{1,J_{m.i}}$), which are used to convert zenith tropospheric delays (m) into slant delays for actual elevation angles of satellites visible at the m-th station at the i-th epoch;

$$\underset{J_{\Sigma.i}\times M}{\boldsymbol{Hw}_i} = \begin{bmatrix} w^1_{1.i} & \cdots & w^{J_{1.i}}_{1.i} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & w^1_{M.i} & \cdots & w^{J_{M.i}}_{M.i} \end{bmatrix}^T;$$

$$\underset{3J_{\Sigma.i}\times 3(M+J_i)}{\boldsymbol{d}_k} = \begin{bmatrix} \underset{J_{\Sigma.i}\times M}{\mathbf{1}} & \underset{J_{\Sigma.i}\times M}{\mathbf{0}} & \underset{J_{\Sigma.i}\times M}{\mathbf{0}} & \underset{J_{\Sigma.i}\times J_i}{-\boldsymbol{U}} & \underset{J_{\Sigma.i}\times J_i}{\mathbf{0}} & \underset{J_{\Sigma.i}\times J_i}{\mathbf{0}} \\ \underset{J_{\Sigma.i}\times M}{\mathbf{0}} & \underset{J_{\Sigma.i}\times M}{\mathbf{1}} & \underset{J_{\Sigma.i}\times M}{\mathbf{0}} & \underset{J_{\Sigma.i}\times J_k}{\mathbf{0}} & \underset{J_{\Sigma.i}\times J_i}{-\boldsymbol{U}} & \underset{J_{\Sigma.i}\times J_i}{\mathbf{0}} \\ \underset{J_{\Sigma.i}\times M}{\mathbf{0}} & \underset{J_{\Sigma.i}\times M}{\mathbf{0}} & \underset{J_{\Sigma.i}\times M}{\mathbf{1}} & \underset{J_{\Sigma.i}\times J_i}{\mathbf{0}} & \underset{J_{\Sigma.i}\times J_i}{\mathbf{0}} & \underset{J_{\Sigma.i}\times J_i}{-\boldsymbol{U}} \end{bmatrix};$$

$$\underset{J_{\Sigma.i}\times M}{\mathbf{1}} = \begin{bmatrix} \underset{J_{1.i}\times 1}{\mathbf{1}} & \underset{J_{1.i}\times 1}{\mathbf{0}} & \cdots & \underset{J_{1.i}\times 1}{\mathbf{0}} \\ \underset{J_{2.i}\times 1}{\mathbf{0}} & \underset{J_{2.i}\times 1}{\mathbf{1}} & \cdots & \underset{J_{2.i}\times 1}{\mathbf{0}} \\ \vdots & \vdots & \ddots & \vdots \\ \underset{J_{M.i}\times 1.i}{\mathbf{0}} & \cdots & \cdots & \underset{J_{M.i}\times 1}{\mathbf{1}} \end{bmatrix};$$

$$\underset{J_{m.i}\times 1}{\mathbf{1}} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \ m=\overline{1,M};$$

$\underset{J_{\Sigma.i}\times J_i}{\boldsymbol{U}_i} = \begin{bmatrix} \underset{J_i\times J_{1.i}}{\boldsymbol{U}^T_{1.i}} & \underset{J_i\times J_{2.i}}{\boldsymbol{U}^T_{2.i}} & \cdots & \underset{J_i\times J_{M.i}}{\boldsymbol{U}^T_{M.i}} \end{bmatrix}^T$ is the matrix consisting of the M submatrices $\underset{J_{m.i}\times J_i}{\boldsymbol{U}_{m.i}}$, $m=\overline{1,M}$ situated one under another. Each submatrix $\underset{J_{m.i}\times J_i}{\boldsymbol{U}_{m.i}}$ is formed from the m-th ($m=\overline{1,M}$) row of the scenario matrix given in Table 1 by splitting it into the $J_{m.i}$ rows, which are all filled with 0 except just one element which is equal to one. Those unit-elements are situated in successive split rows of the submatrices $\underset{J_{m.i}\times J_i}{\boldsymbol{U}_{m.i}}$ at the same places as in the m-th row of the scenario matrix. Figure 1 shows the example of how the submatrix $\underset{J_{1.i}\times J_i}{\boldsymbol{U}_{1.i}}$ is constructed by splitting the 1-st row of the scenario matrix into the $J_{1.i}$=5 rows.

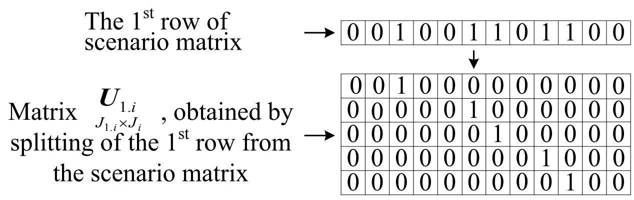


Fig. 1. Splitting the 1-st row of the scenario matrix into $J_{1.i}$=5 rows.

The remaining rows of the scenario matrix are split in the same way;

$$\underset{3J_{\Sigma.i}\times 2J_{\Sigma.i}}{\boldsymbol{\Lambda}_i} = \begin{bmatrix} \underset{J_{\Sigma.i}\times J_{\Sigma.i}}{\mathbf{0}} & \underset{J_{\Sigma.i}\times J_{\Sigma.i}}{\mathbf{0}} \\ -\lambda_{\Delta nifr}\underset{J_{\Sigma.i}\times J_{\Sigma.i}}{\boldsymbol{E}_i} & -\lambda_{n_2ifr}\underset{J_{\Sigma.i}\times J_{\Sigma.i}}{\boldsymbol{E}_i} \\ \underset{J_{\Sigma.i}\times J_{\Sigma.i}}{\mathbf{0}} & -\lambda_{mw}\underset{J_{\Sigma.i}\times J_{\Sigma.i}}{\boldsymbol{E}_i} \end{bmatrix};$$

$\underset{J_{\Sigma.i}\times J_{\Sigma.i}}{\boldsymbol{E}_i}$ is the identity matrix of dimension ($J_{\Sigma.i}\times J_{\Sigma.i}$).

The matrix $\boldsymbol{V}_{net.i}$ whose columns are the basis vectors of the null space of the design matrix $\underset{3J_{\Sigma.i}\times nx_i}{\boldsymbol{H}_{net.i}}$ (5) was obtained:

$$\underset{nx_i\times(2M+2J_i+1)}{\boldsymbol{V}_{net.i}} = \left[\underset{nx_i\times 3}{\boldsymbol{V1}_{net.i}} \quad \underset{nx_i\times M}{\boldsymbol{V2}_{net.i}} \quad \underset{nx_i\times M}{\boldsymbol{V3}_{net.i}} \quad \underset{nx_i\times(J_i-1)}{\boldsymbol{V4}_{net.i}} \quad \underset{nx_i\times(J_i-1)}{\boldsymbol{V5}_{net.i}}\right] \tag{6}$$

where

$$\underset{nx_i\times 3}{\boldsymbol{V1}_{net.i}} = \begin{bmatrix} \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} \\ \underset{M\times1}{\mathbf{1}} & \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} \\ \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{1}} & \underset{M\times1}{\mathbf{0}} \\ \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{1}} \\ \underset{J_i\times1}{\mathbf{1}} & \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{1}} & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{1}} \\ \underset{2J_{\Sigma.i}\times1}{\mathbf{0}} & \underset{2J_{\Sigma.i}\times1}{\mathbf{0}} & \underset{2J_{\Sigma.i}\times1}{\mathbf{0}} \end{bmatrix}, \quad \underset{nx_i\times M}{\boldsymbol{V2}_{net.i}} = \begin{bmatrix} \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} & \cdots & \underset{M\times1}{\mathbf{0}} \\ \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} & \cdots & \underset{M\times1}{\mathbf{0}} \\ \lambda_{\Delta nifr}\underset{M\times1}{\mathbf{1}^1} & \lambda_{\Delta nifr}\underset{M\times1}{\mathbf{1}^2} & \cdots & \lambda_{\Delta nifr}\underset{M\times1}{\mathbf{1}^M} \\ \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} & \cdots & \underset{M\times1}{\mathbf{0}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \cdots & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \cdots & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \cdots & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_{\Sigma.i}\times1}{\mathbf{1}^1} & \underset{J_{\Sigma.i}\times1}{\mathbf{1}^2} & \cdots & \underset{J_{\Sigma.i}\times1}{\mathbf{1}^M} \\ \underset{J_{\Sigma.i}\times1}{\mathbf{0}} & \underset{J_{\Sigma.i}\times1}{\mathbf{0}} & \cdots & \underset{J_{\Sigma.i}\times1}{\mathbf{0}} \end{bmatrix},$$

$$\underset{nx_i\times M}{\boldsymbol{V3}_{net.i}} = \begin{bmatrix} \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} & \cdots & \underset{M\times1}{\mathbf{0}} \\ \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} & \cdots & \underset{M\times1}{\mathbf{0}} \\ \lambda_{n_2ifr}\underset{M\times1}{\mathbf{1}^1} & \lambda_{n_2ifr}\underset{M\times1}{\mathbf{1}^2} & \cdots & \lambda_{n_2ifr}\underset{M\times1}{\mathbf{1}^M} \\ \lambda_{mw}\underset{M\times1}{\mathbf{1}^1} & \lambda_{mw}\underset{M\times1}{\mathbf{1}^2} & \cdots & \lambda_{mw}\underset{M\times1}{\mathbf{1}^M} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \cdots & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \cdots & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \cdots & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_{\Sigma.i}\times1}{\mathbf{0}} & \underset{J_{\Sigma.i}\times1}{\mathbf{0}} & \cdots & \underset{J_{\Sigma.i}\times1}{\mathbf{0}} \\ \underset{J_{\Sigma.i}\times1}{\mathbf{1}^1} & \underset{J_{\Sigma.i}\times1}{\mathbf{1}^2} & \cdots & \underset{J_{\Sigma.i}\times1}{\mathbf{1}^M} \end{bmatrix},$$

are the matrices, where $\underset{M\times1}{\mathbf{1}^m}$ is the column vector made of M nulls excluding the unit in the m-th, $m=\overline{1,M}$ position; $\underset{J_{\Sigma.i}\times1}{\mathbf{1}^m}$ is the $J_{\Sigma.i}$-vector formed from the M subvectors $\underset{J_{m.i}\times1}{\mathbf{1}^m}$, $m=\overline{1,M}$. All the $J_{m.i}$-subvectors $\underset{J_{m.i}\times1}{\mathbf{1}^m}$ are zeroes, except for the m-th consisting of ones.

$$\underset{nx_i\times(J_i-1)}{\boldsymbol{V4}_{net.i}} = \begin{bmatrix} \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} & \cdots & \underset{M\times1}{\mathbf{0}} \\ \underset{3M\times1}{\mathbf{0}} & \underset{3M\times1}{\mathbf{0}} & \cdots & \underset{3M\times1}{\mathbf{0}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \cdots & \underset{J_i\times1}{\mathbf{0}} \\ -\lambda_{\Delta nifr}\underset{J_i\times1}{\mathbf{1}^1} & -\lambda_{\Delta nifr}\underset{J_i\times1}{\mathbf{1}^2} & \cdots & -\lambda_{\Delta nifr}\underset{J_i\times1}{\mathbf{1}^{J_i-1}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \cdots & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_{1.i}\times1}{\mathbf{1s}_1^1} & \underset{J_{1.i}\times1}{\mathbf{1s}_1^2} & \cdots & \underset{J_{1.i}\times1}{\mathbf{1s}_1^{J_i-1}} \\ \underset{J_{2.i}\times1}{\mathbf{1s}_2^1} & \underset{J_{2.i}\times1}{\mathbf{1s}_2^2} & \cdots & \underset{J_{2.i}\times1}{\mathbf{1s}_2^{J_i-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \underset{J_{M.i}\times1}{\mathbf{1s}_M^1} & \underset{J_{M.i}\times1}{\mathbf{1s}_M^2} & \cdots & \underset{J_{M.i}\times1}{\mathbf{1s}_M^{J_i-1}} \\ \underset{J_{\Sigma.i}\times1}{\mathbf{0}} & \underset{J_{\Sigma.i}\times1}{\mathbf{0}} & \cdots & \underset{J_{\Sigma.i}\times1}{\mathbf{0}} \end{bmatrix}$$

$$\underset{nx_i\times(J_i-1)}{\boldsymbol{V5}_{net.i}} = \begin{bmatrix} \underset{M\times1}{\mathbf{0}} & \underset{M\times1}{\mathbf{0}} & \cdots & \underset{M\times1}{\mathbf{0}} \\ \underset{3M\times1}{\mathbf{0}} & \underset{3M\times1}{\mathbf{0}} & \cdots & \underset{3M\times1}{\mathbf{0}} \\ \underset{J_i\times1}{\mathbf{0}} & \underset{J_i\times1}{\mathbf{0}} & \cdots & \underset{J_i\times1}{\mathbf{0}} \\ \underset{J_k\times1}{\mathbf{0}} & \underset{J_k\times1}{\mathbf{0}} & \cdots & \underset{J_k\times1}{\mathbf{0}} \\ \lambda_{mw}\underset{J_i\times1}{\mathbf{1}^1} & \lambda_{mw}\underset{J_i\times1}{\mathbf{1}^2} & \cdots & \lambda_{mw}\underset{J_i\times1}{\mathbf{1}^{J_i-1}} \\ -\frac{\lambda_{n_2ifr}}{\lambda_{\Delta nifr}}\underset{J_{1.i}\times1}{\mathbf{1s}_1^1} & -\frac{\lambda_{n_2ifr}}{\lambda_{\Delta nifr}}\underset{J_{1.i}\times1}{\mathbf{1s}_1^2} & \cdots & -\frac{\lambda_{n_2ifr}}{\lambda_{\Delta nifr}}\underset{J_{1.i}\times1}{\mathbf{1s}_1^{J_i-1}} \\ -\frac{\lambda_{n_2ifr}}{\lambda_{\Delta nifr}}\underset{J_{2.i}\times1}{\mathbf{1s}_2^1} & -\frac{\lambda_{n_2ifr}}{\lambda_{\Delta nifr}}\underset{J_{2.i}\times1}{\mathbf{1s}_2^2} & \cdots & -\frac{\lambda_{n_2ifr}}{\lambda_{\Delta nifr}}\underset{J_{2.i}\times1}{\mathbf{1s}_2^{J_i-1}} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\lambda_{n_2ifr}}{\lambda_{\Delta nifr}}\underset{J_{M.i}\times1}{\mathbf{1s}_M^1} & -\frac{\lambda_{n_2ifr}}{\lambda_{\Delta nifr}}\underset{J_{M.i}\times1}{\mathbf{1s}_M^2} & \cdots & -\frac{\lambda_{n_2ifr}}{\lambda_{\Delta nifr}}\underset{J_{M.i}\times1}{\mathbf{1s}_M^{J_i-1}} \\ \underset{J_{1.i}\times1}{\mathbf{1s}_1^1} & \underset{J_{1.i}\times1}{\mathbf{1s}_1^2} & \cdots & \underset{J_{1.i}\times1}{\mathbf{1s}_1^{J_i-1}} \\ \underset{J_{2.i}\times1}{\mathbf{1s}_2^1} & \underset{J_{2.i}\times1}{\mathbf{1s}_2^2} & \cdots & \underset{J_{2.i}\times1}{\mathbf{1s}_2^{J_i-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \underset{J_{M.i}\times1}{\mathbf{1s}_M^1} & \underset{J_{M.i}\times1}{\mathbf{1s}_M^2} & \cdots & \underset{J_{M.i}\times1}{\mathbf{1s}_M^{J_i-1}} \end{bmatrix}$$

are the matrices, in which $\underset{J_{m.i}\times1}{\mathbf{1s}_{m.i}^j}$ $j=\overline{1,J_i-1}$, $m=\overline{1,M}$ are the column vectors of the submatrices $\underset{J_{m.i}\times J_i}{\boldsymbol{U}_{m.i}}$ (excluding the last column vector) of the matrix $\underset{J_{\Sigma.i}\times J_i}{\boldsymbol{U}_i}$, the construction of which was described earlier. For example, the $J_{1,i}$-1 column vectors $\left[\underset{J_{1.i}\times1}{\mathbf{1s}_1^1} \quad \underset{J_{1.i}\times1}{\mathbf{1s}_1^2} \quad \cdots \quad \underset{J_{1.i}\times1}{\mathbf{1s}_1^{J_i-1}}\right]$ form the matrix $\underset{J_{1.i}\times J_i}{\boldsymbol{U}_{1.i}}$ without the last column. This matrix is shown in Fig. 1, where $J_{1.i}=5$ and $J_i=12$.

The number of columns in the matrix $\underset{nx_i\times(2M+2J_i+1)}{\boldsymbol{V}_{net.i}}$ (6) is $2M+2J_i+1$ and, therefore, the rank deficiency of the design matrix $\underset{3J_{\Sigma.i}\times nx_i}{\boldsymbol{H}_{net.i}}$ (5) is $dfh_i=2M+2J_i+1$, and its rank is $rnkh_i=nx_i-dfhk=2M+J_i+2J_{\Sigma.i}-1$. It can be seen that in the network solution, by contrast to the user solution, the rank deficiency dfh_i of the matrix $\underset{3J_{\Sigma.i}\times nx_i}{\boldsymbol{H}_{net.i}}$ (5) depends on the number J_i of observed satellites and the number of the stations M.

As in the user solution, the linear system (3) is singular, i.e., it has an infinite set of solutions lying in the dfh_i-dimensional solution space parallel-shifted with respect to the null-space $\underset{nx_i\times(2M+2J_i+1)}{\boldsymbol{V}_{net.i}}$ (6). However, as can be seen from (6), the first M elements of the null-space basis vectors $\underset{nx_i \times (2M+2J_i+1)}{\boldsymbol{V}_{net.i}}$ (6) are zeroes. This means that the set of solutions of the system (3) is orthogonal to those axes of the space variables, which correspond to the first M elements of the vector $\underset{nx_i \times 1}{\boldsymbol{x}_{net.i}}$ (4). As it can be seen from (4), these elements compose the vector $\underset{M \times 1}{\boldsymbol{\Delta D}_i}$. Thus, the first M coordinates of the points lying in the dfh_i-dimensional solution space are the same for all points of this space and, hence, all M elements of the vector $\underset{M \times 1}{\boldsymbol{\Delta D}_i}$ can be estimated unambiguously. The remaining elements of the state vector $\underset{nx_i \times 1}{\boldsymbol{x}_{net.i}}$ (4) are co-variables, i.e., they cannot be estimated per se, but only some linear combinations thereof are estimable. The co-variables include the elements of the integer vectors $\underset{J_{\Sigma.i} \times 1}{\boldsymbol{N}1}$, $\underset{J_{\Sigma.i} \times 1}{\boldsymbol{N}_{mw}}$, as well as the vectors $\underset{J_i \times 1}{\boldsymbol{dt}_{\rho.i}}$, $\underset{J_i \times 1}{\boldsymbol{dt}_{L.i}}$, $\underset{J_i \times 1}{\boldsymbol{b}_{mw.i}}$, estimation of which is the purpose of the network solution. Now we are going to derive the expressions for linear combinations, which include these vectors.

As in the case of a user solution, the determination of linear combinations formed by co-variables is performed with the help of the S-transformation theory [2–5], i.e., by projecting all the points of the space of the variables onto the S-subspace along the null-space $\underset{nx_i \times (2M+2J_i+1)}{\boldsymbol{V}_{net.i}}$ (6). The dimension of the S-subspace equals the rank of the matrix $\underset{3J_{\Sigma.i} \times nx_i}{\boldsymbol{H}_{net.i}}$ (5) of the linear system (3). Similar to a user solution, the S-subspace of the network solution can be obtained from the system of normal equations [2, 3]

$$\left(\underset{dfh_i \times nx_i}{\boldsymbol{S}^{\perp}_{net.i}}\right)^T \underset{nx_i \times 1}{\boldsymbol{x}_{net.i}} = \underset{dfh_i \times 1}{\boldsymbol{0}} \tag{7}$$

where $\underset{nx_i \times dfh_i}{\boldsymbol{S}^{\perp}_{net.i}}$ is the matrix of the rank dfh_i, and all the dfh_i column vectors are orthogonal to the S-subspace. By projecting point coordinates, we get a new vector of estimated variables $\underset{nx_i \times 1}{\boldsymbol{x}_{net.s.i}}$ of the same dimension as the original vector $\underset{nx_i \times 1}{\boldsymbol{x}_{net.i}}$ (4). The relation of vectors $\underset{nx_i \times 1}{\boldsymbol{x}_{net.s.i}}$, $\underset{nx_i \times 1}{\boldsymbol{x}_{net.i}}$ is defined by the expression

$$\underset{nx_i \times 1}{\boldsymbol{x}_{net.s.i}} = \underset{nx_i \times nx_i}{\boldsymbol{P}_{net.i}} \underset{nx_i \times 1}{\boldsymbol{x}_{net.i}} \tag{8}$$

where $\underset{nx_i \times nx_i}{\boldsymbol{P}_{net.i}} = \underset{nx_i \times nx_i}{\boldsymbol{E}_i} - \underset{nx_i \times dfh_i}{\boldsymbol{V}_{net.i}} \left(\left(\underset{dfh_i \times nx_i}{\boldsymbol{S}^{\perp}_{net.i}}\right)^T \underset{nx_i \times dfh_i}{\boldsymbol{V}_{net.i}} \right)^{-1} \left(\underset{dfh_i \times nx_i}{\boldsymbol{S}^{\perp}_{net.i}}\right)^T$.

To preserve the integer nature of linear combinations of integer co-variables, which are the components of the integer vectors $\underset{J_{\Sigma.i} \times 1}{\boldsymbol{N}1}$ and $\underset{J_{\Sigma.i} \times 1}{\boldsymbol{N}_{mw}}$ included in the initial state vector $\underset{nx_i \times 1}{\boldsymbol{x}_{net.i}}$ (4), the columns of the matrix $\underset{nx_i \times dfh_i}{\boldsymbol{S}^{\perp}_{net.i}}$ should be set in the same way as in the user solution [1]. Namely, all the elements in the columns of the matrix $\underset{nx_i \times dfh_i}{\boldsymbol{S}^{\perp}_{net.i}}$ should be set to zero, excluding the only element equal to 1.

The estimate $\underset{nx_i \times 1}{\hat{\boldsymbol{x}}_{net.s.i}}$ of a new state vector $\underset{nx_i \times 1}{\boldsymbol{x}_{net.s.i}}$ in the network solution can be found from the solution of the extended system of linear equations obtained by combining the linear systems (3) and (7)

$$\begin{bmatrix} \underset{3J_{\Sigma.i} \times 1}{\boldsymbol{Y}_{net.i}} \\ \underset{dfh_i \times 1}{\boldsymbol{0}} \end{bmatrix} = \begin{bmatrix} \underset{3J_{\Sigma.i} \times nx_i}{\boldsymbol{H}_{net.i}} \\ \left(\underset{dfh_i \times nx_i}{\boldsymbol{S}^{\perp}_{net.i}}\right)^T \end{bmatrix} \cdot \underset{nx_i \times 1}{\boldsymbol{x}_{net.s.i}} + \begin{bmatrix} \underset{3J_{\Sigma.i} \times 1}{\boldsymbol{\Xi}_{net.i}} \\ \underset{dfh_i \times 1}{\boldsymbol{0}} \end{bmatrix}. \tag{9}$$

From the equation $\left(\underset{dfh_i \times nx_i}{\boldsymbol{S}^{\perp}_{net.i}}\right)^T \underset{nx_i \times 1}{\boldsymbol{x}_{net.i}} = \underset{dfh_i \times 1}{\boldsymbol{0}}$ (7), taking into account the specific type of the columns of the matrix $\underset{nx_i \times dfh_i}{\boldsymbol{S}^{\perp}_{net.i}}$, it follows that the elements of the solution vector $\underset{nx_i \times 1}{\hat{\boldsymbol{x}}_{net.s.i}}$ of the system (9) standing on the places defined by the position of the units in the dfh_i columns of the matrix $\underset{nx_i \times dfh_i}{\boldsymbol{S}^{\perp}_{net.i}}$ equal zero. But if it is known in advance that dfh_i of the elements of the solution vector $\underset{nx_i \times 1}{\hat{\boldsymbol{x}}_{net.s.i}}$ of the system (9) are zero, hence, the estimates of the remaining $rnkh_i$ elements of the solution vector of the system (3) can be obtained by solving a simpler system of the linear equations

$$\underset{3J_{\Sigma.i} \times 1}{\boldsymbol{Y}_{net.i}} = \underset{3J_{\Sigma.i} \times (nx_i - dfh_i)}{\boldsymbol{H}_{net.cmpr.i}} \underset{(nx_i - dfh_i) \times 1}{\boldsymbol{x}_{net.cmpr.i}} + \underset{3J_{\Sigma.i} \times 1}{\boldsymbol{\Xi}_i} \tag{10}$$

where $\underset{3J_{\Sigma.i} \times (nx_i - dfh_i)}{\boldsymbol{H}_{net.cmpr.i}}$ is the compressed form of the initial matrix $\underset{3J_{\Sigma.i} \times nx_i}{\boldsymbol{H}_{net.i}}$ (5), where the dfh_i columns are cut from the positions in which the units are located in the dfh_i columns of the matrix $\underset{nx_i \times dfh_i}{\boldsymbol{S}^{\perp}_{net.i}}$; $\underset{(nx_i - dfh_i) \times 1}{\boldsymbol{x}_{net.cmpr.i}}$ is the compressed state vector obtained from the initial state vector $\underset{nx_i\times 1}{\boldsymbol{x}_{net.s.i}}$, where all the zeroes are excluded. Hence, we see that positions of ones in dfh_i columns of the matrix $\underset{nx_i\times dfh_i}{\boldsymbol{S}^{\perp}_{net.i}}$ determine the positions of excluded columns dfh_i of the initial matrix $\underset{3J_{\Sigma.i}\times nx_i}{\boldsymbol{H}_{net.i}}$ (5). Nevertheless, this manipulation of dropping the columns makes the rank of the compressed matrix $\underset{3J_{\Sigma.i}\times(nx_i-dfh_i)}{\boldsymbol{H}_{net.cmpr.i}}$ match the rank $rnkh_i$ of the initial matrix $\underset{3J_{\Sigma.i}\times nx_i}{\boldsymbol{H}_{net.i}}$ (5). This requirement is satisfied if one of the M ground stations with the number r (reference) $1\leq r\leq M$ is defined as reference and if the matrix $\underset{nx_i\times dfh_i}{\boldsymbol{S}^{\perp}_{net.i}}$ is constructed by merging three submatrices $\underset{nx_i\times dfh_i}{\boldsymbol{S}^{\perp}_{net.i}} = \left[\underset{nx_i\times 3}{\boldsymbol{S}^{\perp}_{net.time.i}} \quad \underset{nx_i\times(dfh_i-3)/2}{\boldsymbol{S}^{\perp}_{net.N1.i}} \quad \underset{nx_i\times(dfh_i-3)/2}{\boldsymbol{S}^{\perp}_{net.N_{mw}.i}}\right]$. The units in three columns of the first submatrix $\underset{nx_i\times 3}{\boldsymbol{S}^{\perp}_{net.time.i}}$ are located on the positions $M+r$, $2M+r$, $3M+r$, respectively. Placing the units in the columns of the submatrix $\underset{nx_i\times 3}{\boldsymbol{S}^{\perp}_{net.time.i}}$ this way, we posit them in the same positions as the ionosphere-free code $dT_{\rho.r.i}$ and phase $dT_{L.r.i}$ clock offsets and the biases of the Melbourne-Wübbena combination $b_{mw.r}$ for the r-th ground station in the initial state vector $\underset{nx_i\times 1}{\boldsymbol{x}_{net.i}}$. The location of units in $(dfh_i-3)/2=M+J_i-1$ columns of the submatrices $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N1.i}}$, $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N_{mw}.i}}$ is determined by position of units in the spanning tree matrix $\underset{M\times J_i}{\boldsymbol{STM}}$ elements of which are either zero or one. The matrix $\underset{M\times J_i}{\boldsymbol{STM}}$ has the same dimension $M\times J_i$ as the scenario matrix, while the number of units in $\underset{M\times J_i}{\boldsymbol{STM}}$ equals $M+J_i-1$, and all these units constitute the unit-subspace of the scenario matrix. Hence, $\underset{M\times J_i}{\boldsymbol{STM}}$ can be shown by highlighting in grey some units in the scenario matrix. An example of such highlighting is shown in Table 1.

$\underset{M\times J_i}{\boldsymbol{STM}}$ is obtained from the scenario matrix through a special algorithm, which can be found in [2, 6–9]. In particular, Prim's algorithm applicable to weighted non-directed graphs is briefly described as follows [2]:

1. Take the edge with the highest weight and place two thereby connected nodes in the set V_1;
2. Take the set of edges that connect nodes from $V-V_1$ to nodes from V_1, and select the edge with the highest weight;
3. Add the node belonging to $V-V_1$ of the edges selected in step 2 to V_1;
4. While $V-V_1\neq 0$, go to step 2.

The number of units in $\underset{M\times J_i}{\boldsymbol{STM}}$, i.e., $M+J_i-1$, equals the number of columns of submatrices $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N1.i}}$, $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N_{mw}.i}}$, i.e., each unit of $\underset{M\times J_i}{\boldsymbol{STM}}$ is placed in a certain way within the rows numbered $l=\overline{4M+3J_i+1,\ \ 4M+3J_i+J_{\Sigma.i}}$ of one of the columns of the submatrix $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N1.i}}$ and in the rows with the numbers $l=\overline{4M+3J_i+J_{\Sigma.i}+1,\ \ 4M+3J_i+2J_{\Sigma.i}}$ of one of the columns of the submatrix $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N_{mw}.i}}$ according to the following algorithm. All the units comprising the scenario matrix are indexed by the first lower index in the order of their location from the left to the right and for rows from up to down in the range from 1 to $J_{\Sigma.i}$. Then, the units of the scenario matrix, which comprise the matrix $\underset{M\times J_i}{\boldsymbol{STM}}$, are indexed by the second lower index in the same order in the range from 1 to $M+J_i-1$. As a result, units comprising the matrix $\underset{M\times J_i}{\boldsymbol{STM}}$, will have two indices, and the remaining units will have only one. An example of such indexation is shown in Table 1. Dual-indexed units comprising the matrix $\underset{M\times J_i}{\boldsymbol{STM}}$ are highlighted in grey. Let us denote the first index of dual-indexed units as μ ($\mu=\overline{1, J_{\Sigma.i}}$), while the second index is ν ($\nu=\overline{1, M+J_i-1}$). Each dual-indexed unit corresponds to the column ν of the matrices $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N1.i}}$ and $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N_{mw}.i}}$. The element $4M+3J_i+\mu$ of ν-th column of the matrix $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N1.i}}$ is replaced by one, and the remaining elements of this column are zero. Similarly, the element $4M+3J_i+J_{\Sigma.i}+\mu$ of ν-th column of the matrix $\underset{nx_i\times(M+J_i-1)}{\boldsymbol{S}^{\perp}_{net.N_{mw}.i}}$ is replaced by one, while the remaining elements of this column are zeroes.

A unique solution of the system (10) can be obtained if the number of rows of the compressed matrix $\underset{3J_{\Sigma.i}\times(nx_i-dfh_i)}{\boldsymbol{H}_{net.cmpr.i}}$

is greater or equal to the dimension of the compressed state vector $\underset{3J_{\Sigma.i}\times(nx_i-dfh_i)}{\boldsymbol{H}_{net.cmpr.i}}$, i.e., the condition $3J_{\Sigma.i} \geq (nx_i-dfh_i) = 2M+J_i+2J_{\Sigma.i}-1$ should be fulfilled. Hence, we obtain the following constraint $J_{\Sigma.i} \geq 2M+J_i-1$, which should be met to get a unique network solution.

In general, the analytical calculation of the vector $\underset{nx_i\times 1}{\boldsymbol{x}_{net.s.i}}$ by the formula (8), which will include linear combinations of elements of the vectors $\underset{J_{\Sigma.i}\times 1}{\boldsymbol{N}1}$, $\underset{J_{\Sigma.i}\times 1}{\boldsymbol{N}_{mw}}$, as well as the vectors $\underset{J_i\times 1}{\boldsymbol{dt}_{\rho.i}}$, $\underset{J_i\times 1}{\boldsymbol{dt}_{L.i}}$ and $\underset{J_i\times 1}{\boldsymbol{b}_{mw.i}}$, is cumbersome. To reduce the complexity, let us consider the computation of the vector $\underset{nx_i\times 1}{\boldsymbol{x}_{net.s.i}}$ for an extremely simplified case when $M=3$, $J_i=4$, for which the scenario matrix and $\underset{M\times J_i}{\boldsymbol{STM}}$ are given in Table 2.

Table 2. SV visibility by the ground stations for the simplified case with $M=3$, $J_i=4$, $J_{\Sigma.i}=9$

	m	Numbers of SVs visible by ground stations				J_m
		1	2	3	4	
Ground network stations	1	1_{11}	1_{22}	1_{33}	0	3
	2	1_{44}	0	1_{5}	1_{65}	3
	3	0	1_{7}	1_{8}	1_{96}	3

For the case $M=3$, $J_i=4$, $J_{\Sigma.i}=9$, the vector of the variables $\underset{42\times 1}{\boldsymbol{x}_{net.s.i}}$ (8) and the corresponding projection matrix $\underset{42\times 42}{\boldsymbol{P}_{net.k}}$ will be quite bulky even in the simplified case. Hence, the expressions below are presented only for the subvectors of interest $\underset{M\times 1}{\boldsymbol{\Delta D}_{s.i}}$, $\underset{4\times 1}{\boldsymbol{dt}_{\rho.s.i}}$, $\underset{4\times 1}{\boldsymbol{dt}_{L.s.i}}$, $\underset{4\times 1}{\boldsymbol{b}_{mw.s.i}}$, $\underset{9\times 1}{\boldsymbol{N1}_{s.i}}$, $\underset{9\times 1}{\boldsymbol{N}_{mw.s.i}}$ of the vector $\underset{42\times 1}{\boldsymbol{x}_{net.s.i}}$ (8) obtained by symbolic computations in MATLAB for $r=1$.

$$\underset{M\times 1}{\boldsymbol{\Delta D}_{i.k}} = \begin{bmatrix} \Delta D_{1.i} \\ \Delta D_{2.i} \\ \Delta D_{3.i} \end{bmatrix},$$

$$\underset{4\times 1}{\boldsymbol{dt}_{\rho.s.i}} = \begin{bmatrix} dt^1_{\rho.i} - dT_{\rho.r.i} \\ dt^2_{\rho.i} - dT_{\rho.r.i} \\ dt^3_{\rho.i} - dT_{\rho.r.i} \\ dt^4_{\rho.i} - dT_{\rho.r.i} \end{bmatrix},$$

$$\underset{4\times 1}{\boldsymbol{dt}_{L.s.i}} = \begin{bmatrix} dt^1_{L.i} - dT_{L.r.i} + \lambda_{\Delta nifr} N1^1_1 + \lambda_{n2ifr} N^1_{mw1} \\ dt^2_{L.i} - dT_{L.r.i} + \lambda_{\Delta nifr} N1^2_1 + \lambda_{n2ifr} N^2_{mw1} \\ dt^3_{L.i} - dT_{L.r.i} + \lambda_{\Delta nifr} N1^3_1 + \lambda_{n2ifr} N^3_{mw1} \\ dt^4_{L.i} - dT_{L.r.i} + \lambda_{\Delta nifr}\left(N1^1_1 - N^1_2 + N^4_2\right) + \lambda_{n2ifr}\left(N^1_{mw1} - N^1_{mw2} + N^4_{mw2}\right) \end{bmatrix},$$

$$\underset{4\times 1}{\boldsymbol{b}_{mw.s.i}} = \begin{bmatrix} b^1_{mw} - b_{mw.r} + \lambda_{mw} N^1_{mw.1} \\ b^2_{mw} - b_{mw.r} + \lambda_{mw} N^2_{mw.1} \\ b^3_{mw} - b_{mw.r} + \lambda_{mw} N^3_{mw.1} \\ b^4_{mw} - b_{mw.r} + \lambda_{mw}\left(N^1_{mw.1} - N^1_{mw.2} + N^4_{mw.2}\right) \end{bmatrix},$$

$$\underset{9\times 1}{\boldsymbol{N1}_{s.i}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ N1^3_2 - N1^1_2 + N1^1_1 - N1^3_1 \\ 0 \\ N1^2_3 - N1^4_3 + N1^4_2 - N1^1_2 + N1^1_1 - N1^2_1 \\ N1^3_3 - N1^4_3 + N1^4_2 - N1^1_2 + N1^1_1 - N1^3_1 \\ 0 \end{bmatrix},$$

$$\underset{9\times 1}{\boldsymbol{N}_{mw.s.i}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ N1^3_{mw2} - N1^1_{mw2} + N1^1_{mw1} - N1^3_{mw1} \\ 0 \\ N1^2_{mw3} - N1^4_{mw3} + N1^4_{mw2} - N1^1_{mw2} + N1^1_{mw1} - N1^2_{mw1} \\ N1^3_{mw3} - N1^4_{mw3} + N1^4_{mw2} - N1^1_{mw2} + N1^1_{mw1} - N1^3_{mw1} \\ 0 \end{bmatrix}. \qquad (11)$$

As can be seen from the expressions (11), the elements of the vector $\underset{M\times 1}{\boldsymbol{\Delta D}^T} = \left[\Delta D_{1.i} \quad \Delta D_{2.i} \quad \Delta D_{3.i}\right]^T$ of uncompensated wet component of the zenith tropospheric delays (m) at the locations of three ground stations, as expected, are unbiased. The initial integer ambiguities $\underset{9\times 1}{\boldsymbol{N1}_i}$ and $\underset{9\times 1}{\boldsymbol{N}_{mw.i}}$ are estimated with biases, i.e., as a part of the linear combinations $\underset{9\times 1}{\boldsymbol{N1}_{s.i}}$ and $\underset{9\times 1}{\boldsymbol{N}_{mw.s.i}}$, which are also integer. We are interested in the vectors $\underset{4\times 1}{\boldsymbol{dt}_{\rho.i}}$, $\underset{4\times 1}{\boldsymbol{dt}_{L.i}}$, and $\underset{4\times 1}{\boldsymbol{b}_{mw.i}}$ which are also estimated with biases, i.e., are composed of the linear combinations $\underset{4\times 1}{\boldsymbol{dt}_{\rho.s.i}}$, $\underset{4\times 1}{\boldsymbol{dt}_{L.s.i}}$, and $\underset{4\times 1}{\boldsymbol{b}_{mw.s.i}}$. However, as it is shown in (11), for all J_i SVs, the estimation biases of the variables included in the vector $\underset{4\times 1}{\boldsymbol{dt}_{\rho.i}}$ are the same and equal to ionosphere-free code clock offsets $dT_{\rho.r.i}$ of the reference station; the estimation biases of the variables included in the vector $\underset{4\times1}{\boldsymbol{dt}_{L.i}}$ are accurate within an integer combination of the wavelengths $\lambda_{\Delta nifr}$, λ_{n_2ifr}, and the estimation biases of the variables included in the vector $\underset{4\times1}{\boldsymbol{b}_{mw.i}}$ are accurate within an integer number of the wavelengths λ_{mw}. It leads to respective biases in the residuals of the ionosphere-free code $\Delta\rho^j_{ifr.i} = \rho^j_{ifr.i} - R^j_{c.i} + dt^j_{\rho.ifr.i}$, carrier phase $\Delta L^j_{ifr.i} = L^j_{ifr.i} - R^j_{c.i} + dt^j_{L.ifr.i}$ and Melbourne-Wübbena combinations $\Delta mw^j_i = mw^j_i + b^j_{mw.i}$ measurements included in the left part of the system of equations in the user solution. However, the properties of this system are such that these biases in the left part do not change the estimates of the unambiguously estimated variables Δx, Δy, Δz, ΔD_i. Thus, to obtain estimates of the variables Δx, Δy, Δz, ΔD_i in the user solution instead of estimates of variables which are components of the vectors $\underset{J_i\times1}{\boldsymbol{dt}_{\rho.i}}$, $\underset{J_i\times1}{\boldsymbol{dt}_{L.i}}$, $\underset{J_i}{\boldsymbol{b}_{mw.i}}$, one can use their biased equivalents, which are components of the vectors $\underset{J_i\times1}{\boldsymbol{dt}_{\rho.s.i}}$, $\underset{J_i\times1}{\boldsymbol{dt}_{L.s.i}}$, $\underset{J_i\times1}{\boldsymbol{b}_{mw.s.i}}$ (11).

The algorithms of solving the linear equation system (10) taking into account the integer nature of the part of its variables, are the basis of the algorithms of estimating the variables which are the part of the vectors $\underset{J_i\times1}{\boldsymbol{dt}_{\rho.s.i}}$, $\underset{J_i\times1}{\boldsymbol{dt}_{L.s.i}}$ and $\underset{J_i\times1}{\boldsymbol{b}_{mw.s.i}}$, which are the aim of solving the network solution of the Integer PPP. Unfortunately, the restrictions for the article's volume do not allow us to explore these algorithms in this paper. We can only refer a reader to existing literature on the methods of linear recurrent estimating [10, 11] and on phase integer ambiguity resolution [12–15].

Examples of determination of precise corrections and their features

Two versions of the network solution were implemented. The first solution was obtained with using 5 European stations assuming the permanence of the SV constellation (all stations of the ground network receive measurements from the same set of 6 SVs). Figure 2 shows decoupled code and carrier phase satellite corrections calculated for one of the 6 satellites in the first version of the network solution.

According to (11), the bias between code and phase corrections shown in Fig. 2 may differ from the true one by an integer number of the wavelengths $\lambda_{\Delta nifr}$, λ_{n_2ifr}. It can be seen that this bias is constant during the permanent observation scenario.

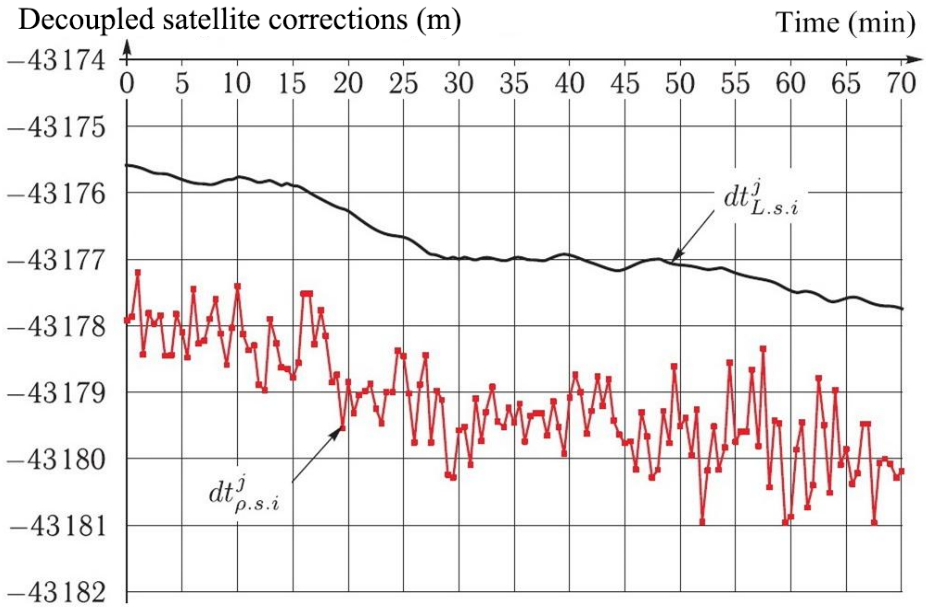


Fig. 2. Decoupled satellite corrections (code and phase) for one of 6 SVs calculated for the first version of the network solution.

Fig. 3. SDCM network stations used in the second version of a network solution.

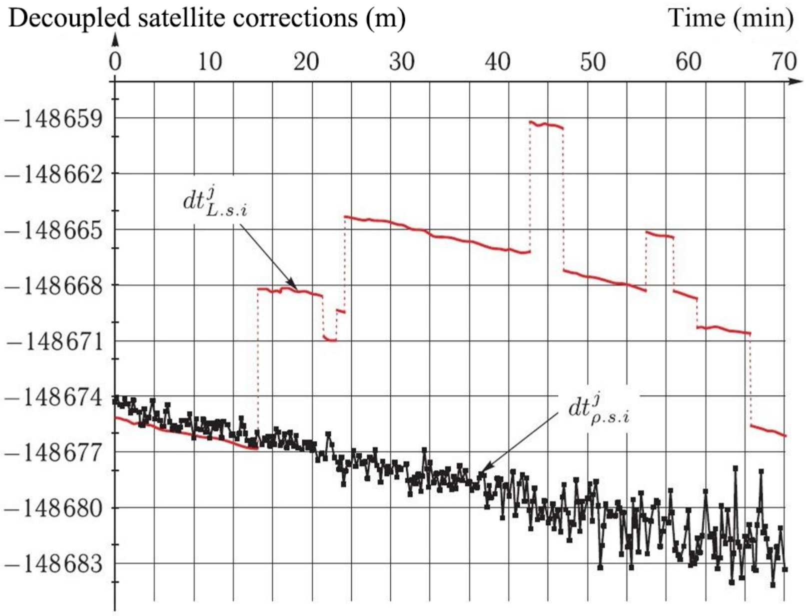


Fig. 4. Decoupled satellite corrections (code and phase) for one of the SVs calculated in the second version of a network solution.

Within the second version of a network solution, measurements from 10 stations of Russian SBAS (SDCM) network were used (highlighted in Fig. 3 by green circles), taking into account changing satellite constellation. The receivers installed at different stations of the network had dissimilar characteristics; hence, the accuracy of the estimation of decoupled satellite corrections was somewhat degraded which caused overall reduction of the accuracy of the user solution.

Figure 4 [16] shows the decoupled code and carrier phase satellite corrections calculated for one of the satellites in the second version of the network solution. The graph shows jumps in phase corrections dt^j_L at the moments of change in the observation scenario and/or change of the matrix $\underset{M\times J_i}{\boldsymbol{STM}}$ associated with changes in the estimated linear combinations in the components of the vector $\underset{J_i\times 1}{\boldsymbol{dt}_L}$.

The results of the user solution applying the evaluated corrections shown in Figs. 2 and 4 were already presented and discussed in [1].

Conclusion

Algebraic principles of network solutions for PPP including ambiguity resolution of carrier phase measurements in GNSS with CDMA are considered.

Examples of determination of precise satellite corrections for the GPS Integer PPP are presented. Significant reduction in convergence time required to achieve high-precision positioning when using precise satellite corrections in comparison with the Float PPP is demonstrated.

As it follows from the expression for the vector $\underset{4\times 1}{\boldsymbol{dt}_{\rho.s.i}}$ that is included into (11), in the case of a timely determination of the ionosphere-free code clock offset $dT_{\rho.r.i}$ of a reference ground station relative to the GNSS time scale, there appears a possibility of rapid evaluation of the ionosphere-free code clocks offsets $dt^{j}_{\rho.i}\ \ j=\overline{1,J_i}$ of all the J_i SVs which are visible by ground stations at the i-th epoch. This data can be used to increase the accuracy of broadcast clock corrections transmitted in the SV navigation messages.

The experience gained so far shows the urgent need for the methods of receiver calibration to deal with the algorithms of determination of precise decoupled satellite corrections.

References

1. Povalyayev A.A., Podkorytov A.N., Nikitin S.A., Filimonova D.V. Algebraicheskiye osnovy obrabotki izmereniy pri vysokotochnom absolyutnom mestoopredelenii po signalam GNSS s kodovym razdeleniyem kanalov [Algebraic Principles of Precise Point Positioning with CDMA GNSS Signals]. *Raketno-kosmicheskoye priborostroyeniye i informatsionnyye sistemy* [Rocket-Space Device Engineering and Information Systems], 2019, Vol. 6, No. 1, pp. 4–16. (in Russian)
2. De Jonge P.J. A Processing Strategy for the Application of the GPS in Networks. *Netherlands Geodetic Commission*, *Publications on Geodesy*, 1998, 225 p.
3. Teunissen P.J.G. Generalised Inverses, Adjustment the Datum Problem and S-transformations. Preprint. *Delft University of Technology, Reports of the Department of Geodesy Mathematical and Physical Geodesy*, 1984.
4. Bisnath S., Collins P. Recent Developments in Precise Point Positioning. *Geomatica*, 2012, Vol. 66, No. 2, pp. 103–111.
5. Teunissen P.J.G., Khodabandeh A. Review and principles of PPP-RTK methods. *Journal of Geodesy*, 2014. Available at: https://www.researchgate.net/publication/268520221_Review_and_principles_of_PPP-RTK_methods (accessed April 29, 2019).
6. Cormen T., Leiserson Ch., Rivest R., Stein C. *Glava 23. Minimal'nyye ostovnyye derev'ya. Algoritmy: postroyeniye i analiz* [Chapter 23. Minimal spanning trees. Introduction to Algorithms]. Ed. Krasikov I.V., 2-nd edition. Moscow, Vil'yams, 2005, pp. 644–662, ISBN 5-8459-0857-4. (in Russian)
7. Haray F. *Graph theory*. Ed. Gavrilov G.P. 2-nd edition. Moscow, Editorial URSS, 2003, 296 p. (in Russian) (translated from English by Kozyreva V.P)
8. Lannes A., Gratton S. GNSS Networks in Algebraic Graph Theory. *Journal of Global Positioning Systems*, 2009, Vol. 8, No. 1, pp. 53–75.
9. Lannes A., Teunissen P.J.G. GNSS algebraic structures. *Journal of Geodesy*, 2011, Vol. 85, No. 5, pp. 272–290.
10. Sage A.P., Melse J.L. *Teoriya otsenivaniya i eye primeneniya v svyazi i upravlenii* [Estimation theory with application to communication and control]. Moscow, Svyaz', No. 6, 1976, 496 p. (in Russian)
11. Tikhonov V.I., Kharisov V.N. *Statisticheskiy analiz i sintez radiotekhnicheskikh ustroystv i sistem* [Statistical analysis and synthesis of radio engineering devices and systems]. Moscow, Radio i svyaz', 2004, 608 p. (in Russian)
12. Povalyayev A.A. *Sputnikovyye radionavigatsionnyye sistemy. Vremya, pokazaniya chasov, formirovaniye izmereniy i opredeleniye otnositel'nykh koordinat* [Satellite radio navigation systems. Time, indications of clocks, measurement formation, and determination of relative coordinates]. Moscow, Radiotekhnika, 2008, 328 p. (in Russian)
13. Hassibi A., Boyd S. Integer Parameter Estimation in Linear Models with Applications to GPS. *Proceedings of IEEE Conference on Decision and Control*, 1996, Vol. 3, pp. 3245–3251. Available at: http://www.stanford.edu/~boyd/int_est.html
14. De Jonge P.J. and Tiberius C.C.J.M. *The LAMBDA method for integer ambiguity estimation: implementation aspects*. August 1996. LGR. Delft Geodetic Computing Centre. Delft University

of Technology. Faculty of Geodetic Engineering. Thisseweg 11. 2629 JA Delft. The Netherlands. Available at: http://www.tudelft.nl/en/ceg/about-faculty/departments/geoscience-remote-sensing/ (accessed April 29, 2019)

15. Arkhangel'skiy V.A. Veytsel' V.A. Volkovskiy A.S. et al. *Radiosistemy i kompleksy upravleniya. Uchebnik* [Radio systems and control complexes. Manual]. Ed. Veytsel' V.A. Moscow, Vuzovskaya kniga, 2016, 574 p. (in Russian)
16. Podkorytov A.N. Povysheniye operativnosti vysokotochnogo absolyutnogo mestoopredeleniya v GNSS [Increase in operability of precise point positioning in GNSS]. *Sbornik trudov IX Vserossiyskoy nauchno-tekhnicheskoy konferentsii "Aktual'nyye problemy raketno-kosmicheskogo priborostroyeniya i informatsionnykh tekhnologiy" (5–7 Iyunya 2018)* [Proceedings of IX All-Russian scientific and technical conference "Current problems of rocket and space device engineering and information technologies" (June 5–7, 2018)]. Moscow, Joint Stock Company "Russian Space Systems", 2018, pp. 37–43. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 14–20

The Possibility of Independent Self-Positioning of a Vessel Based on Signals from the AIS System

S.V. Trusov, *Cand. Sci. (Engineering), trusov_sv@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
O.I. Baraboshkin, *baraboshkin_oi@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
A.M. Kuznetsov, *Cand. Sci. (Engineering), kuznetzov_am@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
S.A. Bobrovskiy, *bobrovsky_sa@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation

Abstract. The work aims to find an alternative method of calculating the position of a vessel in order to detect the deliberate local distortion of the GLONASS and GPS navigation signals (spoofing). A method for determining the coordinates of the vessel based on the combination of measurements of the AIS signal arrival angles from neighboring vessels, as well as the coordinates of the vessels contained in these signals, is proposed. A variant of the receiver circuit for implementing this method based on phase direction finding by a simple two-antenna system is considered. The possibility of determining the direction of arrival of a signal with an average error of 2^o is shown with the distance between the two antennas equal to λ (with each antenna $\lambda/4$ long). Based on the radio direction finding errors, the accuracy of the methods for determining the coordinates was estimated from two and three signal sources. The maps of the applicability of this method of location for the Sea of Okhotsk region have been constructed. It is shown that within a 2^o radio direction finding error, the error in determining the coordinates of the vessel will be no more than 4 km, depending on the number and relative position of the sources of AIS signals. The proposed solution can be used as part of the technical means of control of the fishing fleet.

Keywords: AIS, Fishing Industry Monitoring System, GNSS-spoofing, radio direction finding.

Introduction

Since 1999, in order to prevent the illegal, unreported and unregulated fishing of marine biological resources, the Russian Federation has been operating an industry-based system for monitoring aquatic biological resources, surveillance and control of the activities of fishing vessels (IMS). The system controls 3,500 vessels [1].

Control over the movement of vessels is carried out using technical control means (TCM), mandatory for installation on fishing vessels. The TCM, according to [2], should include a GLONASS /GPS signal receiver, a ship satellite communication terminal (Gonets, Inmarsat), and an AIS system transceiver. TCM should ensure the accuracy of determining the position of the vessel of not worse than 100 m and the transmission of this non-correctable information through the satellite terminal to the regional monitoring center (RMC) of the IMS [2].

The problematic point when using this method of monitoring the position of vessels is a possible distortion of the actual movements of the vessel due to local substitution of the navigation GLONASS/GPS signals (spoofing). Equipment for spoofing attacks is becoming more accessible and can become a common tool for poaching, since its application does not violate the requirements of the law [2]: the integrity of the TCM is not violated; the information is not corrected during the transmission.

The ARGOS system, the transmitters of which were previously used as the TCM, has an alternative possibility of determining the location based on Doppler measurements of the signal characteristics with an accuracy of 150-300 m, but since 2016, ARGOS transmitters cannot be used as the TCM on Russian ships. In this regard, to find an alternative way to determine and automatically register the coordinates of vessels is an urgent problem.

In this paper, we consider an approach to determining the coordinates of a vessel using special receiving equipment for AIS signals, which is proposed to be included in the TCM. The main difference between this equipment and a conventional AIS signal receiver is an antenna system that it can determine the angles of arrival of signals. This information, in combination with data on the location of the sources of these signals, makes it possible to determine the position of the vessel. The problems of accuracy of positioning of the vessel depending on a number of factors are considered.

Fundamental possibility of determining the ship's location using AIS signals

The proposed approach to the determination of coordinates is based on the following provisions:

1. Obtaining the coordinates of the signal sources and the direction of arrival of these signals make it possible to determine the position of the receiver and, accordingly, the vessel.

2. Around a fishing vessel in the radio visibility zone of the AIS receiver (20–40 miles), there is at least one AIS signal source (vessel or coast station) transmitting its coordinates (AIS messages of types 1, 2, 3, 4, 11, 18, 21, 27) and not vulnerable to spoofing.

3. A number of radio direction finding methods are known [3], that make it possible to determine the directions of arrival of signals from each of the vessels.

The source data for choosing the method of determining the location of the target vessel can be: information about the speed, course of the vessel and the number of vessels in the radio visibility zone. Depending on the set of available source data, the following known calculation methods can be used to determine the location of the vessel:

1. If data on the speed and course of a given vessel ~~and~~ registration of angles of arrival of AIS signals from a single source are available~~,~~ it is possible to determine the location of this vessel by the cruise-bearing method, by processing measurements at different times [4];`

2. if data on the course and angles of arrival of AIS signals from two sources are available, the bearing method by two landmarks can be used [4];

3. when registering the angles of arrival of AIS signals from three sources, it is possible determine the location using the single resection method [5];

4. when registering the angles of arrival of AIS signals from four sources or more, the multiple resection method is used [5].

In this paper, we consider the characteristics of the second and third methods, since the first method requires the vessel being positioned to be in motion while the source of the AIS signal remains immobile, which, according to expert estimates, is an unlikely situation, and the fourth method requires a large number of AIS signal sources.

In the general case, the accuracy of the methods is determined by the accuracy of the coordinates of the AIS signal sources and the accuracy of determining the angles

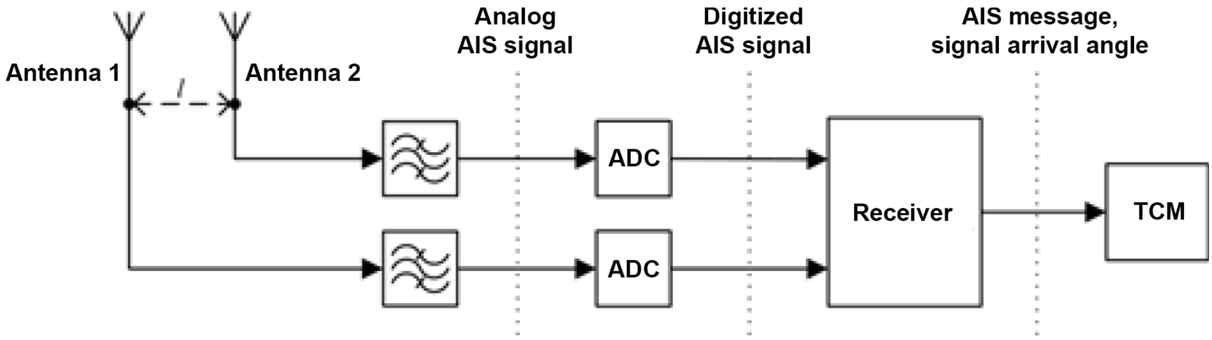


Fig. 1. Diagram of the AIS receiver with the function of determining the angle of arrival of the AIS signal

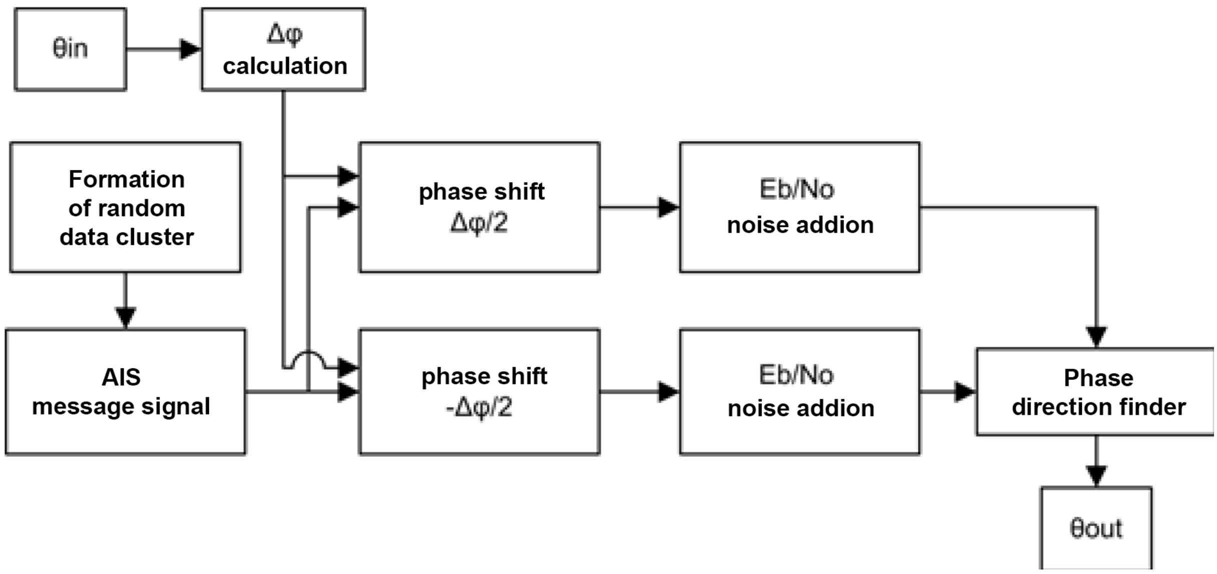


Fig. 2. The block diagram of the model of the AIS signal direction finding system

of arrival of these signals. The accuracy of the coordinates of the sources is 10–100 m and is not considered in this paper.

The accuracy of determining the angles of arrival of AIS signals depends on the signal registration system and the relative location of the sources of AIS signals. These factors are discussed in more detail in the following sections.

System for determining the angles of arrival of AIS signals

According to [3], methods for determining the angles of arrival of an AIS signal are divided into amplitude, phase, and complex methods. For a method using the amplitude characteristics of the signal (the amplitude and complex methods), an antenna system consisting of antennas with a non-circular radiation pattern is required. With a wavelength of λ = 1.85 m at AIS frequencies, such an antenna system will be very cumbersome, difficult to install on a vessel, and quite expensive. In this connection, it is preferable to use the phase method of direction finding.

In [3], a phase direction finding method using two antennas is presented. Antennas are located at a distance from each other, called the base. Each of the antennas can have a circular radiation pattern and represent a $\lambda/4$ long pin. The accuracy of direction finding and the complexity of placement on a vessel depend on the length of the base. The larger the base, the more accurately it is possible to determine the direction, but the more difficult is the installation. The disadvantage of the method under consideration is the ambiguity in determining the direction — for a base length $l<\lambda/2$ there are two such directions, for $l<\lambda/2<l<\lambda$ there are four directions, etc. To solve the problem of ***monitoring*** the information received from another vessel, ambiguity does not play a significant role, because AIS information from both vessels resolves this ambiguity.

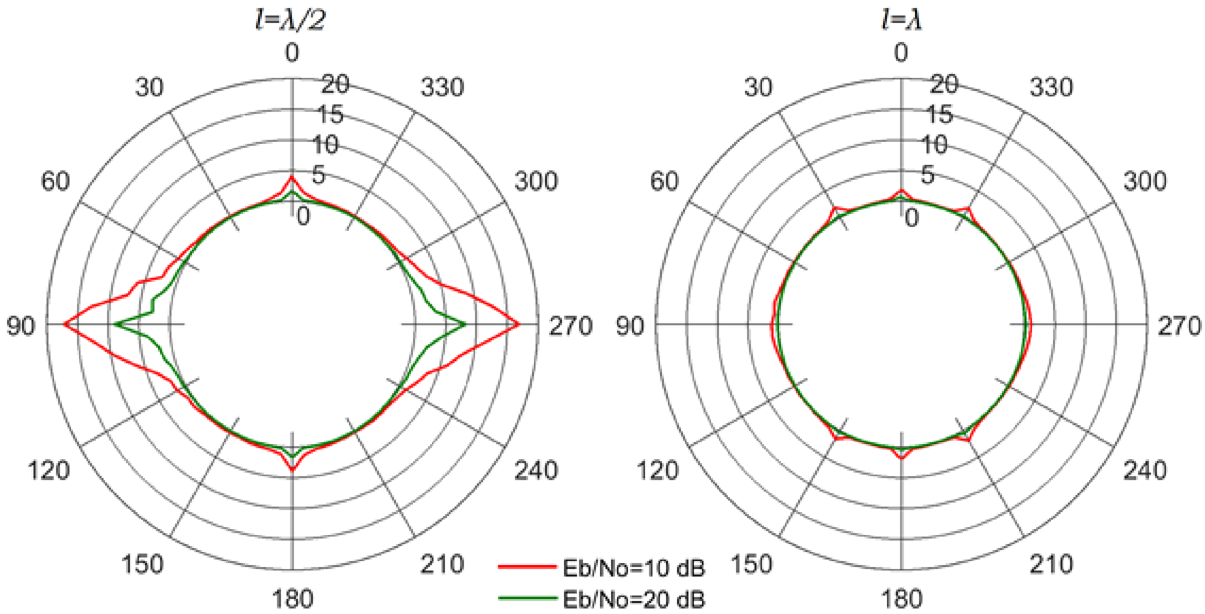


Fig. 3. The average error of the calculated direction of arrival of the signal depending on the actual direction for the AIS signal

Thus, the following scheme of an AIS receiver combined with a radio direction finding system is proposed. The signals from two antennas located at a distance *l* from each other are transmitted to the RF path of the receiver, where they are independently filtered and digitized by two analog-to-digital converters (ADCs) with a common clock generator. One of the signals is processed with the same method as in a conventional AIS receiver, and the beginning and end of each AIS message are determined. This information is transmitted to a receiver module containing a phase direction finder. This module simultaneously processes two signals containing the same AIS message, demodulates the message and finds the angle (Fig. 1).

Calculation of the accuracy of determining the angles of arrival of the AIS signal

The accuracy of determining the angles of arrival of AIS signals when using the proposed AIS receiver circuit is influenced by many factors: the length of the base of the antenna system, noise characteristics of the receiver, the format of the AIS messages, the distance to the transmitter, vessel pitching, errors in the length of the base of the antenna system (temperature fluctuations, installation accuracy), the error in determining the course of the vessel, etc. This section discusses the impact of the first three factors.

To calculate the direction finding accuracy, numerical modeling was used according to the following diagram (Fig. 2). An AIS signal is generated with GMSK modulation, NRZI coding, and a data transfer rate of 9600 bps, containing a random data block [6]. This signal is duplicated, each signal is independently summed with white noise, specified by the ratio of bit energy to power spectral density *Eb/No*. The lower threshold *Eb/No*, at which AIS messages can be received, is defined in [6] and amounts to 10 dB. Next, the first signal is shifted by $\Delta\varphi/2$ and the second is shifted by $-\Delta\varphi/2$. $\Delta\varphi$ is calculated based on a predetermined angle of arrival of the radio signal $\theta_{in}=[-\pi/2..\pi/2]$ according to the formula

$$\Delta\varphi=2\pi l/\lambda\cdot\sin(\theta_{in}).$$

The phase direction finder is constructed according to the design with sum-difference processing [3]. The signal at which the simulation of the angle of arrival of the radio signal θ_{out} is calculated during the modeling, has a length of 224 bits and consists of a training sequence (24 bits), a start indicator (8 bits), a data block (184 bits) and an AIS message end indicator (8 bits).

The simulation results for determining θ_{out} for 75 thousand random messages with a base length $l = \lambda/2$ and $l = \lambda$, $Eb/N0 = 10$ dB and 20 dB for each θ_{in} with a step of 5° are presented in Fig. 3. The graphs show the average errors ζ_{out} of determining the angle θ_{out} depending on θ_{in}.

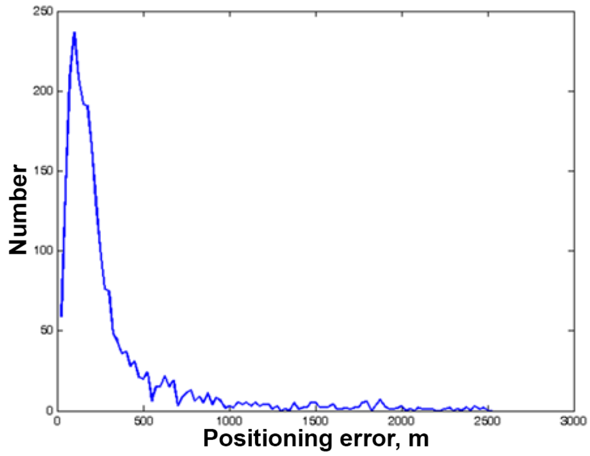


Fig.4. The dependence of the number of measurements on the error in determining the coordinates with a random error of the angle of arrival of the AIS signal within 2 for 10°4 measurements

With the base length $l = \lambda/2$, the direction finding accuracy strongly depends on the direction to the signal source - the errors increase strongly when θ_{in} approaches 90° and 270°. Even with a high signal to noise ratio, errors make it impossible to confidently distinguish between different directions of arrival of AIS signals at $70° < \theta_{in} < 110^{o}$ and $250° < \theta_{in} < 290^{o}$. For other directions with $Eb/N0 = 10$ dB, the average error is less than 5º, and for $Eb/N0 = 20$ dB - less than 2º.

Using the signal base $l = \lambda$ gives a smaller error and more stable characteristics of the direction finding system in the entire range of angles. For $Eb/N0 = 10$ dB, the average error is less than 2º, and for $Eb/N0$ = 20 dB it is less than 1º.

The results obtained show that an antenna system with a base length $l = \lambda$ is more preferable for constructing a vessel position control system.

Modeling errors in determining coordinates depending on the accuracy of determining the angles of arrival of a signal

Locating a ship using three AIS signals. To assess the influence of the accuracy of determining the angles of arrival of signals on the determination of the coordinates of the vessel, a numerical simulation of the determination of coordinates by the method of single resection was performed. The identified vessel was in a circle 36 km (~ 20 miles) in diameter. This, on average, corresponds to the maximum distance for the propagation of AIS signals between ships. Three vessels with known coordinates were randomly placed in this area, arrival angles were

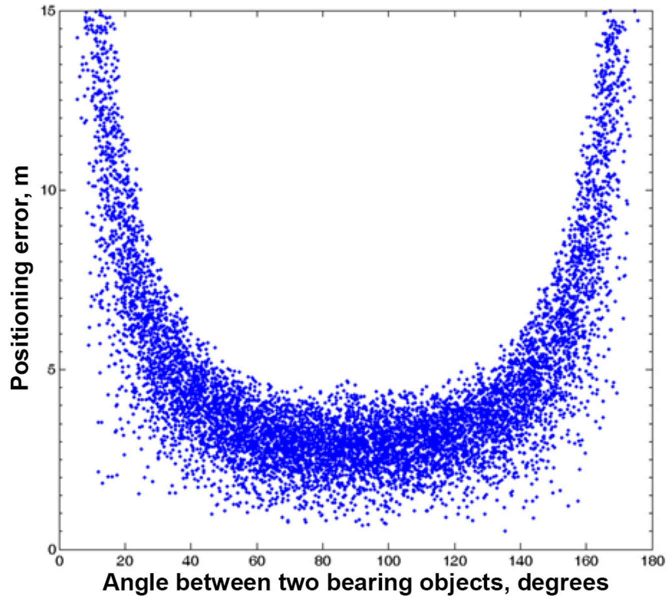


Fig. 5. Dependence of the error in determining the position of the vessel by the bearing of two reference points on the angle between the objects with an error in measuring directions 2°

determined, a random error of up to 2° was introduced into each of the angles, and a resection problem was solved [5]. After the calculation, the error in determining the coordinates was determined as the distance between the known and calculated points.

The simulation results of 10^4 cases of ship distribution (Fig. 4) indicate that registration of three AIS signal sources with an error of up to 2° allows in 85 % of cases to get an error in determining the location of the vessel of up to 2 km.

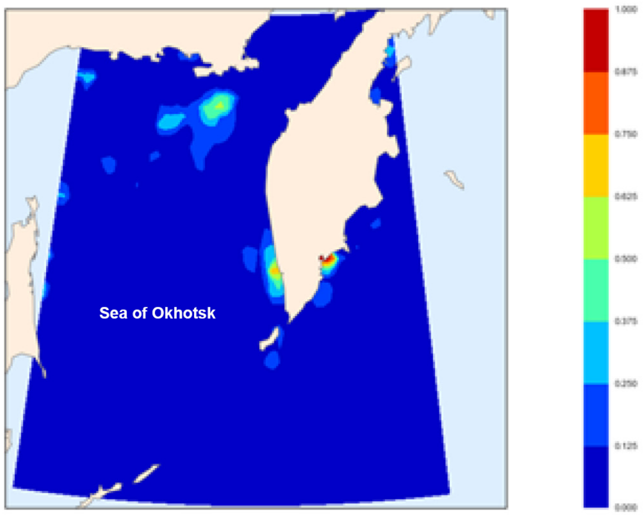
The error in determining the coordinates decreases with increasing accuracy of determining the angles of arrival and a decrease in the average distance between the signal sources and the receiver.

Locating a ship using two AIS signals.

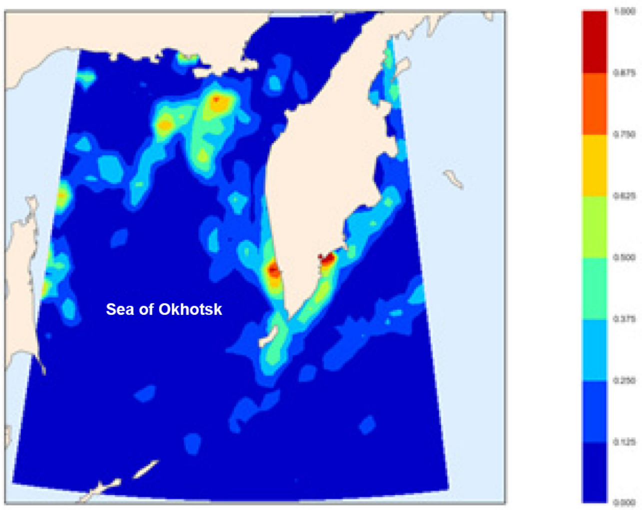
The positioning of the vessel is possible if there are two AIS signals and the known course of the vessel. In this case, the accuracy of determining the location depends on the accuracy of determining the course and the accuracy of determining the angles of arrival of AIS signals. In this method, reference points (AIS signal sources) that are 30–150° apart from each other can be used. The dependence of the error in determining the location on the angle between the reference points (with an accuracy of 2(°), shown in Fig. 5, shows that for the working range of angles, the determination error on average is 4 km.

This method is less accurate than resection, but the conditions for its use are more common than these required for resection (see Fig. 6).

Detection of a false course on one source of AIS signal. If there is one AIS signal, it is possible to check the correctness of the coordinates and course of the vessel, determined by a GLONASS/GPS receiver. To do this, based on the coordinates of the transmitter vessel contained in the AIS message, it is enough to calculate the angle of arrival of the signal and compare it with the measured angle. If the measured angle does not coincide with the calculated angle within the allowed error, then the coordinates of one of the two vessels are incorrect. It is possible to unambiguously identify a vessel with incorrect coordinates when other vessels appear in the vicinity: on a vessel with incorrect coordinates, there will be a mismatch between the calculated and determined signal arrival angles for other vessels.



a



b

Fig. 6. The distribution map of the probability of the presence of AIS signals from N vessels in the Sea of Okhotsk: *a*) N=3, *b*) N=2:

Estimation of the probability of the presence of other vessels in the vicinity of an arbitrary vessel

To assess the likelihood of the presence of two or more vessels in the vicinity of an arbitrary vessel, which are necessary for applying the proposed method, data were processed on the location of vessels in the waters of the Okhotsk and Bering Seas within latitude limitations of 45 ÷ 60°n. and in longitude 144 ÷ 164°e. The data for May – July 2016 were used [7]. The calculation was made for the grid in increments of 0.33°. For each grid point, the ratio of days on which at least N vessels were registered within a radius of 20 miles from this point to the total number of observation days was calculated; N = [2,3]. The calculation did not use signals from coastal AIS stations with a range of about 40 miles.

The calculation results (Fig. 6) show that the 100% probability of recording AIS signals from three vessels at the same time in the region under consideration exists only in the Petropavlovsk-Kamchatsky region. High probability (> 50%) is seen in the main fishing areas of the North Okhotsk and West Kamchatka subzones of the Sea of Okhotsk (Fig. 6, a). In the remaining water areas, the probability of the presence of three AIS signals from different vessels is less than 12%.

A similar calculation for the probability of receiving AIS signals from two vessels (Fig. 6, *b*) shows that the area of regions with a corresponding probability increases by an average of 4.5 times. Areas with a probability of more than 0.33 correspond to the main fishing areas of the Okhotsk Sea basin [8].

Conclusion

The results of the studies confirmed the authors' assumption that the use of AIS signals from other vessels and a special receiving system make it possible to determine the location of the vessel under a number of conditions.

Numerical simulation of the receiving device showed that a system with two antennas spaced by l allows the phase-direction finding method to determine the angle of arrival of the AIS signal with an accuracy of 1–2°. The directions of arrival of AIS signals, measured with such errors, make it possible, in the presence of two signals, to determine the position of the vessel with an accuracy of 2–4 km, and within 1–2 km in the presence of three signals. In areas of intensive fishing, the probability of determining the position of the vessel by the signals of two vessels is more than 33 %.

In cases where there is only one AIS signal, the false course detection method can be used. All methods should be implemented in a single device, be applied depending on the situation and give to the TCM data on the position of the vessel within appropriate accuracy.

The proposed method for determining the coordinates of the vessel can be used in the design of an improved TCM for use in fiscal systems for monitoring the location of vessels, in particular, for detecting local substitution of the GNSS navigation field, as well as for determining the location of a vessel when the AIS ship transponder is turned off.

References

1. Pyrkov V.N., Solodilov A.V., Degay A.Yu. Otraslevaya sistema monitoringa vodnykh biologicheskikh resursov, nablyudeniya i kontrolya za deyatel'nost'yu sudov rybopromyslovogo flota (OSM) [Sectoral system for monitoring aquatic biological resources, monitoring and controlling the activities of vessels of the fishing fleet (OSM)]. Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa [Modern problems of remote sensing of the Earth from space]. 2015, Vol. 12, No. 5. pp. 251–262. (in Russian)
2. Order of the Ministry of Agriculture of the Russian Federation of July 13, 2016 No. 294 On approval of the Procedure for equipping vessels with technical means of control and their types. (not in effect). Available at: http://www.garant.ru/products/ipo/prime/doc/71439614/ (Last accessed: 14/05/2019). (in Russian)
3. Timofeev V.A. Amplitudnye i fazovye metody opredeleniya uglovogo polozheniya istochnika elektromagnitnykh voln: metod. ukazaniya po vypolneniyu laboratornoy raboty [Amplitude and phase methods for determining the angular position of a source of electromagnetic waves: method, instructions for performing laboratory work]. Yaroslavl State University, 2006, 56 p. (in Russian)
4. Rul'kov D.I. Navigatsiya i lotsiya [Navigation and pilot charts]. Moscow, Transport, 1973, 232 p. (in Russian)
5. Kusov V.S. Geodeziya i osnovy aerometodov [Geodesy and fundamentals of airborne methods]. Moscow, Publishing House of Moscow University, 1995, 160 p. (in Russian)
6. MSE-R M.1371-5 (02/2014) Technical characteristics of an automatic identification system using time-division multiple access in the VHF maritime mobile service band. Available at: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.1371-5-201402-I!!PDF-R.pdf (Last accessed: 05/14/2019). (in Russian)
7. Kuznetsov A.M., Trusov S.V., Baraboshkin O.I., Bobrovskiy S.A., Romanov A.A., Romanov A.A. Analiz rezul'tatov, poluchennykh za tri goda raboty apparatury monitoringa sudov BRK AIS, razmeshchennoy na KA «Resurs-P» № 2 [Analysis of the Results Obtained over Three Years of Operation of AIS Vessel Monitoring Equipment Based on the Resurs-P No. 2 Spacecraft]. Raketno-kosmicheskoe priborostroenie i informatsionnye sistemy [Rocket-Space Device Engineering and Information Systems]. 2018. Vol. 5, No. 4. pp. 80–87. (in Russian)
8. Obzor snyurrevodnogo promysla v period s 5 po 11 marta 2012 goda. [Review of the snurrevaad fishing in the period from 5 to 11 March 2012]. Available at: http://www.kamniro.ru/presscenter/obzory_promysla/promysel/obzory_promysla3 (Last acessed: 05/14/2019). (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 21–31

Simulation of the Low-Power Part of Transceiving Modules of the C-Band for AESA of the MMIC Form

V.A. Budnyaev*, engineer, VABudnyaev@sevsu.ru, Engineering Center of Micro- and Nanoelectronics Devices, Sevastopol State University, Sevastopol, Russian Federation*
I.V. Kravchenko*, engineer, IVKravchenko@sevsu.ru, Engineering Center of Micro- and Nanoelectronics Devices, Sevastopol State University, Sevastopol, Russian Federation*
D.A. Snegur*, postgraduate student, engineer, DASnegur@sevsu.ru, Engineering Center of Micro- and Nanoelectronics Devices, Sevastopol State University, Sevastopol, Russian Federation*
V.V. Tumko*, engineer, VVTumko@sevsu.ru, Engineering Center of Micro- and Nanoelectronics Devices, Sevastopol State University, Sevastopol, Russian Federation*
I.F. Filippov*, postgraduate student, engineer, IFFilippov@sevsu.ru, Engineering Center of Micro- and Nanoelectronics Devices, Sevastopol State University, Sevastopol, Russian Federation*
V.V. Vertegel*, Cand. Sci. (Engineering), director, VVVertegel@sevsu.ru, Engineering Center of Micro- and Nanoelectronics Devices, Sevastopol State University, Sevastopol, Russian Federation*

Abstract. The paper presents the development results of electric circuits of functional units and layout of the integrated transceiver module of active electronically scanned arrays (AESA) of the C-band based on the 0.18 μm SiGe BiCMOS technology. The integrated circuit consists of control units for the amplitude and phase of a microwave signal, the reception and transmission modes switchers, amplifiers, digital control unit, and units for parameter temperature corrections. The root mean square (RMS) deviation in amplitude does not exceed 0.4 dB. The RMS deviation in phase is below 4.2 degrees. The noise factor value of the reception path does not exceed 5.2 dB. The output power in the compression point 1 dB in transmission mode is 8.5 dBm. The power consumption in the reception and transmission modes is below 195 mW and 365 mW, respectively. The layout area is 6 mm2.

Keywords: monolithic microwave integrated circuit (MMIC), transceiver module, active electronically scanned array (AESA), SiGe, BiCMOS, C-band.

Introduction

Active electronically scanned arrays (AESA) are widely used in radar systems, in radio navigation, electronic warfare, and telecommunication systems for special purposes. One of the key areas of development of domestic electronics is the increase of the production of electronic components, both dual-purpose and for civil use. AESA systems are promising in terms of their possible future implementation in fifth-generation telecommunication systems, satellite communication systems, etc. Currently, transceiving modules (TM) of the AESA are built, as a rule, based on monolithic microwave integrated circuits (MMIC) of various technology groups. The worldwide sales of MICs for AESA systems in recent years have demonstrated significant growth rates. Yet, despite the colossal functionality of such systems, they have several disadvantages. The key demerit limiting their widespread use in the civilian sector is the extremely high cost of MICs for TMs, which are manufactured primarily using GaAs-technology.

A reduction of the cost of TMs may be achieved via implementing integrated transceivers (Core Chip) that are manufactured with the use of silicon (Si) or silicon-germanium (SiGe) technologies. The implementation of the SiGe BiCMOS-technological process is capable of solving the problem of integrating the digital and analog parts of the transceiving modules, which will positively affect the final cost of the system while maintaining acceptable characteristics of TMs for civilian applications. The article presents the results of developing electric circuits and layouts of functional units of the integrated C-band TM based on 0.18 μm SiGe BiCMOS-technology. The relevance of the present study lies in the achievement of such important characteristics for AESA systems as miniaturization, stability of parameters under the influence of external climatic factors as well as reduction of series production costs.

MIC block diagram

Based on an analytical overview of mass-produced products [1–3] and literary sources [4–9] containing results of studying AESA TMs, a block diagram of an intellectual property (complex functional, IP) MIC was developed and is given in Fig. 1.

Unlike the typical block diagram of IP MIC transceiving modules, it contains a digital temperature sensor along with a correcting phase shifter and attenuator, which help to reduce the errors in setting the amplitude and signal phases conditioned by changes in temperature.

The controlled phase shifter CPS-1 and controlled attenuator CATT-1 are common units of the reception/transmission path. Switching of operation modes (reception/transmission) of the IP MIC is performed with the help of the SW-1, SW-2 and SW-3 on-off switches.

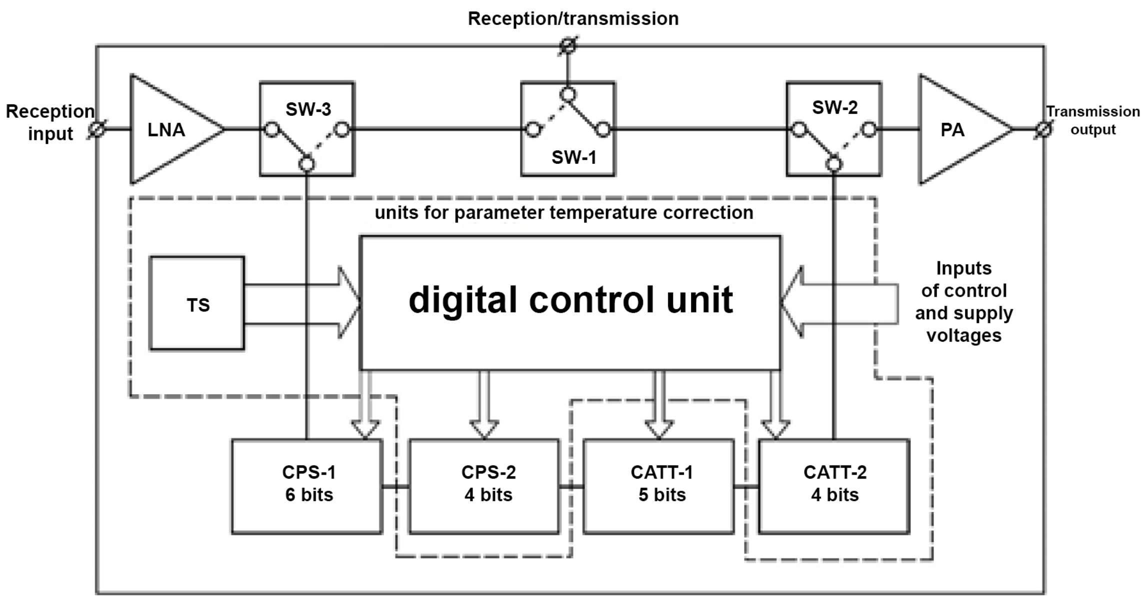


Fig. 1. IP MIC block diagram.

The low-noise amplifier (LNA) ensures an increase in the signal level at the input of the reception path to the required values, having the lowest possible value of the noise factor. The power amplifier (PA) provides the required level of the microwave signal at the output of the transmission path.

The integrated temperature sensor (TS) generates a temperature-dependent voltage and converts it into digital codes. Based on these codes, the digital control unit (DCU) corrects the phase shift and attenuation with the use of the CPS-2 phase shifter and CATT-2 attenuator. For data exchange between the IP MIC and the control microcontroller, a serial peripheral interface is included into the DCU. With the help of external digital signals, the DCU ensures switching between operation modes, controls the parameters of the CPS-1 phase shifter and CATT-1 attenuator.

The block diagram does not show the functional unit of the reference voltage source (RVS), with the help of which the given operation mode is ensured by the direct current of all of the units of the IP MIC.

MIC functional nodes and their characteristics

Low-noise amplifier

A simplified electrical diagram of the LNA is given in Fig. 2.

The VT1 transistor is connected with the common base in the circuit for ensuring coordination of the input impedance of the first stage of the LNA with a resistance of 50 Ω in a wide bandwidth. A cascode amplifier based on VT2 and VT3 field-effect transistors, due to the partial elimination of the influence of the Miller effect, has a high voltage gain without narrowing the operating frequency band. The low-pass filter in the base circuit of the transistor VT1 is necessary to prevent the ac component of the voltage from entering the output of the reference voltage source. Thus, stabilization of the operating point of the VT1 transistor, as well as the reduction of the noise factor of the stage are achieved.

Figures 3 – 4 give the results of LNA general circuit simulation with account for parasitic elements of the layout.

The gain in the frequency range 4– 6 GHz is 21-22 dB and the noise factor does not exceed 4.7 dB. The value of the return losses at the input in the operating frequency range exceeds 15 dB.

The compression point of 1 dB at the input (IP_{1dB}) in the operating frequency range exceeds 22 dBm. The input third-order intermodulation intercept point (IIP3) is at least 12 dBm.

Controlled phase shifter

The CPS-1 includes (Fig. 5):

- a balancing device (BD; balancer), which performs the conversion of an asymmetric input signal into a symmetric (balanced) output signal;

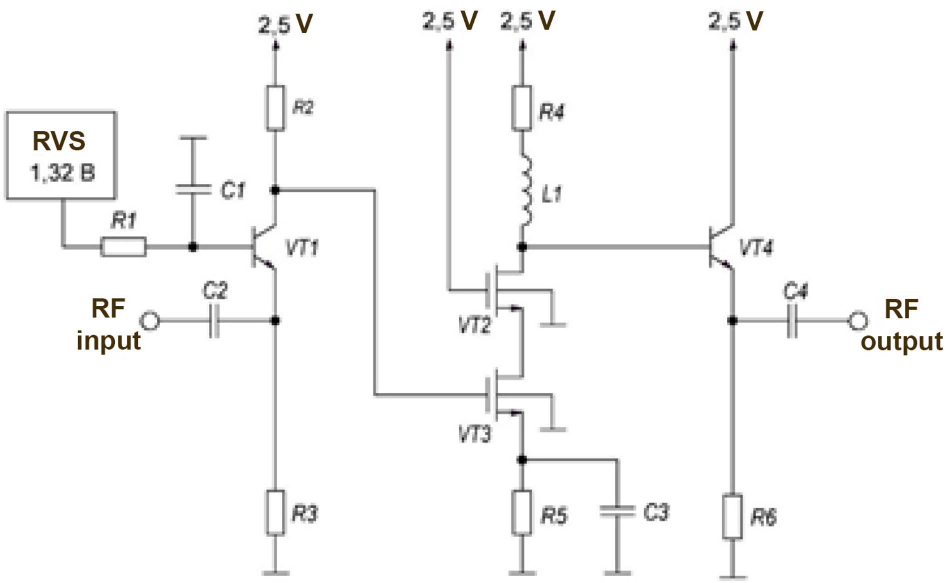


Fig. 2. Electric circuit of the LNA.

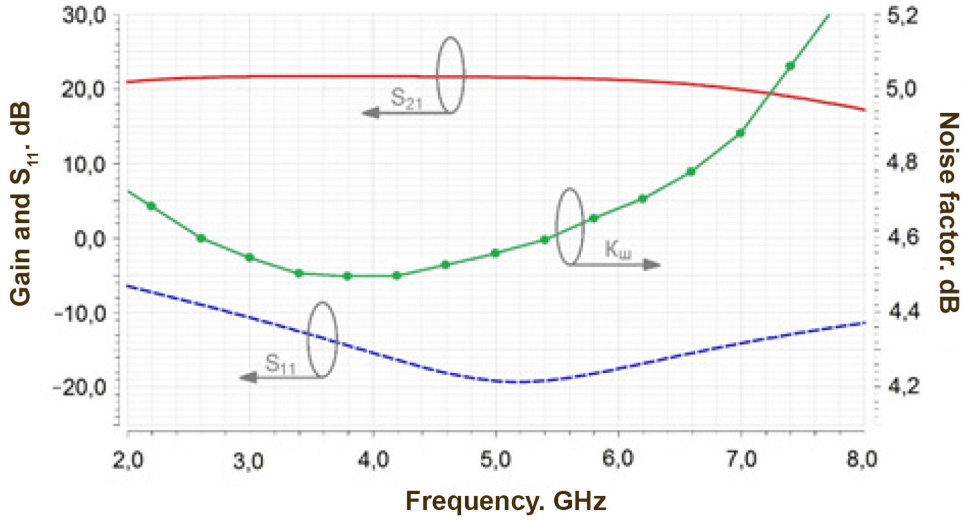


Fig. 3. Frequency dependences of small-signal LNA parameters.

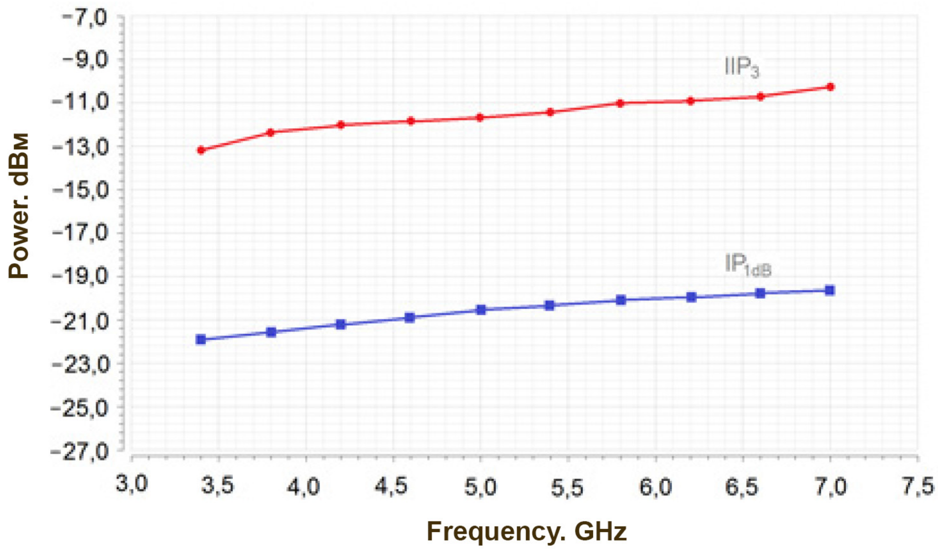


Fig. 4. Dependence of IP_{1dB} and IIP_3 on the frequency.

- a polyphase filter (PPF), which divides the balanced signal into orthogonal components;
- a dual adder based on Gilbert cells (DA) responsible for the weighted summation of the components;
- a controlled attenuator (ATT);
- a current digital-to-analog converter (DAC), which is a circuit of the driver of analog control signals by an adder (summator);
- a buffer amplifier (BA) at the output of the circuit, which is necessary for the compensation of the attenuation introduced by the indicated CPS-1.

The key feature of the given phase shifter is the implementation of an additional stage of a controlled attenuator, which allows reducing the amplitude error introduced when switching phase samples. The use of frequency and temperature–dependent feedback allows us to adjust the amplitude-frequency and temperature characteristics of the phase shifter.

The adder and attenuator are controlled simultaneously using circuits of current DACs, which are a binary current matrix controlled by transistors in key mode. Furthermore, the inputs of every current DAC are connected in parallel. Thus, the adjustment (correction) does not require additional control bits [10]. Fig. 6 demonstrates the frequency dependences of the transmission factor and phase states of the CPS-1 when switching phase samples.

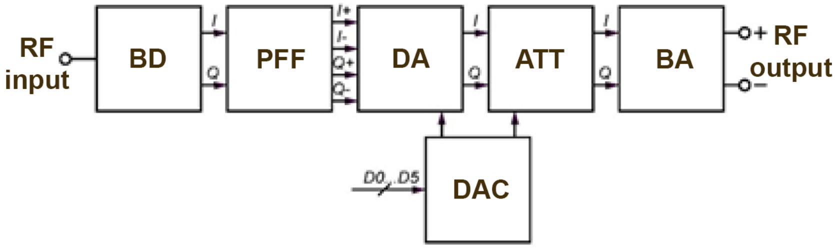
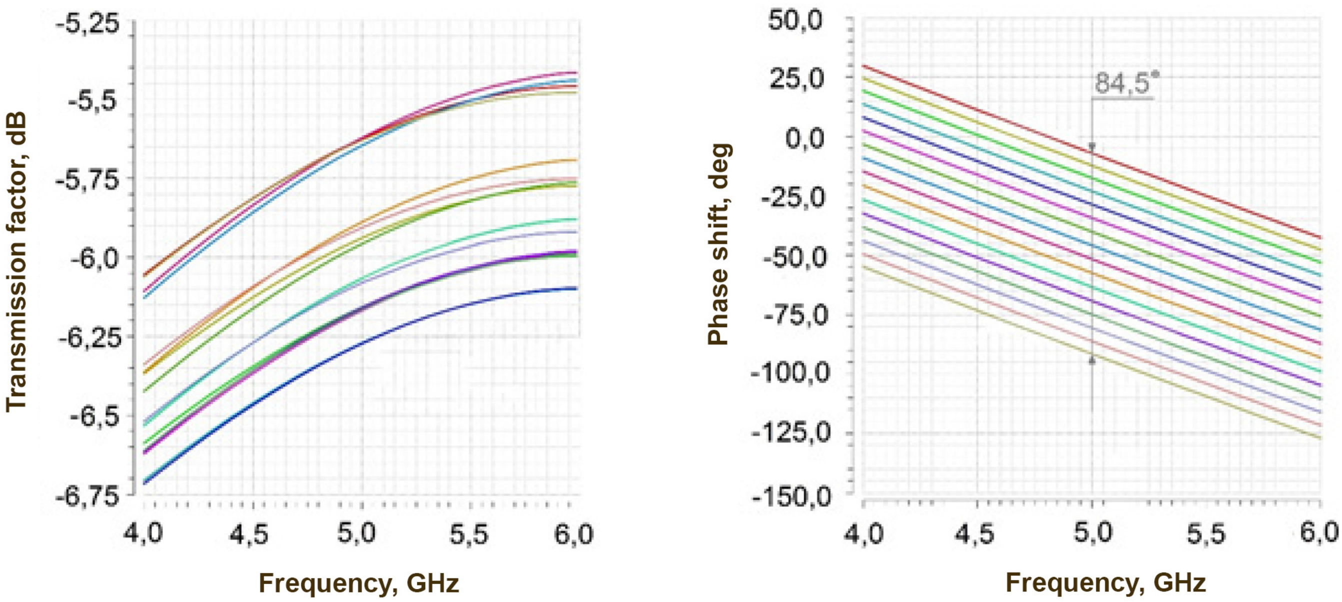


Fig. 5. CPS-1 block diagram.



a) b)

Fig. 6. Amplitude (a) and phase (b) frequency characteristics of the CPS-1 during phase adjustment within one quadrant.

Table 1. Parameters of the developed phase shifters

Parameter, unit	CPS-1	CPS-2
Range of operating frequencies, GHz	4—6	
Range of operating temperatures, °C	−60—85	
Range of phase adjustment, deg	360	8
Adjustment step size, deg	5.625	0.5
Phase-setting error, deg	<1.005	<0.25
Transmission factor, dB	–6	5.9
Amplitude error, dB	<0.703	<0.1
Supply voltage, V	2.5; 5	2.5; 5
Current consumption, mA	<35	<27

The circuit of the correcting phase shifter CPS-2 is analogous to the circuit of the main CPS-1. The correcting phase shifter does not include a controlled attenuator circuit, since it has a relatively small value of the amplitude error within the limit of the range of operating frequencies and temperatures.

The key parameters of the phase shifter obtained as a result of their simulation are set forth in Table 1.

Controlled attenuator

When developing an IP MIC based on silicon-germanium technologies for controlling the amplitude of the output signal, it is reasonable to use an active controlled attenuator (amplifier with a variable transmission factor) because it has a relatively low amplitude error, permits to eliminate the insertion losses and occupies a relatively small area on the chip. The block diagram of the developed CATT-1 is given in Fig. 7.

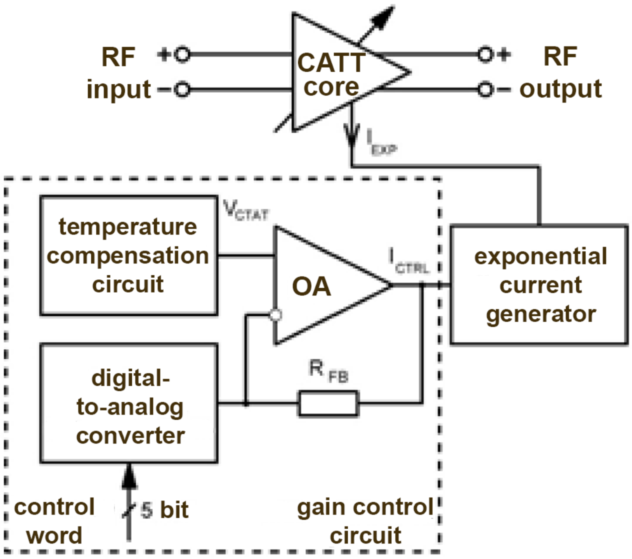


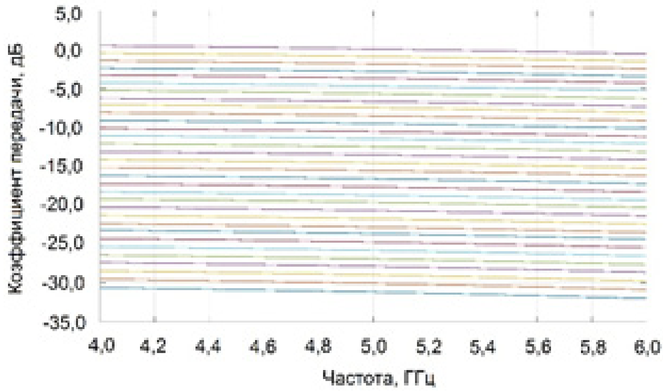
Fig. 7. CATT-1 block diagram.

The key element of the circuit is the CATT core, which changes the amplitude of the differential microwave signal at the input of the attenuator. The core consists of a classic differential stage with a pair of oppositely connected transistors with “a floating emitter” [11]. Such a connection allows us to widen the adjustment range of the attenuation factor of the circuit. Emitter followers are implemented at the output of the circuit in order to match the CATT-1 with the next stage of TMs.

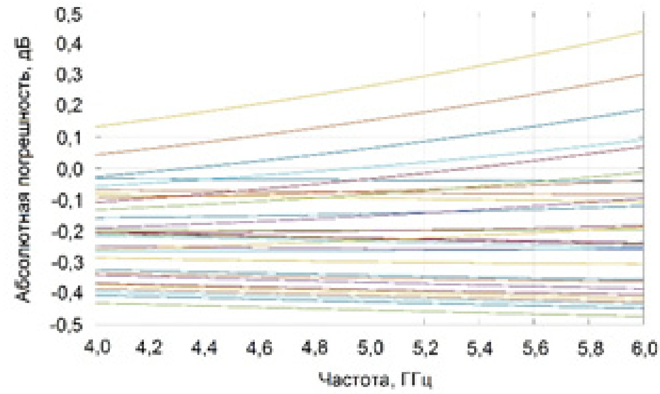
To ensure a uniform step size in changing the transmission factor of the attenuator, a nonlinear current generator, which uses an exponential form of the transmission characteristic of the bipolar transistor, is introduced.

The control circuit generates the control signal for the exponential current generator and includes the DAC, a temperature compensation circuit and an operational amplifier (OA). The voltage at the output of the temperature compensation circuit is inversely proportional to the temperature (V_{STAT}), which allows us to compensate the increase in the current via the differential stage of the core and ensure the temperature stability of the attenuator. In Fig. 8a, the frequency dependences of the attenuator gain are presented for various states of the control code (from 0 to 31). Fig. 8b shows the frequency dependences of the absolute error of the attenuator.

The range of variation of the transmission factor is 31.3 dB on the central frequency of 5 GHz. In this case, the



a)



b)

Fig. 8. Frequency dependences of the transmission factor (a) and the absolute amplitude error (b) of CATT-1.

non-uniformity of the transmission factor at maximum attenuation is 1.8 dB and at minimum – is 1.08 dB. Across the entire range of operating frequencies, the absolute error of setting the transmission factor does not exceed 0.48 dB and the root mean square deviation is 0.178 dB.

The functional unit of the correcting CATT-2, in terms of the general circuit solution, copies the main CATT-1. The difference lies in the magnitude of the reference currents of the DAC. In the DAC of the CATT-2, they are reduced in order to obtain a smaller adjustment range of the circuit transmission factor. The main parameters of the developed main and correcting attenuators are given in Table 2.

Microwave switch

In the developed circuit of a single-pole switch in two directions (Fig. 9), to increase the isolation, parallel switching of shunt field-effect transistors is used.

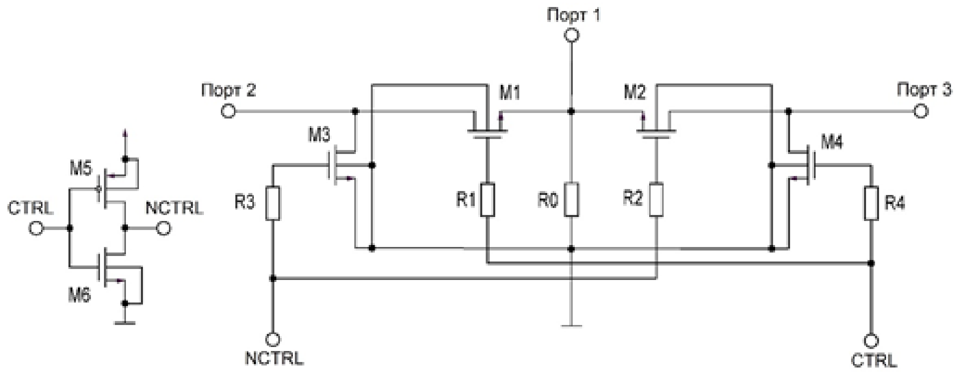
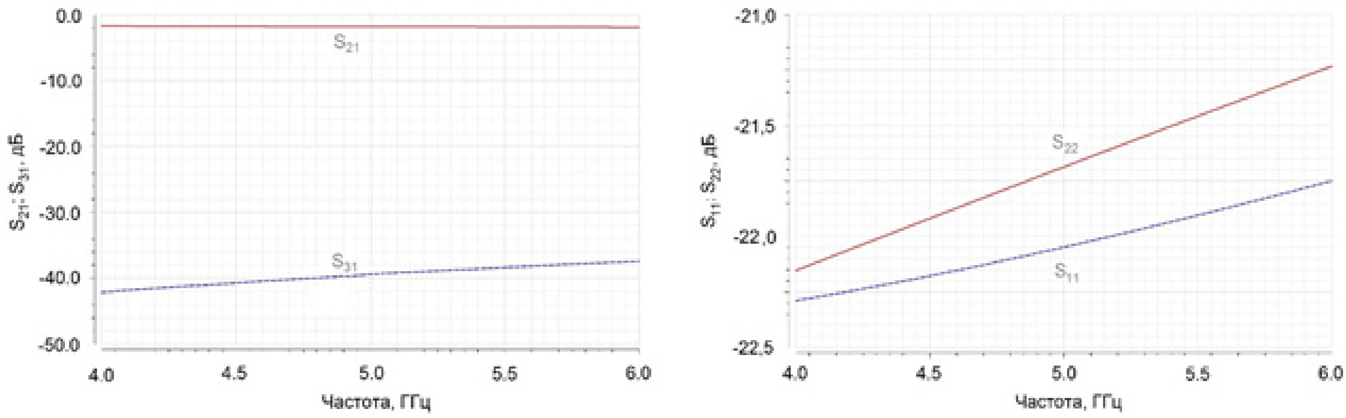


Fig. 9. Electric circuit of the microwave switch.



a) b)

Fig. 10. Frequency dependences of parameters S_{21}, S_{31} (a) and S_{11}, S_{22} (b) of microwave switches.

Table 2. Parameters of the developed attenuators

Parameter, unit	CATT-1	CATT-2
Range of operating frequencies, GHz	4—6	
Operating temperature range, °C	−60—85	
Range of insertion losses, dB	−31—0	−3.56—0.21
Attenuation step size, dB	1	0.25
Amplitude error, dB	<0.48	<0.028
Phase error, deg.	<5.3	<0.37
Compression point of 1 dB at output, dBm	−30	−27
HDF at max. signal amplitude, %	0.72	0.8
Noise factor, dB	<18	
Supply voltage, V	2.5; 5	
Power consumption, mW	17.3	15.45
Footprint, mm^2	0.009	

The M1 and M4 transistors are controlled in-phase by the CTRL voltage; the NCTRL voltage controls transistors M2 and M3. Thus, when the M4 is in the ON-state, the part of the signal, which passed the closed M2 transistor, is shunted to the ground. In the other channel of the switch, the M3 transistor will be in the OFF-state and practically will not affect signal passage. The M5 and M6 transistors generate an inverse control signal for the NCTRL.

The transmission characteristics of the open (S21) and closed (S31) channels of the switch are given in Fig. 10a. Insertion losses in the operating frequency band do not exceed 1.9 dB, and isolation is more than 37 dB. The return losses at the input and output exceed 21 dB (Fig. 10b).

Power amplifier

A simplified electric circuit of the power amplifier (PA) is given in Fig. 11. The PA provides the required signal power level at the input of the IP MIC under development with a minimal level of non-linear distortions.

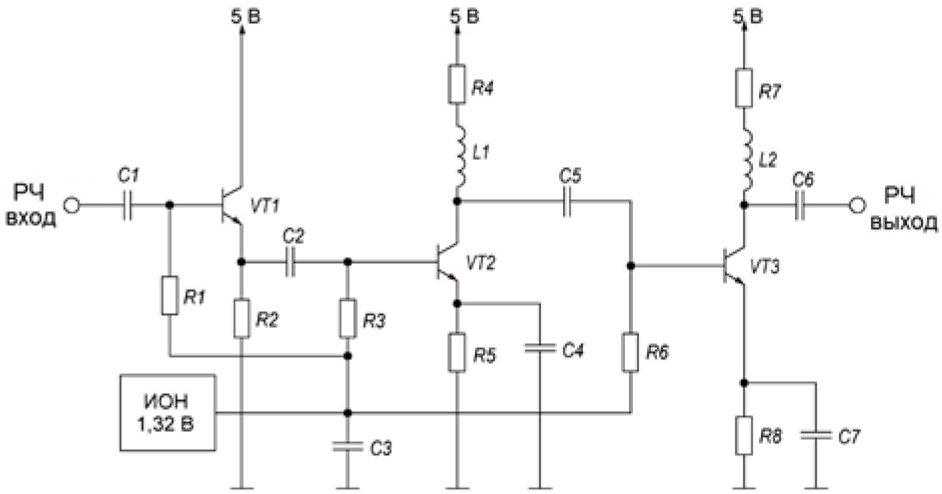


Fig. 11. PA electric circuit.

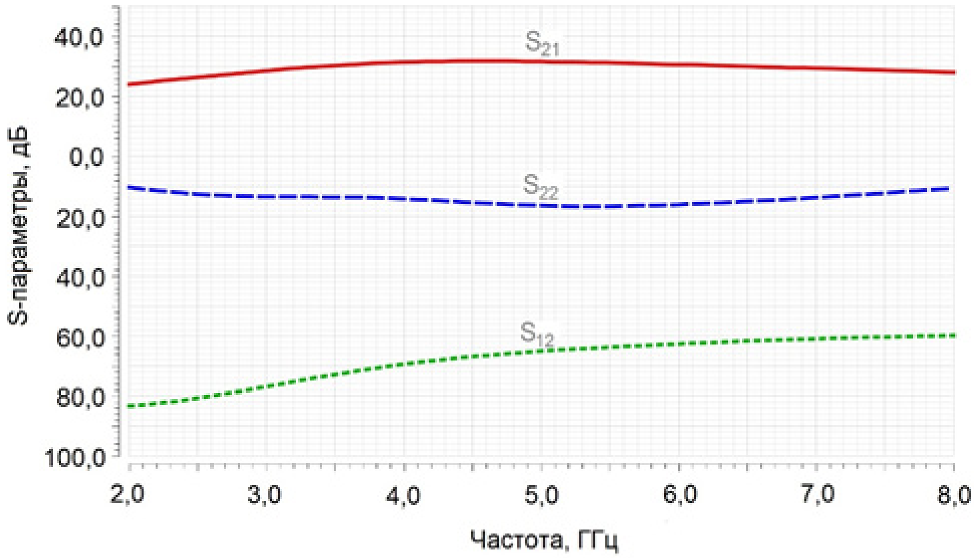


Fig. 12. Frequency dependences of low-signal parameters of PA.

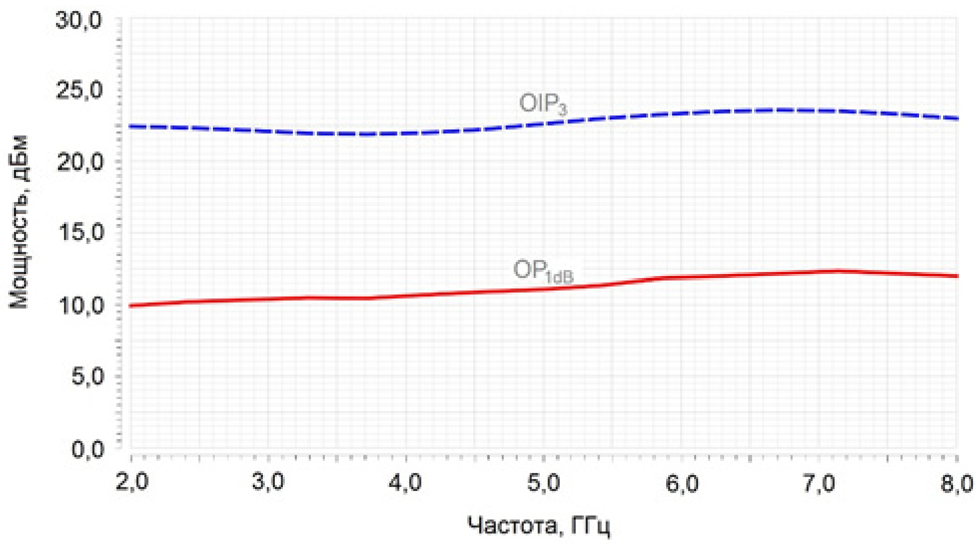


Fig. 13. Dependence of OP_{1dB} and OIP_3 of the PA on the frequency.

The first stage in the circuit with a common collector ensures the coordination of the input impedance of the PA with the output impedance of the previous stage. The second and third stages in the circuit with a common emitter provide the necessary value of the gain. The bias voltage of the VT1–VT3 transistors is 1.3 V and is set with the help of the RVS. For stabilizing the operating points of the transistors and preventing the alternating current component of the voltage from reaching the RVS output, elements R1, R3, R6 and C3 are employed.

The amplifier ensures a gain of 26.5–28 dB in the 4–6 GHz frequency range. Return losses at the output exceed 13 dB (Fig. 12).

Fig. 13 demonstrates the dependences of the 1 dB output compression point (OP_{1dB}) and the point of intersection of the 3rd order output intermodulation (OIP_3) on the frequency. The consumed power from the power supply with a voltage of 5 V is 295 mW.

Dependences of the harmonic distortion factor (HDF) and the energy conversion efficiency of added power (ECE_{AP}) at the input power level are presented in Fig. 14. At a PA input power level of 17.6 dBm (compression point of 1 dB at the input), the HDF is 5.8%, and the ECE_{AP} ¾ 5.6%.

Temperature sensor

The integrated temperature sensor (TS) includes three functional units: a temperature sensor, an analog-to-digital converter and a voltage regulator. The sensor uses the principle of first-order compensation, which involves the summation of temperature factors of the

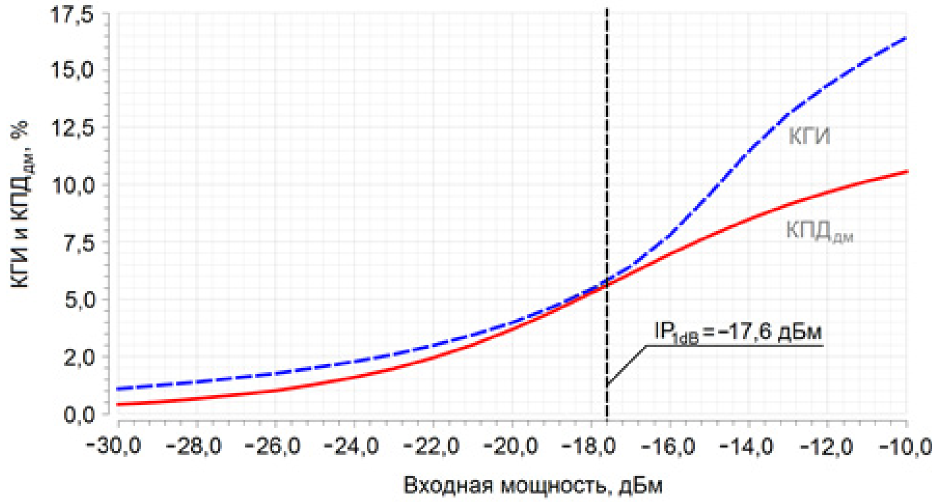


Fig. 14. Dependence of HDF and ECE_{AP} on the input power level.

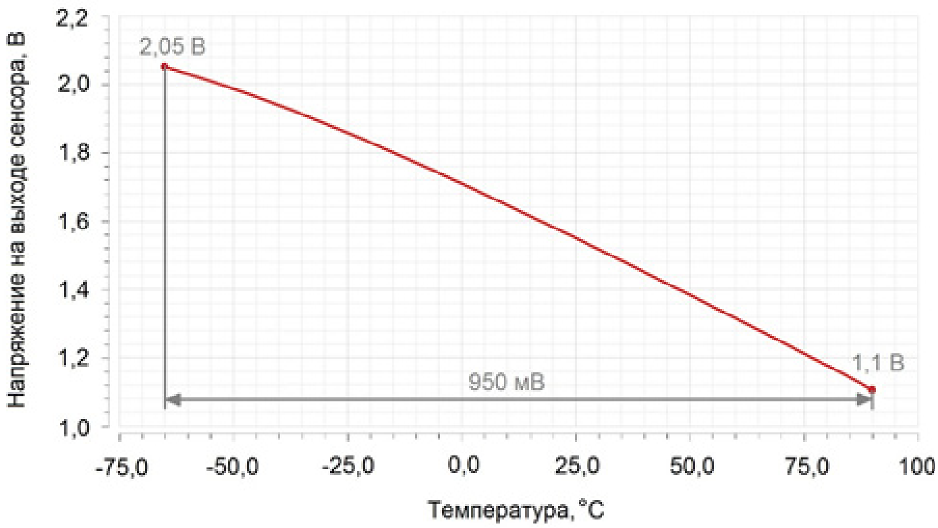


Fig. 15. Dependence of the output voltage of the sensor.

resistance of the p-n junction of the bipolar transistor with the temperature factors of additional transistors. The sensor circuit includes a pair of bipolar transistors, which are controlled by the voltage at the output of the divider formed by resistors selected in such way that the voltage at the collector of one of the transistors depends on the temperature as much as possible (Fig. 15).

In order to provide power for the temperature sensor, a stabilizer based on an operational amplifier with feedback is used. This stabilizer generates a voltage of 2.51 V at its output when the supply voltage changes in the range from 4.5 to 5.5 V. The output voltage of the stabilizer weakly depends on the temperature of the chip. The temperature dependence of the output voltage of the sensor is converted into a parallel digital code with the help of a counter-type analog-to-digital converter. The results of sensor simulation are set forth in Table 3.

Table 3. Parameters of the integrated TS

Parameter, unit	Value
Range of measured temperatures, °C	–65—90
Bit depth of the sensor, bit	5
Resolution, °C	5.156
Measurement error, °C	<5
Measurement time, μs	<1 ($f_{cl.f.}$=50 MHz)
Consumption current, mA	6
Layout area, mm²	0.025

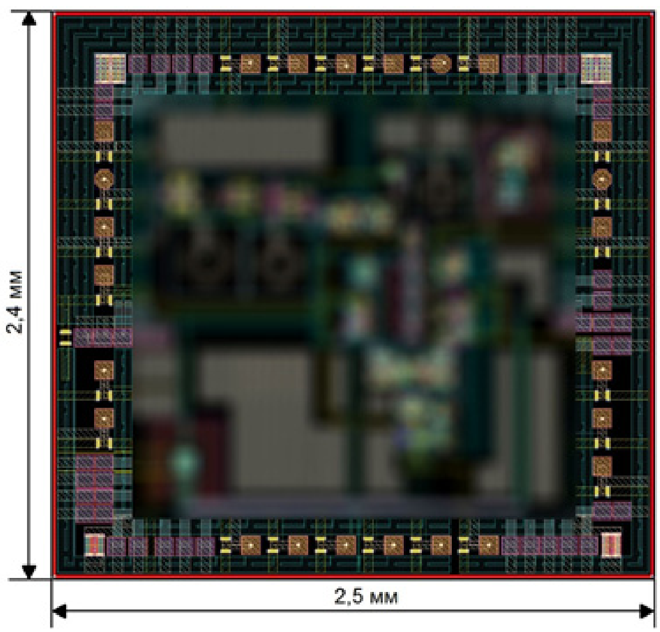


Fig. 16. IP MIC layout.

Result of IP MIC simulation

Fig. 16 presents the developed layout of the IP MIC. The linear dimensions of the layout are 2.5×2.4 mm. The layout area is 6 mm^2.

The IP MIC characteristics that were obtained with the help of circuit simulation with account for parasitic elements of the layout are given in Table 4. The chosen approaches to the design, circuit and layout solutions, as well as the stability of the technological process allows us to count on a satisfactory result after the production of a test batch of IP MICs [12,13].

Table 4. IP MIC parameters

Parameters, unit	IP MIC under development	
	Reception mode	Transmission mode
Range of operating frequencies, GHz	4—6	
Range of operating temperatures, °C	–65—85	
CPS-1 bit depth, bit	6	
CPS-1 least significant bit, deg	5.625	
CATT-1 bit depth, bit	5	
CATT-1 least significant bit, dB	1	
RMS deviation in phase, deg	<4.2	
RMS deviation in amplitude, dB	<0.4	
Transmission factor at a frequency of 5 GHz, dB	35	33
Non-uniformity of the transmission factor in the operating frequency band, dB	<1.9	<3.5
Upper limit of output linearity, dBм	–19.7	8.5
Noise factor at a frequency of 5 GHz, dB	<5.2	—
VSWR in output, units.	<1.8	<1.3
VSWR in input, units	<1.33	<1.8
Isolation between reception and transmission channels, dB	>50.5	>62.5
Bit depth of the control code of the DCU, bit	19	
"0"/"1" of the DCU, V	0/1.8	
Supply voltage, V	2.5; 5	
Consumption current, mA	<39	<73
Power consumption, mW	<195	<365
Layout area, mm²	6	

Conclusion

The article sets forth the results of developing an IP MIC transceiving module of a C-band AESA. The integrated circuit contains one reception/transmission channel. Switching between operating modes of the IP MIC is performed by changing the state of microwave switches. The amplitude and output signal phase are controlled by changing the state of the controlled phase shifter and attenuator. The implementation of a circuit for the correction of IP MIC parameters based on an integrated TS, correcting CPS-2 и CATT-2 allows ensuring the stability of the parameters of the transceiving module in a wide range of operating temperatures. The noise factor in reception mode does not exceed 5.2 dB. The 1 dB compression point in output in transmission mode is 8.5 dB. The power consumed in reception and transmission modes does not exceed 195 mW and 365 mW, respectively. Considering the obtained characteristics, the IP MIC can be widely implemented as part of a low-power base station of fifth-generation telecommunication systems.

Sources of financing

The article was prepared based on the results of a study conducted with the support of the Ministry of Science and Higher Education of the Russian Federation, project No. 8.3962.2017/4.6.

References

1. CGY2175AHV/C1 3-Port C-Band Integrated Core Chip. OMMIC. URL: http://www.ommic.fr/produits/w2175ah-46 (accessed 15 April 2019).
2. METDA Mnogotselevyye monolitnyye IS (kristally) [Multipurpose monolithic IC (chips)]. Active Supply. URL: http://activesupply.ru/files/downloads/multifunctional%20chips.pdf (accessed 15 April 2019). (in Russian)
3. RFCore Multifuction ICs. RFCore. URL: http://www.rfcore.com/index.php/topic/product?id=8&v=7&topicid=136 (accessed 15 April 2019).
4. Oppermann M., Rieger R. Multifunctional MMICs – Key Enabler for Future AESA Panel Arrays. 2018 IMAPS Nordic Conference on Microelectronics Packaging, Oulu, 2018, pp. 77—80.
5. Rebeiz G.M., Paulsen L.M. Advances in low-cost phased arrays using silicon technologies. 2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting, San Diego, CA, 2017, pp. 1035—1036.
6. Timoshenkov V.P., Efimov A.G. Use of Silicon-Germanium Technology for the Development of Active Microwave Units of Active Electronically Scanned Arrays. Russian Microelectronics, 2017, Vol. 46, No. 7, pp. 516—522.
7. Lohmiller P., Reber R., Schuh P., Oppermann M., Chartier S. SiGe BiCMOS X-band transceiver-chip for phased-array systems. 2017 47th European Microwave Conference (EuMC), Nuremberg, 2017, pp. 1253—1256.
8. Hagiwara T., Shinjo S., Kagano M., Taniguchi E. An L-band SiGe BiCMOS core chip MMIC for transmit/receive modules. 2014 9th European Microwave Integrated Circuit Conference, Rome, 2014, pp. 73—76.
9. Jeong J.-Ch., Shin D., Ju I., Yom I.-B. An S-Band Multifunction Chip with a Simple Interface for Active Phased Array Base Station Antennas. ETRI Journal, 2013, Vol. 35, No 3, pp. 378—385.
10. Cetindogan B., Ozeren E., Ustundag B. A 6-Bit Vector-Sum Phase Shifter with a Decoder Based Control Circuit for X-Band Phased-Arrays. IEEE Microwave and wireless components letters, 2016, pp. 64—66.
11. Kumar T., Ma K., Yeo K., Yang W. A 35 mW 30 dB Gain Control Range Current Mode Programmable Gain Amplifier with DC Offset Cancellation. 2014 IEEE MTT-S International Microwave Symposium (IMS2014), Tampa, FL, 2014, pp. 1—4.
12. Saint C., Saint J. IC Mask Design: Essential Layout Techniques. McGraw-Hill, New York, 2002.
13. TSMC 0.18 μm BiCMOS Mixed Signal SiGe 3P6M/3P5M/3P4M Salicide 1.8/3.3V Process Design Kit (PDK) Design Rules. Document No. T-018-BI-DR-001.

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 32–37

Digital Signal Processing Using an RTL-SDR Dongle by the Example of FM Signal

I.R. Gubaydullin, *contact@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
T.T. Mamedov, *contact@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation

Abstract. The article describes the reception of frequency-modulated signals using a software-defined radio (SDR) based on hardware and software platforms of the RTL-SDR project and the following demodulation of the signal in the MATLAB. The operation principle and the mathematical model of the SDR dongle are considered. The mathematical model describes the process of signal receiving and processing that occurs in the receiver prior to the sampling of the input signal. The experimental part describes the process of receiving and processing the frequency-modulated signal of one of the Moscow radio stations presented as discrete samples of the in-phase and quadrature components in the MATLAB. The use of RTL-SDR dongle in conjunction with the MATLAB mathematical software allows software-based digital signal processing, which resulted in the reproduction of the transmitted audio information in the live broadcast.

Keywords: software-defined radio (SDR), digital signal processing, FM signal

Introduction

Software defined radio (SDR) is the technology allowing one to change software radio frequency parameters of the receiver. Over the past 20 years, SDR has attracted a lot of interest among researchers and developers [1]. Such devices differ in a possibility of implementation of functions of the physical layer in a program way that provides implementation of processing of different signal types without change of the hardware. When cheap devices appeared, interest in this technology increased even more as it permits creating a cheap radio operating in the range from tens megahertz to units of gigahertzes.

This paper considers usage of RTL-SDR and MATLAB to receive and demodulate frequency-modulated signals.

RTL-SDR description

In general case, SDR includes: a receiving antenna, analog-to-digital converter (ADC), frequency converter, and single processing device. A single processing device can be programmable logic device (PLD) or signal processor. In the present case, the signals were processed using a personal computer.

During creation of SDR, it is important to understand what functions are performed hardware and what is possible to carry out software. The simplified block diagram of the RTL-SDR receiver can be presented in the successive joining of the two chips: Raphael Micro R820T and Realtek RTL2832U (Fig. 1).

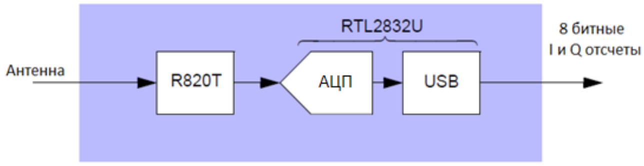


Fig. 1. The block diagram of RTL-SDR. (Antenna / ADC / 8-bit I and Q readings)

The chip R820T performs as radio receiver operating by superheterodyne principle in the frequency range from 24 MHz to 1850 MHz. The chip performs the following: amplification, bandpass filtration, transmission to the intermediate frequency 3.57 MHz [2], and low-frequency filtration. When working with R820T, it is necessary to set the central frequency f_c and amplification coefficient *G*.

It is known, that the chip RTL2832 contains the ADC with the sample rate of 28.8 MHz, digital oscillators, and decimation filters of lower frequencies. On the output of the chip there are formed and transmitted through USB interface 8-th bit readings of in-phase (I) and quadrature (Q) components of a signal. Sample rate of output reports is set software and can be up to 2.8 MHz.

Here the hardware part of SDR comes to an end and following operations such as filtration, demodulation, and signal detection are software realized.

Fig. 2 depicts a functional scheme of the RTL-SDR dongle [3].

The input signal $r(t)$ is transmitted into the system:

$$r(t) = s(t) + N(t) \tag{1}$$

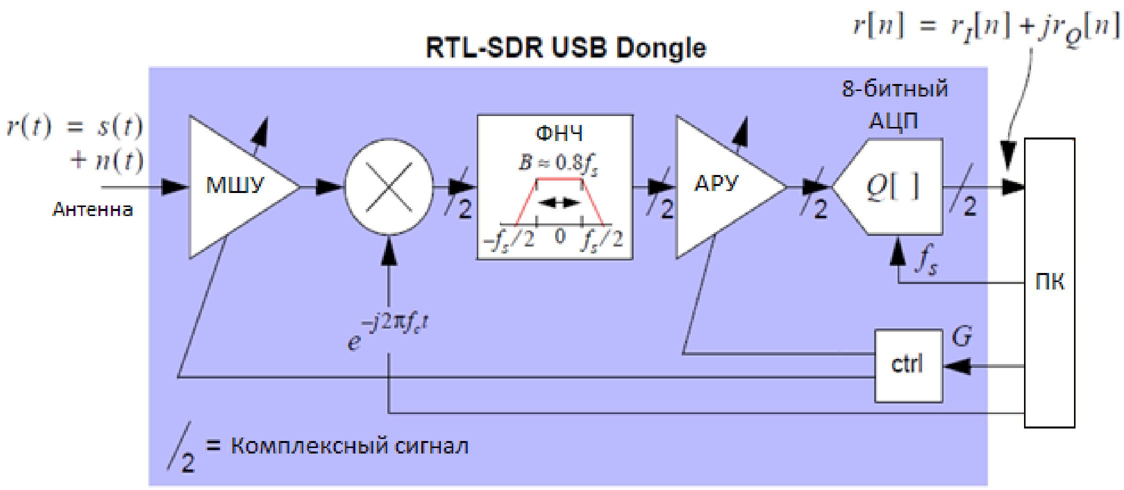


Fig. 2. The functional scheme of the RTL-SDR dongle. (Antenna / LNA / LPF / AGC / 8-bit ADC / Computer / Complex signal)

where $N(t)$ is the noise component of the signal, and $s(t)$ is the useful signal.

The input signal is transmitted onto the low-noise amplifier, which main objective consists in signal amplification without distortions. The amplified signal is mixed with the heterodyne signal s_r:

$$s_r = e^{-j2\pi fct}$$

Hence, there is a multiplication of the amplified input signal with a heterodyne signal operating at the negative carrier frequency. The illustration of this process is given in Fig. 3.

Fig. 3. The transmission of a signal into the zero frequency. (Shift to the left / Effective bandpass / discrete signal spectrum)

At the output of the mixer, the signal can be presented in a complex form, since [4]:

$$e^{j\theta} = \cos\theta + j\sin\theta.$$

After the mixer, a signal is transmitted to the built-in low-pass filter (LPF), where the actual and imaginary part of the signal is separately filtered, i.e., the filter discards the high-frequency part generated by multiplying the two signals.

In the discrete form, the signal can be described as follows [5]:

$$r[n] = G \cdot LP\{r(t)e^{-j2\pi fct}\}_{t=n/fs},$$

where LP is the transfer function of a LPF.

In the frequency domain, the signal is described as

$$R(f) = G \cdot R(f+f_c) \cdot H_{LP}(f).$$

The last element of signal processing to receive reports on in-phase and quadrature components is an 8-bit ADC. An output signal from an ADC is described by the sum of signals:

$$r[n] + N[n] = (r_I[n] + jr_Q[n]) + (N_I[n] + jN_Q[n]).$$

An output signal, in the form of a flow of 8-bit readings is transmitted from a RTL-SDR dongle via the USB interface, which can be processed in the MATLAB.

The analysis of capabilities and prospects of the RTL-SDR application is described by the example of frequency-modulated signal reception and processing with MATLAB [6].

MATLAB used an optional installation package to work with RTL-SDR dongle – “Communications System Toolbox Support Package for RTL-SDR Radio”, where there are the means to carry out interaction with RTL-SDR, and digital signal processing library. This package allows receiving a digitized signal for which receiving it is necessary to set the carrier frequency and sampling rate. Then, a signal in the form of an array of complex numbers is transmitted to the discriminator, where the above operations are performed.

Frequency modulation is a type of analog modulation at which the information signal controls the frequency of the carrier fluctuation, at the same time a derivative of a phase deviation is proportional to an information signal [7]. A frequency-modulated signal is described by the expression:

$$x_c(t) = A_c \cos(2\pi f_c t + \varphi(t)) = A_c \cos\left(2\pi f_c t + 2\pi k_d \int_{-\infty}^{t} m(t)dt\right)$$

where k_d is the modulation coefficient, $m(t)$ is the modulating signal.

The received frequency-modulated signals are transmitted to a complex bandpass discriminator, also known as a quadricorrelator. At the output of the quadricorrelator, real and imaginary parts of the received signal are generated:

$$x'_c(t) = \cos(2\pi\Delta ft + \varphi(t)) + j\sin(2\pi\Delta ft + \varphi(t)) =$$
$$= x_I(t) + jx_Q(t),$$

where Δf is the error in frequency occurring at the signal transmission to the “zero” frequency.

The above expressions relate to continuous time systems, as they relate to discrete systems, further statements relate to the implementation of a digital signal processing system from the RTL-SDR dongle.

The following demodulator algorithm was implemented in the MATLAB.

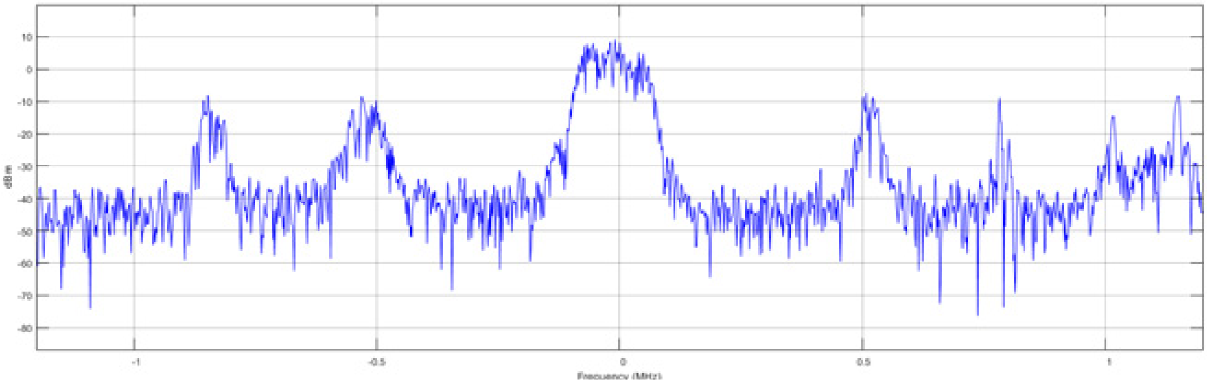


Fig. 4. The input signals spectrum.

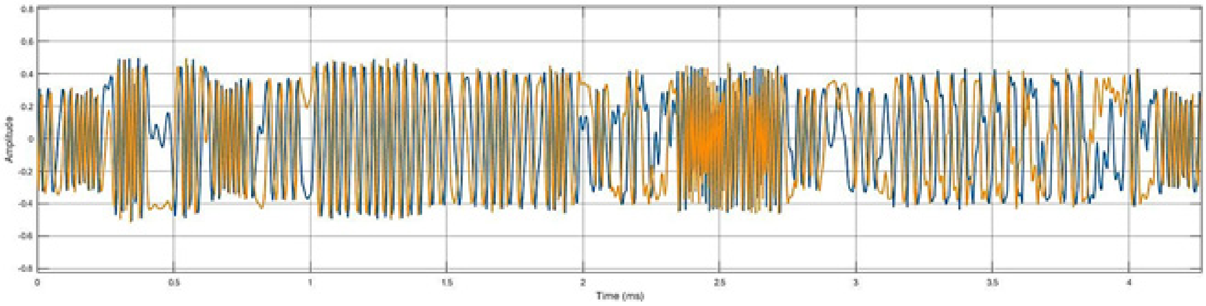


Fig. 5. *I* and *Q* components of the input signals.

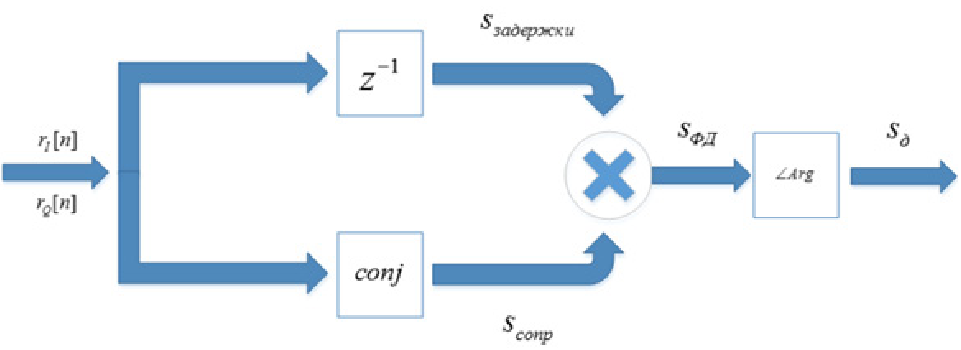


Fig. 6. The discriminator scheme. (S_{delay}/ S_{PD}/ S_d)

The signals output from the RTL-SDR can be expressed as follows:

$$r[n] = (r_I[n] + N_I[n]) + j(r_Q[n] + N_Q[n])).$$

The spectrum of input signals is shown in Fig. 4 [8]. The frequency of one of a popular radio stations in Moscow was used as the center frequency.

Two streams of discrete signals are transmitted to the discriminator input: quadrature and in-phase components of the signal (Fig. 5).

The discriminator performs two operations in parallel: generation of complex conjugated signals and generation of delayed signals per one cycle (Fig. 6).

In MATLAB, the conjugated signal is generated by the ***conj*** function, and the delayed signal is formed by the ***delayseq*** function [6]. Then the signals s_d and s_{conj} are multiplied. In the result, a signal of a phase detector s_{FD} described by the expression [9]:

$$s_{FD}[n] = s_d[n] \times s_{conj} .$$

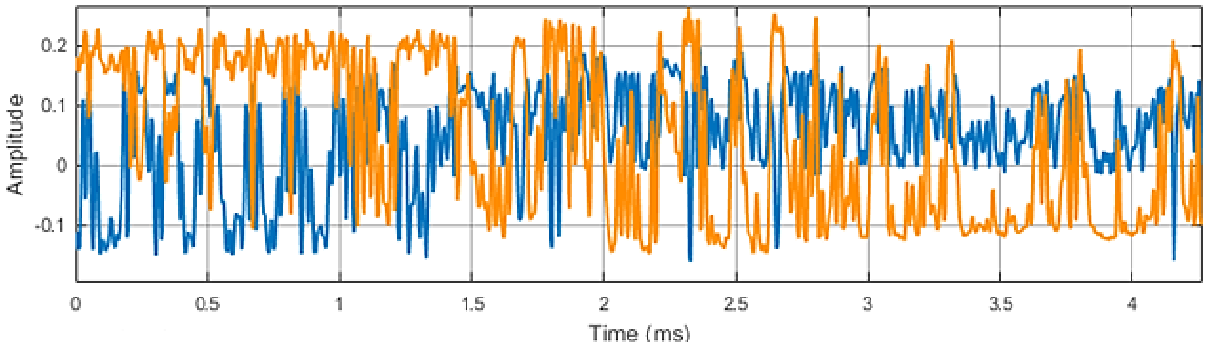


Fig. 7. The S_{FD} signal.

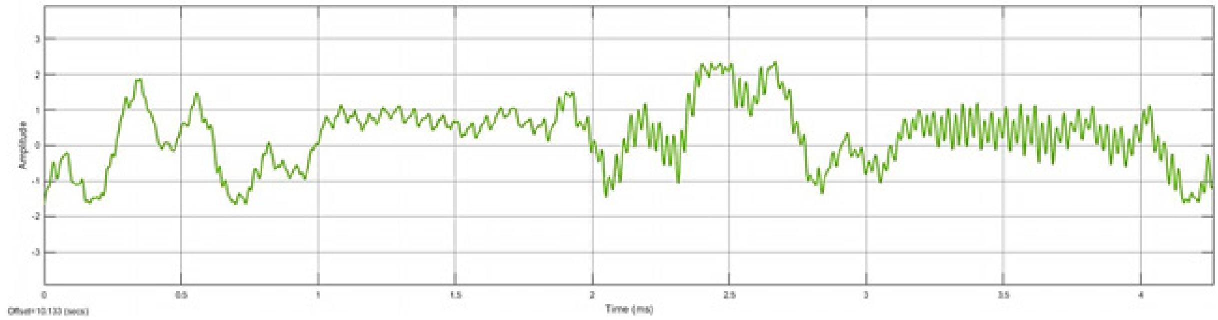


Fig. 8. The signal in the discriminator output.

The arctangent from the ratio of the imaginary component s_{FD} to the real one equals the instantaneous phase and corresponds to the information component of the received. MATLAB allows one to calculate the arctangent using the function ***atan***, to which real and imaginary components of s_{FD} is transmitted as arguments. This operation is performed for each member of the set. In the output of the discriminator, the current signal value can be recorded as follows:

$$s_d(t) = < s_{FD}(t) .$$

As a result, a discrete signal will be generated at the output of the system $s_d[n]$:

$$s_d[n] = <\{(r_I[n] - r_Q[n]) \times (r_I[n - 1] - r_Q[n - 1])\} .$$

Then the signal is remodified and reproduced on audio speakers using the function ***sound*** [10]. The output signal of the system is shown in Fig. 8.

As a result when carrying out the above-stated actions during the experiment from PC loudspeakers the music corresponding to that radio station, which carrier frequency was selected as central for RTL-SDR receiver, was reproduced.

Conclusion

In this article, the device and mathematical model of RTL-SDR dongle are considered. The demodulation algorithm of the FM signal generated by the radio station and received with the help of RTL-SDR dongle is considered with the following digital signal processing in the MATLAB modeling system. The results showed that the use of RTL-SDR dongle in combination with MATLAB allows employing configurable algorithms of digital signal processing for radio communication needs.

References

1. Stewart R., Kenneth B., Atkinson D., Crockett L. *Software Defined Radio using MATLAB & Simulink and the RTL-SDR*. Glasgow: Department of Electronic and Electrical Engineering, University of Strathclyde, 2015.
2. Kadman W. *R820T Rafael Microchip - High Performance Low Power Advanced Digital TV Silicon Tuner*. March 2013. Available at: http://radiofiction.com/showarticle/176.php

3. *Mathworks, Communication System Toolbox Support Package for RTL-SDR Radio*. User guide. 2013.
4. Kharkevich A.A. *Osnovy radiotekhniki* [Basics of Radio Engineering]. 3rd ed., Moscow, Fizmatlit, 2007. (in Russian)
5. Rice M. *Digital Communications: A Discrete-Time Approach*. New Jersey, Pearson Education, 2009.
6. Dyakonov V.P. *MATLAB R2007/2008/2009 dlya radioinzhenerov* [MATLAB R2007 / 2008/2009 for radio engineers]. Moscow, DMK Press, 2010. (in Russian)
7. Shima J. *FM Demodulation using digital radio and digital signal processing*. University of Florida, 1995.
8. Poularikas A. *Understanding Digital Signal Processing with MATLAB and Solutions*. Boca Raton, CRC Press, 2018.
9. Sergienko A.B. *Tsifrovaya obrabotka signalov: Uchebnik dlya vuzov* [Digital Signal Processing: A Textbook for Universities]. 2nd ed., St. Petersburg, Piter, 2007. (in Russian)
10. Dyakonov V.P. *MATLAB. Polnyy samouchitel'* [MATLAB. Full self-study guide]. Moscow, DMK Press, 2012. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 38–44

Passage of Electromagnetic Waves of Elliptical Polarization through a Flat Dielectric Plate

O.A. Kurdyumov, *Cand. Sci. (Phys.-Math.), associate professor, mail.orkkniikp/inkosp.ru*
A branch of Joint Stock Company "United Rocket and Space Corporation" – "Institute of Space Device Engineering", Moscow, Russian Federation
V.E. Sagach, *mail.orkkniikp/inkosp.ru*
A branch of Joint Stock Company "United Rocket and Space Corporation" – "Institute of Space Device Engineering", Moscow, Russian Federation

Abstract. The article analyzes the passage of electromagnetic waves of elliptical polarization through a flat dielectric plate based on the elliptical wave represented as a sum of horizontal and vertical polarization waves with a quadrature phase shift. In this case, not only amplitude but also phase characteristics of the waves transmitted through the plate are investigated. The results obtained allow one to study the influence of dielectric plates on the amplitudes and phases of electromagnetic waves transmitted through them at random values of the thicknesses and permittivity.

The formulae are obtained for calculating the attenuation and phase of electromagnetic waves of linear and elliptical polarization transmitted through the plate. The change in ellipticity coefficient is examined for elliptical polarization.

The results of the calculation for the 1.6 GHz frequency of the plate of 5 mm in thickness and several values of permittivity are presented.

The calculation is compared with the experiment, and it showed the correctness of the calculations.

The influence of the blisters located in the near zone on the characteristics of a microstrip antenna of the L-band is experimentally studied. It is shown that such blisters practically do not affect the diagrams of an antenna.

Keywords: electromagnetic waves, dielectric plate, blister, attenuation, phase, ellipticity coefficient

Introduction

In today's application of satellite navigation systems, antennas are widely used to receive electromagnetic waves of elliptical polarization. As a rule, these antennas are equipped with radiotransparent protective caps – blisters. Usually, consumer antenna blisters are designed such way that their parts that influence the reception of electromagnetic waves are in the form of flat dielectric plates. When blisters are located as close to the radiators as possible, they enter the near wave zone. In this case, antenna radiators are adjusted to the blisters, and it makes no sense to consider the passage of electromagnetic waves through blisters. Instead, one needs to analyze an antenna with a blister taken into consideration.

The influence of blisters located in the far zone of antennas on their characteristics is studied in sufficient detail in monographs [1–5]. However, most publications refer to antennas and linear polarization waves. The effect of blisters on the properties of rotating polarization antennas has not been studied yet in the literature.

Near receiving antennas, especially on aircraft, there may be structural elements that may affect antenna performance. The antennas are particularly heavily influenced by metal elements. Therefore, it is common to formulate a requirement that antennas should not be shadowed at a certain degree to the horizon. However, dielectric elements can also change antenna characteristics. Hence, it is interesting to investigate passage of electromagnetic waves through dielectric plates when they are located in the far zone of antennas. Passage of linear polarization waves through the dielectric plate is studied in the papers [6–8], but the phases of the passed waves are not determined.

This paper analyzes the diffraction of electromagnetic waves of elliptical polarization on a flat dielectric plate based on the representation of an elliptical wave as a sum of horizontal and vertical phase-shift quadrature polarizations. At the same time not only amplitude, but also phase characteristics of waves passed through the plate are examined. The results obtained allow studying the influence of dielectric plates on amplitudes and phases of the electromagnetic waves which passed through them at any values of thickness and dielectric permeability.

Horizontal polarization

The structure under study is shown in Fig. 1.

We introduce the vector magnetic potential **A** (hereinafter the vector values are indicated in bold). Components of magnetic and electric field are determined from the ratios [7]:

$$\boldsymbol{H} = \frac{1}{\mu} rot\, \boldsymbol{A} \tag{1}$$

$$\boldsymbol{E} = \frac{1}{j\omega\varepsilon} rot \boldsymbol{H} \tag{2}$$

Vector potentials in the regions are represented as incident and reflected waves:

$$\boldsymbol{A_1} = \boldsymbol{A_{1п}} e^{-j(k_{1x}x+k_{1z}z)} + \boldsymbol{A_{1о}} e^{j(k_{1x}x+k_{1z}z)} \tag{3}$$

$$\boldsymbol{A_2} = \boldsymbol{A_{2п}} e^{-j(k_{1x}x+k_{1z}z)} + \boldsymbol{A_{2о}} e^{j(k_{1x}x+k_{1z}z)} \tag{4}$$

$$\boldsymbol{A_3} = \boldsymbol{A_{3п}} e^{-j(k_{1x}x+k_{1z}z)} \tag{5}$$

where the lower indexes correspond to the area numbers in Fig. 1.

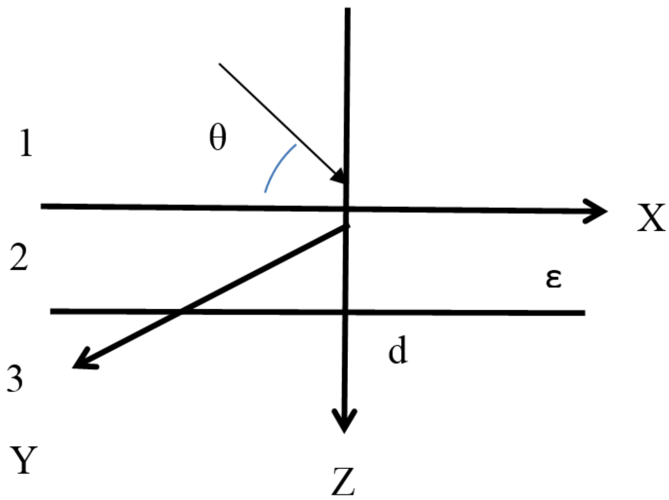


Fig. 1. Fall of electromagnetic wave at the angle θ onto a dielectric plate with the thickness *d* with the dielectric constant ε. Medium *1* and *3* is air, medium *2* is dielectric.

Using formulae (1) and (2), the following expressions are obtained for the magnetic and electric field components necessary for further actions:

$$H_{x1} = j\frac{k_{1z}}{\mu_0}(A_{1п}e^{-jk_{1z}z} - A_{1о}e^{jk_{1z}z})e^{-ik_{1x}x} \tag{6}$$

$$H_{x2} = j\frac{k_{2z}}{\mu_0}(A_{1п}e^{-jk_{2z}z} - A_{1о}e^{jk_{2z}z})e^{-ik_{1x}x} \tag{7}$$

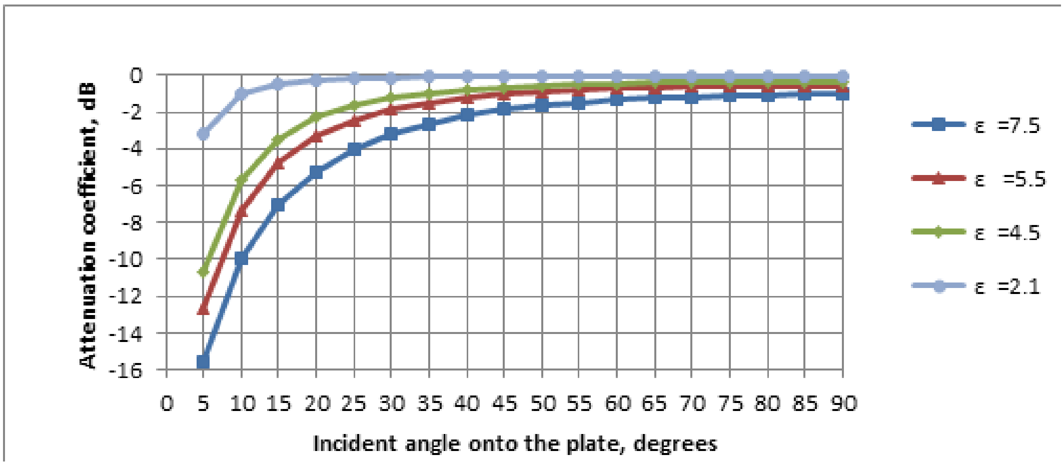


Fig. 2. Attenuation coefficient of electromagnetic wave of linear horizontal polarization, λ=188 mm, by a dielectric plate with the thickness of *d*=5 mm.

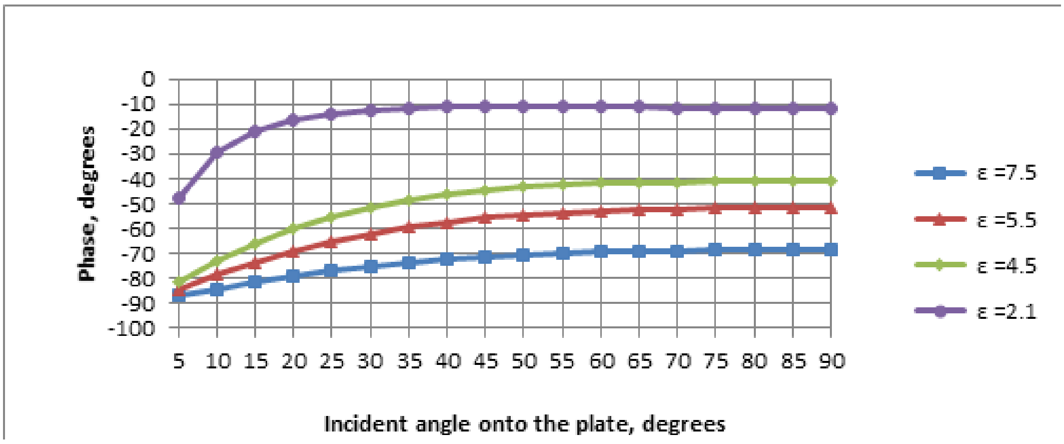


Fig. 3. The phase of linear horizontal polarization electromagnetic wave, λ=188 mm, passed through a dielectric plate with the thickness of *d*=5 mm.

$$H_{x3} = j\frac{k_{1z}}{\mu_0}A_{3п}e^{-jk_{1z}z}e^{-ik_{1x}x} \quad (8)$$

$$E_{y1} = \frac{k_1^2}{j\omega\varepsilon_0\mu_0}(A_{1п}e^{-jk_{1z}z} + A_{1о}e^{jk_{1z}z})e^{-ik_{1x}x} \quad (9)$$

$$E_{y2} = \frac{k_2^2}{j\omega\varepsilon\varepsilon_0\mu_0}(A_{2п}e^{-jk_{2z}z} + A_{2о}e^{jk_{2z}z})e^{-ik_{1x}x} \quad (10)$$

$$E_{y3} = \frac{k_1^2}{j\omega\varepsilon_0\mu_0}A_{3п}e^{-jk_{1z}z}e^{-ik_{1x}x} \quad (11)$$

where ε_0 and μ_0 are electric and magnetic permittivity of vacuum, $k_1^2 = \omega\varepsilon_0\mu_0$ and $k_2^2=\omega\varepsilon\varepsilon_0\mu_0$ are the squares of wave numbers of the corresponding regions, k_{1x}, k_{1z}, k_{2x}, k_{2z} are the constants propagation along the *x* and *z* axes of the corresponding regions.

We will require fulfillment of boundary conditions at the boundaries of the areas representing equality of tangential components of electric and magnetic fields: at $z = 0$, $E_{y1} = E_{y2}$, $H_{x1} = H_{x2}$, at $z = d$, $E_{y2} = E_{y3}$, $H_{x2} = H_{x3}$.

We will obtain a system of algebraic equations:

$$A_{1n} + A_{1o} = A_{2n} + A_{2o} \quad (12)$$

$$k_{1z}(A_{1n} - A_{1o}) = k_{2z}(A_{2n} - A_{2o}) \quad (13)$$

$$A_{2n}e^{-jk2zd} + A_{2o}e^{jk2zd} = A_{3n}e^{-jk1zd} \quad (14)$$

$$k_{2z}(A_{2n}e^{-jk2zd} - A_{2o}e^{jk2zd}) = k_{1z}A_{3n}e^{-jk1zd} \quad (15)$$

Solving this system of equations, we will find the coefficient of passage from the medium *1* into the medium *3*, K_g, and corresponding phase incursion ϕ_g:

$$K_g = \frac{A_{3п}}{A_{1п}} = \frac{e^{jk_{1z}d}}{\cos k_{2z}d + 0.5j(\frac{k_{1z}}{k_{2z}}+\frac{k_{2z}}{k_{1z}})\sin k_{2z}d} \quad (16)$$

$$\phi_g = k_{1z}d - \tan^{-1}[0.5\,(\frac{k_{1z}}{k_{2z}} + \frac{k_{2z}}{k_{1z}})\tan k_{2z}d] \quad (17)$$

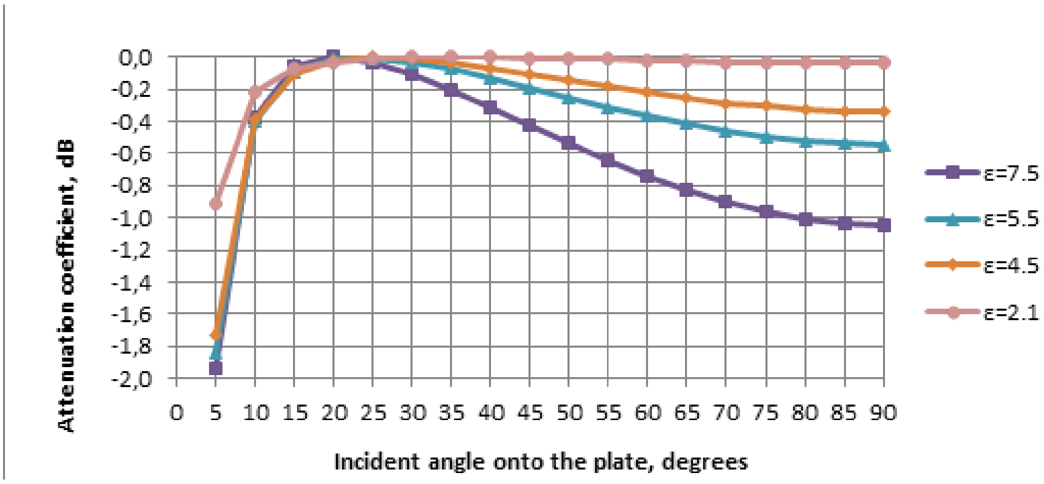


Fig. 4. Attenuation coefficient of electromagnetic wave of linear vertical polarization, λ=188 mm, by a dielectric plate with the thickness of *d*=5 mm.

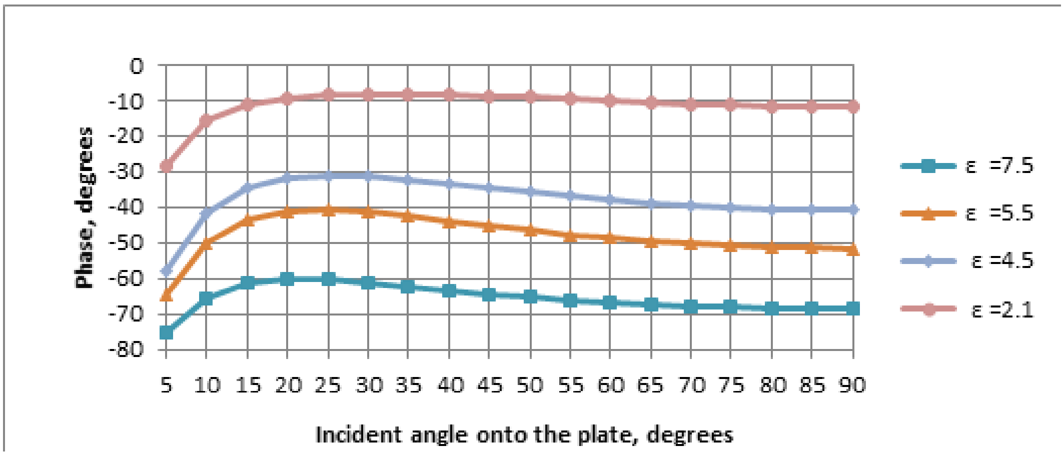


Fig. 5. The phase of electromagnetic wave of linear vertical polarization, λ=188 mm, passed through a dielectric plate with the thickness of *d*=5 mm.

where, according to [4]:

$$k_{2z}d = \frac{2\pi d}{\lambda}\sqrt{\varepsilon - cos^2\,\theta} \tag{18}$$

$$\frac{k_{2z}}{k_{1z}} = \frac{\sqrt{\varepsilon - cos^2\,\theta}}{\sin\theta} \tag{19}$$

Figs. 2 and 3 show the results of calculation of attenuation factor and phase according to the formulae (16) and (17) for the wavelength λ=188 mm, *d*=5 mm and several values of dielectric permittivity.

From these figures it can be seen that the attenuation in the cases considered can be as high as 16 dB and the effect on the phase properties of the antenna can also be significant. In particular, changes in the phase of the signal arriving at the antenna in the upper hemisphere can reach –90°.

Vertical polarization

For this case, it is more convenient to introduce a vector electric potential ***F*** through which the components of the electric and magnetic field are defined as follows [7]:

$$\boldsymbol{E} = \frac{1}{\varepsilon} rot\,\boldsymbol{F} \tag{20}$$

$$\boldsymbol{H} = \frac{1}{j\omega\mu} rot\boldsymbol{E} \tag{21}$$

Saving behind electric vector potentials in the areas *1–3* the expressions (3)–(5) and having done conversions similar to the previous section, we will receive the following system of the algebraic equations concerning amplitudes of fields:

$$A_{1n} + A_{1o} = \varepsilon\,(A_{2n} + A_{2o}) \quad (22)$$

$$k_{1z}\,(A_{1n} - A_{1o}) = k_{2z}\,(A_{2n} - A_{2o}) \quad (23)$$

$$\varepsilon\,(A_{2n}e^{-jk2zd} + A_{2o}e^{jk2zd}) = A_{3n}e^{-jk1zd} \quad (24)$$

$$k_{2z}\,(A_{2n}e^{-jk2zd} - A_{2o}e^{jk2zd}) = k_{1z}A_{3n}e^{-jk1zd} \quad (25)$$

Hence, we get the expressions for the transmission factor and the phase incursion:

$$K_v = \frac{A_{3п}}{A_{1п}} = \frac{e^{jk_{1z}d}}{\cos k_{2z}d + 0.5j(\varepsilon\frac{k_{1z}}{k_{2z}} + \frac{k_{2z}}{\varepsilon k_{1z}})\sin k_{2z}d} \quad (26)$$

$$\phi_v = k_{1z}d - \tan^{-1}[\,0.5\,(\varepsilon\frac{k_{1z}}{k_{2z}} + \frac{k_{2z}}{\varepsilon k_{1z}})\tan k_{2z}d] \quad (27)$$

The ratios (18) and (19) remain true.

Similar to the previous section, Figs. 4 and 5 show the results of calculation of attenuation factor and phase by the formulae (26) and (27) for the wavelength λ=188 mm, *d*=5 mm and several values of dielectric permittivity.

The general nature of the dependences of these values on the incidence angle remains the same as for horizontal polarization. Variation of attenuation in the upper hemisphere of the antenna does not exceed 2 dB, and of the phase –75°.

Elliptical polarization

Let us present the elliptically polarized incident wave on the plate as a sum of quadrature phase-shifted horizontal and vertical components:

$$A_1 = A_g + jA_v = A_g\,(1 + jR_1) \quad (28)$$

where $R_1 = A_v / A_g$ is the coefficient of ellipticity of incident wave. Then the wave which passed through a plate will be defined by the amplitude:

$$A_3 = K_gA_g + jK_vA_v = A_g\,(K_g + jK_vR_1) \quad (29)$$

The passage coefficient of the elliptically polarized wave will be determined as:

$$K_e = \frac{K_g + jK_vR_1}{1 + jR_1} \quad (30)$$

The ellipticity factor of the passed wave will be equal to:

$$R_3 = \frac{K_v}{K_g}R1 \quad (31)$$

and its phase is defined by the expression:

$$\phi_e = k_{1z}d - \tan^{-1}((R_1\tan\alpha + b) / (a\,R_1 - \tan\alpha)) - \tan^{-1}(((a+b)\tan\alpha) / (1 - ab\tan^2 k_{2z}d)) - \tan^{-1} R1 \quad (32)$$

where

$$a = 0.5\left(\frac{k_{1z}}{k_{2z}} + \frac{k_{2z}}{k_{1z}}\right) \quad (33)$$

$$b = 0.5(\varepsilon\frac{k_{1z}}{k_{2z}} + \frac{k_{2z}}{\varepsilon k_{1z}}) \quad (34)$$

$$\alpha = k_{2z}d \quad (35)$$

Figs. 6–8 show the results of calculation of attenuation coefficient, ellipticity coefficient and phase by the formulae (30)–(32) for the wavelength λ=188 mm, d=5 mm, and several values of dielectric permittivity.

From these figures it can be seen that the attenuation of the elliptically polarized wave in the upper hemisphere is the mean between horizontal and vertical polarizations. Phase change does not exceed –60°. A dielectric plate has a much stronger influence on the ellipticity of the passed wave: in the cases considered, the ellipticity can decrease from 1 to 0.2.

Attenuation of electromagnetic wave at the frequency of 1.6 GHz by a dielectric plate 5 mm thick with dielectric permittivity 4.5 was experimentally checked. When the plate is located in the far zone of the antenna (at the distance of 10 cm from it), the attenuation of ~0.5 dB is obtained, which corresponds well enough to the calculation (0.35 dB).

Dielectric plate in near antenna field

The effect of the blister located in the near field of the antenna on its characteristics was studied on the electrodynamic model shown in Fig. 9.

The antenna consists of a base with the diameter of 120 mm, substrate from material with the dielectric constant 9.6 and height of 6 mm, diameter of 42 mm, radiator with the diameter of 30 mm of the blister with the thickness of 5 mm located at the height of 3 mm above a radiator (side walls of the blister are not shown). The excited state of the radiator is single-point.

The radiation pattern calculations of this model were performed by the decomposition method followed by the moment method.

At the change ε of the blister, the radiator dimensions changed so that the resonance was observed near the frequency of 1.6 GHz. The results obtained are shown in Table 1.

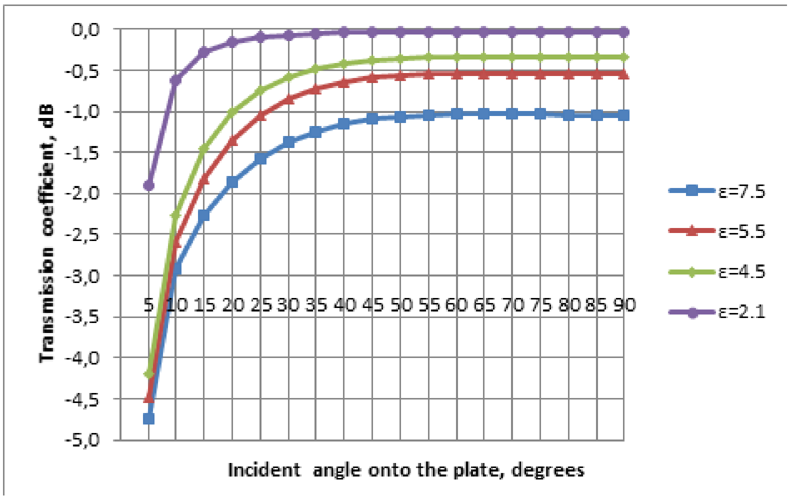


Fig. 6. Attenuation coefficient of electromagnetic wave of elliptical polarization by a dielectric plate with the thickness of d=5 mm, λ=188 mm, R_1=1.

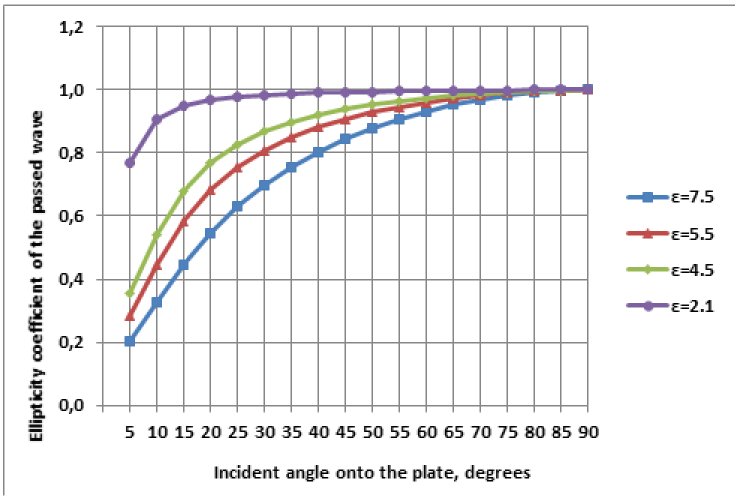


Fig. 7. Ellipticity coefficient of electromagnetic wave of elliptical polarization passed through a dielectric plate with the thickness of d=5 mm, λ=188 mm, R_1=1.

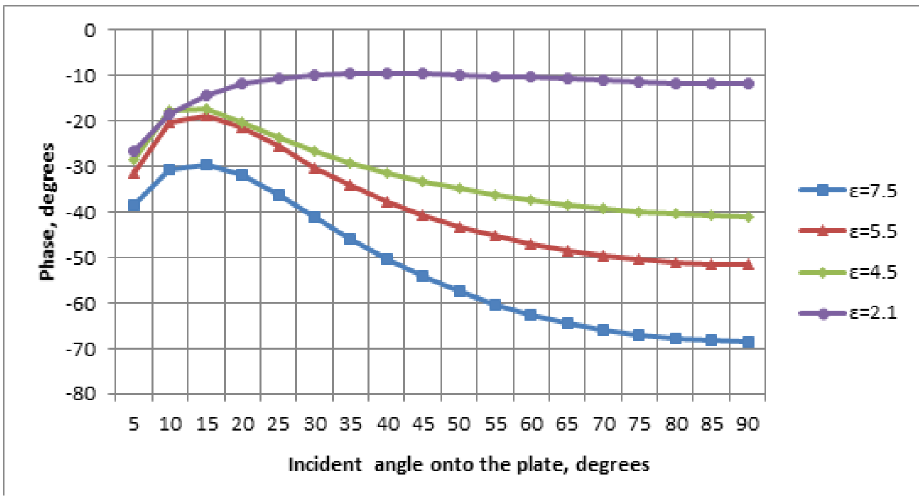


Fig. 8. The phase of electromagnetic wave of elliptical polarization passed through a dielectric plate with the thickness of d=5 mm, λ=188 mm, R_1=1.

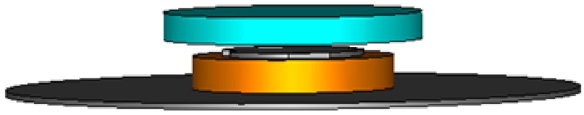


Fig. 9. Electrodynamic model: frequency band microstrip antenna L1.

Table 1. The results of calculation of a microstrip antenna with blisters made of materials with different ε.

ε of the blister	Antenna gain, dB	Ellipticity coefficient at θ = ±90°	Phase change ° at the change of the angle Θ from −90° to 90°
Without blister	6.7	5	8.3
2.1	6.6	4.4	5.7
4.5	6.5	4.5	5.9
5.5	6.3	5	5.8
7.5	6.2	6.4	10.6

As can be seen from the table materials, a blister located in the near field of the antenna practically does not affect its characteristics.

Conclusions

Formulae for calculation of attenuation and phase of electromagnetic waves of linear and elliptical polarizations passed through dielectric plate are obtained. For elliptical polarization, the change in ellipticity coefficient as the plate passes has also been investigated.

Results of calculation for the frequency of 1.6 GHz, a plate 5 mm thick and several values of dielectric permittivity are presented.

A comparison of the calculation with the experiment was made, which showed the correctness of the calculations.

The influence of blisters located in the near zone on the characteristics of the L-band microstrip antenna was experimentally investigated. It is shown that such blisters practically do not affect the diagram properties of the antenna.

References

1. *Obtekateli antenn* [Antenna blisters]. Ed. Shpuntov A.I. Moscow, Izdatel'stvo "Sovetskoye radio", 1950. (in Russian) (translated from English)
2. Kaplun V.A. *Obtekateli antenn SVCH* [Blisters of the microwave antennas]. Moscow, Izdatel'stvo "Sovetskoye radio", 1974, 240 p. (in Russian)
3. Nguayya M.L. *Raschetnyye metody kontrolya sloistykh materialov i izdeliy na osnove issledovaniya koeffitsiyentov otrazheniya i prokhozhdeniya elektromagnitnykh voln: 05.11.13. Dis.... kand. tekhnich. Nauk* [Calculated methods of control of multilayer materials and items based on analysis of reflection and passage coefficients of electromagnetic waves: 05.11.13. (Candidate of Engineering Sciences thesis]. Kazan', KGEU, 2011, 122 p. (in Russian)
4. Minokin M.N. Difraktsiya ploskoy elektromagnitnoy volny na polom dielektricheskom konuse [Diffraction of a flat electromagnetic wave on a hollow dielectric cone]. *Zhurnal radioelektroniki: Elektronnyy zhurnal* [Radio Electronics Journal: Electronic Journal], 2000, No. 6. Available at: http://jre.cplire.ru/jre/jun00/index.html (accessed: May 25, 2019).
5. Kabalin S.V. Issledovaniye kharaktera zavisimosti koeffitsiyenta prokhozhdeniya elektromagnitnoy energii cherez radioprozrachnuyu kryshku ot chastoty radiosignala i tipa polyarizatsii [Study of transmission factor of radio-transparent cap depending on radio wave frequency and type of polarization]. *Antenny* [Antennas]. 2013, No. 1, pp. 41–45. (in Russian)
6. Antonets L.V., Kotov L.N., Shavrov V.G., Shcheglov V.I. Raschet koeffitsiyenta otrazheniya i pogloshcheniya pri naklonnom padenii elektromagnitnykh voln na plastiny [Reflection, transmission, and absorption coefficients calculated for the oblique influence of an electromagnetic wave on a plate]. *Radiotekhnika i elektronika* [Journal of Communications Technology and Electronics], 2008, Vol. 53, No. 4, pp. 389–402. (in Russian)
7. Kugushev A.M., Golubeva N.S. *Osnovy radioelektroniki* [Electronics fundamentals]. Moscow, Izdatel'stvo Energiya, 1969. (in Russian)
8. Gorelov V.V. *Sposob umen'sheniya chastotnoy zavisimosti oslableniya pri prokhozhdenii elektromagnitnoy volny cherez dielektricheskuyuobolochku odnovremenno dlya parallel'noy i perpendikulyarnoy polyarizatsiy i ustroystvo dlya ego realizatsii (vektornyy antennyy obtekatel')* [Method of reducing frequency dependence of attenuation when electromagnetic wave passes through dielectric shell simultaneously for parallel and perpendicular polarizations and device for its implementation (vector antenna blister)]. Patent of the Russian Federation No. 2334315 dated 20.09.2008. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 45–51

Radio Frequency Provision and International Legal Protection of Frequency Assignments for Small Satellites

A. A. Talanov *Cand. Sci. (Engineering), contact@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
S.A. Fedotov *Cand. Sci. (Engineering), contact@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation

Abstract. The article deals with the issues related to the allocation and assignment of radio frequencies, the procedure of their international legal protection, which should be performed by the head developer of a small satellite (SSL) during research, development and production of elements of space systems and complexes.
The analysis of radio frequency ranges of service and target radio lines of well-known Russian SSL is carried out. The data on the radio frequency bands of command and telemetric systems (CTS) and the target information radio line (TIRL) of small satellites developed by Russian companies are presented.
It is noted that the radio frequency bands currently used for the SSL operation impose rather strict restrictions on the construction of a space segment and/or cluster on their basis, which will inevitably have a negative impact on the possibilities of their extended use. Proposed are alternative frequency bands and radio link designs to meet the requirements of the Table of Distribution of Frequency Bands between radio services of the Russian Federation and the Radio Regulations.

Keywords: small satellite, space platform, frequency range, Radio Regulations

Introduction

The number of launched small spacecraft (SSC) [1] of the mini-, micro- and nano- (CubeSat) class [2-5] today reaches several dozens per year, and the demand for them is constantly growing. During the creation of the SSC, the issues of radio-frequency support for the functioning of radio-electronic equipment (REE) and its electromagnetic compatibility, are not always timely and fully solved by developers today due to their complexity and volume. In this regard, those issues become more relevant.

Consider the regulatory framework for the allocation and assignment of radio frequencies, as well as the radio frequency bands used today for the operation of service and target radio links of the SSC.

Legal framework for frequency assignments

All issues related to the allocation and assignment of radio frequencies are regulated by the Federal Law "On Communications" [6] and a number of other legal documents. The law defines the permissive procedure for access to the radio frequency spectrum. The right to use the radio frequency spectrum is granted through the official allocation of radio frequency bands and the assignment of radio frequencies or radio frequency channels. Use of the radio frequency spectrum without permission is not permitted.

An application for the allocation of radio frequency bands should be prepared and submitted during the planning of implementation, development, production, modernization, research, development and experimental work for a period of more than 6 months [7].

The allocation of radio frequency bands is carried out by the State Commission on Radio Frequencies (SCRF) in accordance with the procedure for reviewing materials and making decisions on the allocation of radio frequency bands, rescheduling and amending them [7].

It should be noted that the application for the allocation of radio frequency bands for the development of REE of space complexes and/or systems is submitted by the applicant (lead contractor) no later than the stage of completion of the preliminary or outline technical design [7].

The assignment of a radio frequency or radio frequency channel for civilian electronic equipment is carried out by the federal executive body in the field of communications on the basis of applications, taking into account the results of the radio-frequency examination service, of the possibility of using the declared electronic equipment and its electromagnetic compatibility with existing and planned electronic equipment (expert examination of the electromagnetic compatibility). Consideration of materials and decision making by the State Committee for Radio Frequencies on the assignment (designation) of radio frequencies or radio frequency channels within the allocated radio frequency bands is carried out in accordance with the approved Examination Procedure [8].

The SCRF decision on the assignment (designation) of radio frequencies is not always positive [9, 10], and sometimes it contains significant limitations.

As an example: the applicants were denied the allocation of the radio frequency band 1610.115-1621.185 MHz, 2483.773-2494.845 MHz (space-to-space) for using the GLOBALSTAR ASP satellite modem, intended for installation on the AIST-2D SSC in 2015 [9] and the radio frequency bands 432-433 MHz for the REE of the space complex 197KS of the Aist-2 SSC in 2016 [10] on the basis of negative conclusions about the possibility of allocating radio frequency bands submitted by the members of the state commission on radio frequencies.

At the same time, the SCRF allowed the use of the radio frequency band 8025-8393 MHz (space-to-Earth) by the earth station of the monitoring and control system TsPOI-SAMARA (radio frequency band of shared use by REE of any purpose), provided that no claims are made for possible unintentional interference from governmental and special purpose REE and development of technical measures to ensure electromagnetic compatibility with military REE and coordination with the Ministry of Defense of Russia [10, 12].

In accordance with Article 23 of the Federal Law "On Communications" and Decree of the Government of the Russian Federation dated December 21, 2011 No. 1049-34 SCRF, a Table for the Allocation of Radio Frequency Bands between the Radio Services of the Russian Federation [13] has been developed. The Table is the main document governing the distribution and conditions of use of radio frequency bands in the Russian Federation by citizens of the Russian Federation and Russian legal entities who order, develop or import electronic equipment into the territory of the Russian Federation, as well as plan the use and operation of electronic equipment.

The procedure for using the 144-146 MHz and 435-438 MHz frequencies of the amateur-satellite service (AMSAT) in the Russian Federation is determined by the Decision of the State Committee for Radio Frequencies "On the Allocation of Radio Frequency Bands for Radio-Electronic Means of the Amateur and Amateur Satellite Services" dated 07.15.2010 No. 10-07-01, taking into account the changes approved by the Decision of the State Committee for Radio Frequencies dated 04.16.2018 No. 18-45-02, in accordance with which it is not required to issue separate Decisions of the State Committee for Radio Frequencies and permits for the use of frequencies, subject to the fulfillment of technical and operational restrictions (limitation of the power of radio lines, radiation classes and radio frequency bands, and others.).

The international legal protection of frequency assignments for radio electronic means is carried out in order to ensure the national interests of the Russian Federation in the international regulation of the use of the radio frequency spectrum, international recognition of frequency assignments for radio electronic means, as well as to create favorable conditions for the development and use of radio electronic means, various radio services in the Russian Federation [14].

In the Russian Federation, the procedure of international legal protection of the assignment (designation) of radio frequencies or radio frequency channels, including work related to the declaration, coordination and registration with the International Telecommunication Union of such radio frequencies or radio frequency channels and the corresponding positions of satellites in geostationary orbit or corresponding characteristics of satellites in other orbits, is defined in the following rules introduced by the Decree of the Government of the Russian Federation of November 14, 2014 No. 1194 [14]:

- "Rules for conducting work on international legal protection of the assignment (designation) of radio frequencies or radio frequency channels in the Russian Federation";
- "Rules for the use in the territory of the Russian Federation of satellite communications networks under the jurisdiction of foreign states".

Thus, in the Russian Federation, issues related to the assignment and designation of radio frequencies, the procedure for their international legal protection, which should be carried out by the lead contractor in carrying out scientific research, development and production of elements of space systems and SSC complexes, are normatively fixed and strictly regulated.

Characteristics of the service and target SSC radio links

To analyze the radio-frequency ranges of the service and target radio links of the SSC, let us consider the radio frequencies used in the space-to-Earth and Earth-to-space directions as their main parameter.

Table 1 shows the radio frequency bands of the command-telemetry system (CTS), as well as the radio links for the exchange of target information (RLTI) of small spacecraft designed to solve the following problems:

- operational optical-electronic surveillance of water areas, water bodies of the earth's surface, various emergencies, as well as for operational monitoring of the state of engineering facilities: AIST-2D [9, 10, 15], Kanopus-V [16, 17];
- scientific experiments: Universitetskiy-Tatyana-2 [18], Yubileiny-2 [19, 20], Chibis-M (AI) [10, 21, 22], Baumanets-2 [10], SamSat-218 [10, 23].

SSC for communication, relay, and radio surveillance tasks for detecting, identifying, determining the location and course of movement of ships, rivercraft, and aircraft by the radiation of their radio electronic equipment were not considered in this table.

In accordance with the above data, the frequency range of the CTS operation is within the primary distribution of the AMSAT of 435–438 MHz and 144–146 MHz [13, 24–26]. Note that the frequency range 144-146 MHz is quite heavily loaded, since amateur and airborne mobile communications are also in this range [13, 26].

The amateur, radar services, as well as the Earth exploration-satellite service (EESS) (active) are also in the 435–438 MHz frequency range [13, 26].

The radio frequency band 435–438 MHz can be used by AMSAT stations on a secondary basis, provided that interference with other services is eliminated [13, 24, 25]. This radio frequency band is used by military and special purpose radio electronic equipment, which may impose restrictions on its usage and create unacceptable interference for the operation of other radio links in these frequency bands, so protection against interference when using this range cannot be guaranteed.

Note that the SamSat-218 SSC radio link developed by the SSAU is not compatible in the direction of transmission with the rest of the given radio links.

As an alternative, the use of radio frequency bands 267-273 MHz (space-to-Earth) specially allocated for the Space Operation Service (SOS) can be considered [13, 24, 25].

Table 1 - the Main characteristics of the SSC radio lines [34]

Organization	SSC (launch date)	Bus	RLTI (space-to-Earth) MHz	CTS (space-to-Earth) (Earth-to-space) MHz
JSC "RSS" Progress	AIST-2D (28.04.2016)	AIST-2 [15]	8025-8393 [10, 12]	435.3065-435.3235 435.3556-435.3735/ 145.831-145.849 145.861-145.879 [9,10]
JSC VNIIEM Corporation	*Kanopus-V* (22.07.2012)	Kanopus-V [17]	8066.3-8189.7 8258.3-8381.7 [16]	2205.44-2224.06/ 2031.52-2048.23 [16]
JSC VNIIEM Corporation	Universitetskiy – Tatyana-2 (17.09.2009)	UMP-70 [18]	1.7 GHz, [18, 31, 32]	435.3-435.6/ 145.8-146.0 [18, 31, 32]
JSC ISS	Yubileiny-2 (28.07.2012)	Yubileiny [19, 20]	2.4 GHz, [19, 20, 33]	435.3-435.6/ 145.8-146.0 [19, 20, 33]
SRI RAS	Chibis-M (30.10.2011)	Chibis-M [21]	2.27 GHz, [21]	435.3-435.6/ 145.8-146.0 [21]
SRI RAS	Chibis-AI (project)	Chibis-M [21]	2269.5-2270.5, 8395-8400 [10, 22]	435.2065-435.2235 435.3065-435.3235/ 145.804-145.816 145.849-145.861 [10, 22]
MSTU named after Bauman	Baumanets-2 (28.11.2017)	———	8195-8255 94980-95020 [10]	435.395-435.435/ 145.815-145.835 and 145.845-145.865 [10]
SSAU	SamSat-218 (28.04.2016)	CubeSat (3U).	———	145.85-145.89/ 435.59-435.61 [10, 23]

Thus, in the 145-146 MHz and 435-438 MHz radio frequency bands, the sources of interference will increase over time, and, as a result, the electromagnetic environment in the area of operation of ground-based control and reception of information from the SSC will only worsen over time. It is necessary to work out an option of switching to a different frequency range.

As an option, the 258-261 MHz (Earth-to-space) radio frequency bands can be used to build a CTS, while consumers should not cause harmful interference to systems in the mobile-satellite service [13, 24, 25].

The frequency bands used by the RLTI are assigned in accordance with the requirements of the table of the distribution of radio frequency bands of the Russian Federation (given in table 1) and do not require changes.

It should be noted as an unfortunate choice of 2.4 GHz by the Information Satellite Systems JSC for its SSC [19, 20], because the radio frequency band 2300-2450 MHz [13, 24, 25] is the radio frequency band of primary use of radio electronic equipment intended for the needs of state administration, including presidential communications, government communications, the needs of the country's defense, state security and the rule of law, and for fixed, mobile, amateur, radar services. Also, the frequencies 2300-2400 MHz and 2500-2690 MHz or parts of these bands are identified for use by administrations wishing to implement the International Mobile Telecommunications (IMT) in accordance with Resolution 223 (Rev. WRC-15) [13, 26].

In addition, the frequency of 2.4 GHz (Earth-to-space) on a secondary basis [13, 24, 25] can only be used for space communications, subject to the development and coordination with the Russian Ministry of Defense of organizational and technical measures to ensure electromagnetic compatibility with REE of special and military purpose. The above conditions for the use of the 2.4 GHz frequency impose serious restrictions on the possibility of obtaining permission and using this frequency in the construction of space systems and complexes of spacecraft.

Conclusions and suggestions

The issues of radio-frequency support for the functioning of the space complex (system) based on the SSC are normatively rather strictly regulated, but the developers are not fully implemented. Currently used radio frequency bands impose rather stringent restrictions for the construction of the space segments and/or clusters based on the SSC, which will inevitably negatively affect their expanded use. The predicted further increase in the number of launched SSCs will complicate the functioning of command-telemetry systems in the (space-to-Earth) and (Earth-to-space) directions, as well as to the loading of the radio links for the exchange of target information (RLTI) in the (space-to-Earth) direction.

Unfortunately, it should be noted that at present there is practically no necessary coordination and interaction regarding the use of the radio frequency band between civil and military radio services.

Under these conditions, it seems preferable to use the following radio frequency bands:

- 258 ... 261 MHz (Earth-to-space) can be used to build the command and telemetry system of the SSC, while consumers should not cause harmful interference to the systems of mobile and mobile satellite services operating in this frequency band, or require protection from them, or limit the use and development of such systems [13, 24, 25];
- 267 ... 273 MHz (space-to-Earth) is allocated for the space operation service, taking into account that this is the band of radio frequencies of primary use for REE intended for the needs of public administration, including presidential communications, government communications, the needs of the country's defense, state security and law enforcement [13, 24, 25] with restrictions of 5.254, 5.257 [26] and 95, 125, 127 [13];
- 2025 ... 2110 MHz used by the space research service (Earth-to-space, space-to-space), the space exploitation service (Earth-to-space, space-to-space), the satellite earth exploration service (Earth-to-space, space-to-space) with the restrictions 5.392 [26] and 193, 200, 211, 219, 221, 222, 223, 224, 225 [13];
- 2200 ... 2290 MHz used by the space research service (space-to-space, space-to-Earth), the space exploitation service (space-to-space, space-to-Earth), the satellite earth exploration service (space to space, space-to-Earth) [13];
- 8025 ... 8400 MHz for the Earth exploration-satellite service (space-to-Earth), meteorological satellite (Earth-to-space), and fixed satellite (Earth-to-space) taking into account the limitation of 5.462A of the radio frequency distribution table [13].

At the same time, it is necessary to consider abandoning the 8400-8500 MHz space research (space-to-Earth) service radio frequency band, designed for the needs of government, including presidential communications, government communications, the needs of the country's defense, state security and law enforcement [13].

Implementation of control and reception of information from spacecraft of satellite communication systems related to the fixed satellite and mobile satellite services (SSSS Turaiya [27], SPSS Iridium [27], multifunctional system of personal satellite communications Gonets-D1M [28, 30], the advanced low-orbit satellite communications system "Gonets-WEB" [29, 30]) will be extremely complicated due to the unresolved issues of frequency assignments and will require amendments to the relevant regulatory documents.

The recommended option for constructing electronic equipment and ensuring the electromagnetic compatibility of this radio electronic equipment for space systems (complexes) based on the SSC is a combined control system for receiving telemetry and information from onboard payloads operating in the radio frequency band allocated for solving a particular objective in accordance with the table of the distribution of radio frequency bands between the radio services of the Russian Federation.

It is proposed to use the radio frequency ranges [13, 26]: 2025-2110 MHz (space-to-space, Earth-to-space) to solve the objectives of operational optical-electronic surveillance of water areas, water bodies of the earth's surface, emergency situations of various nature, as well as for operational monitoring of the condition of engineering objects, 7190-7250 MHz (Earth-to-space), 8025-8400 MHz (space-to-Earth), 25.5-27.0 GHz (space-to-Earth), 37.5-38.0 GHz (space-to-Earth on a secondary basis), 40.0-40.5 GHz (Earth-to-space), 65.0-66.0 GHz.

For scientific experiments, it is advisable to use the radio frequency ranges [13, 26]: 2025-2100 MHz (space-to-space, Earth-to-space), 7145-7190 MHz (deep space, Earth-to-space), 7190-7235 MHz (Near space, Earth -cosmos), 8400-8450 MHz (Deep space, space-to-Earth), 8450-8500 MHz (space-to-Earth), 22.55-23.15 (Earth-to-space), 31.8-32.3 GHz (Far space, space-to-Earth), 34.2-34.7 GHz (deep space, Earth-to-space), 37.0-38.0 GHz (space-to-Earth), 40.0-40.5 GHz (Earth-to-space), 65.0-66.0 GHz, 94.0-94.1 GHz.

References

1. GOST 53802-2010 *Sistemy i kompleksy kosmicheskie. Terminy i opredeleniya [Space systems and complexes. Terms and Definitions]*. Moscow, Standardinform, 2011, 27 p. (in Russian)
2. GOST R 56526-2015 *Trebovaniya nadezhnosti i bezopasnosti kosmicheskikh sistem, kompleksov i avtomaticheskikh kosmicheskikh apparatov edinichnogo (melkoseriynogo) izgotovleniya s dlitel'nymi srokami aktivnogo sushchestvovaniya [Requirements of reliability and safety of space systems, complexes and automatic spacecraft of a single (small-scale) production with long periods of active existence]*." Moscow, Standardinform, 2015, 50 p. (in Russian)
3. *Small Spacecraft Technology State of the Art*. Mission Design Staff Ames Research Center, Moffett Field, California, 2014, 200 p. (NASA / TP – 2014–216648 / REV1).
4. *Small Spacecraft Technology State of the Art.* Mission Design Division Ames Research Center, Moffett Field, California, 2015, 168 p. (NASA / TP – 2015–216648 / REV1).
5. *Characteristics, definitions and spectrum requirements of nanosatellites and picosatellites, as well as systems composed of such satellites.* Electronic Publication Geneva, ITU, 2014, 15 p. (Report ITU-R SA.2312-0 (09/2014)).
6. *Federal Law of the Russian Federation* from July 7, 2003 No. 126-FZ "On Communications" with amendments and additions dated August 4, 2017. (in Russian)
7. *Decree of the State Committee of Radio Frequencies* dated December 20, 2011 No. 11-13-01 "Procedure for reviewing materials and making decisions on allocating radio frequency bands, re-issuing decisions and making changes to them". (in Russian)
8. *Decree of the State Committee for Radio Frequencies* dated December 20, 2011 No. 11-13-02 "The procedure for the examination of the possibility of using the declared radio-electronic means and their electromagnetic compatibility with the existing and planned radio-electronic means, reviewing materials and making decisions about assignment of radio frequencies or RF channels within allocated radio frequency bands." (in Russian)
9. *Decree of the State Committee for Radio Frequencies* dated October 16, 2015 No. 15-35-09-4. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=40711 (Last accessed: 01.20.2019). (in Russian)
10. *Decree of the State Committee for Radio Frequencies* dated February 29, 2016 No. 16-36-11-4. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=41867 (Last accessed: 01.20.2019). (in Russian)
11. OST-134-1020-2008 *Space systems and complexes. Terms and Definitions*. Moscow, TsKBS, FSUE Central Research Institute of Mechanical Engineering, 2008, 56 p. (in Russian)
12. Appendix No. 12 to the *Decree of the State Committee for Radio Frequencies* dated February 29, 2016 No. 16-36-11-4. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=41867 (Last accessed: 01.20.2019). (in Russian)
13. *Table of distribution of radio frequency bands between the radio services of the Russian Federation*, Decree of the Government of the Russian Federation of December 21, 2011 No. 1049-34, Extract, 98 p. (in Russian)
14. *Decree of the Government of the Russian Federation* of November 14, 2014 No. 1194 On the international legal protection of the assignment of radio frequencies or radio frequency channels and the procedure for using satellite communications networks in the Russian Federation under the jurisdiction of foreign states, as well as on introducing changes to some acts of the Government of the Russian Federation. With amendments and additions dated August 15, 2017. (in Russian)
15. Kirilin A.N., Akhmetov R.N., Shakhmatov E.V., Tkachenko S.I. et al. *Opytno-tekhnologicheskiy malyy kosmicheskiy apparat* «AIST-2D» [Experimental technological small spacecraft AIST-2D]. Samara, SamNTs RAS, 2017, 324 pp. (in Russian)

16. *Decree of the State Committee for Radio Frequencies* dated April 28, 2008 No. 08-24-05-009. URL: http:// https://digital.gov.ru/ru/documents/4035/#tag5 (Last accessed: 01.20.2019). (in Russian)
17. Makridenko L.A., Volkov S.N., Gorbunov A.V., Justrin O.S., Ilyina I.Yu. *Istoriya sozdaniya malykh kosmicheskikh apparatov «KANOPUS-V» № 1 i belorusskogo KA*]. URL: https://www.roscosmos.ru/media/files/docs/3/kanopus.pdf (Last accessed: 17.10.2018). (in Russian)
18. Volkov S.N., Makridenko L.A., Khodnenko V.P. Malye kosmicheskie apparaty NPP VNIIEM. Ot kontseptsii do voploshcheniya v «metalle» [Small spacecraft NPP VNIIEM. From concept to embodiment in the metal]. *Voprosy elektromekhaniki* [Questions of Electromechanics]. 2011, V. 121, No. 2. pp. 3-8. (in Russian)
19. Zimin I. I., Valov M. V., Yakovlev A. V., Galochkin S. A. Malyy kosmicheskiy apparat «Mikhail Reshetnev». Rezul'taty raboty [The small spacecraft Mikhail Reshetnev. Results of work]. *Trudy MAI* [Proceedings of MAI]. 2013, No. 65., 10 p. URL: http://www.mai.ru/science/trudy/published.php?ID=35908 (Last accessed: 10.17.2018). (in Russian)
20. Zimin I. I. et al. Malyy kosmicheskiy apparat «Mikhail Reshetnev». Rezul'taty raboty [Small spacecraft "Mikhail Reshetnev". Results of work]. *Aviatsiya i kosmonavtika* [Aviation and Cosmonautics]. 2012, 11th Intern. conf. (November 13-15, 2012, Moscow). St. Petersburg, Masterskaya pechati, 2012, p. 97. (in Russian)
21. Zelenyy L.M. et al. Akademicheskiy mikrosputnik ChIBIS-M [Academic microsatellite CHIBIS-M]. *Kosmicheskie issledovaniya* [Space research]. 2014, V. 52, No. 2. pp. 93-105. (in Russian)
22. Appendix No. 6 to the *Decree of the State Committee for Radio Frequencies* of February 29, 2016 No. 16-36-11-4. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=41867 (Last accessed: 01.20.2019). (in Russian)
23. Appendix No. 6 to the *Decree of the State Committee for Radio Frequencies* of July 4, 2017 No. 17-42-08-2. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=46587 (Last accessed: 01.20.2019). (in Russian)
24. Appendix No. 2 to the *Decree of the State Committee for Radio Frequencies* of July 15, 2010 No. 10-07-01. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=511 (Last accessed: 01.20.2019). (in Russian)
25. Appendix to the *Decree of the State Committee for Radio Frequencies* of October 16, 2015 No. 15-35-02. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=40675 (Last accessed: 01.20.2019). (in Russian)
26. *Reglament radiosvyazi* [Radio Regulations]. Articles. 2016 edition, ITU, 442 p. URL: http://search.itu.int/history/HistoryDigitalCollectionDocLibrary/1.43.48.ru.601.pdf (Last accessed: 01.20.2019). (in Russian)
27. *Decree of the State Committee for Radio Frequencies* of October 2, 2012 No. 12-15-05-7. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=607 (Last accessed: 01.20.2019). (in Russian)
28. *Decree of the State Committee for Radio Frequencies* of February 29, 2016 No. 16-36-11-4. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=41867 (Last accessed: 01.20.2019). (in Russian)
29. *Decree of the State Committee for Radio Frequencies* of May 24, 2013 No. 13-18-06-5. URL: http://www.grfc.ru/grfc/norm_doc/resheniya-gkrch/?PAGEN_1=6 (Last accessed: 01.20.2019). (in Russian)
30. *Decree of the State Committee for Radio Frequencies* of November 7, 2016 No. 16-39-05-2. URL: http://www.grfc.ru/grfc/norm_doc/detail/?ID=44061 (Last accessed: 01/20/2019). (in Russian)
31. Makridenko L.A., Volkov S.N., Gorbunov A.V., Kozhevnikov V.A., Khodnenko V.P. Malyy kosmicheskiy apparat «Universitetskiy – Tat'yana-2» [Small University-Tatiana-2 spacecraft]. *Voprosy elektromekhaniki* [Questions of Electromechanics]. Proceedings of VNIIEM, Vol. 160, 2017, pp. 45-54. (in Russian)
32. *Tatiana-2 / Universitetsky-2* URL: https://directory.eoportal.org/web/eoportal/satellite-missions/t/tatiana-2 (Last accessed: 01.01.2019). (in Russian)
33. *Yubileiny-2 / MiR (Mikhail Reshetnev) Microsatellite* URL: https://directory.eoportal.org/web/eoportal/satellite-missions/v-w-x-y-z/yubileiny-2 (Last accessed: 01.01.2019). (in Russian)
34. Talanov, A.A., Fedotov, S.A., Stepanov, A.M. redlozheniya po razrabotke unifitsirovannykh kosmicheskikh platform malykh kosmicheskikh apparatov [Proposals for the development of unified space busses for small spacecraft]. *Raketno-kosmicheskoe priborostroenie i informatsionnye tekhnologii* [Rocket and space device engineering and information technology]. 2018. Proceedings of the IX All-Russian Scientific and Technical Conference "Actual Problems of Rocket and Space Instrumentation and Information Technologies" (June 5-7, 2018). Ed. by Doctor of Technical Sciences, prof. A.A. Romanov. Moscow, JSC Russian Space Systems, 2018, Pp. 470-482. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 52–60

Concept of Constructing a Technological Model for Solving Semi-structured Problems on the Basis of Set Theory

V. V. Betanov, *Dr. Sci. (Engineering), Prof., betanov_vv@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
V. K. Larin, *Cand. Sci. (Engineering), vklarin@mail.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation

Abstract. This article discusses one of the approaches to solving semi-structured problems using the apparatus of set theory. The problem is defined as a system of inhomogeneous elements, the structural analysis of which allows identifying incorrect elements. The further solution is presented in two stages. At the first stage, a block diagram of the solution of the problem is developed with the last level of the hierarchy detailed to the maximum. At the second stage, the description of the structural blocks is given in symbols of set theory. As an example, the problem of processing trajectory measurements is considered. The incorrect blocks of the problem are pinpointed and the technology of the general solution is given taking into account the representation of these blocks in symbols of set theory.

Keywords: systems approach, structuring, system, sets, model

Introduction

One of the problems related to the field of systems analysis is the search for a solution to the problem of semi-structured problems [2, 4]. Solving this problem is complicated by the absence of boundaries that would establish the belonging of tasks to the semi-structured type, as well as by the lack of specific criteria for determining the degree of structuredness. Allowing that semi-structured problems are located in the range from structured to non-structured problems, it can be assumed that the solution–result may vary from quantitative to qualitative and, in some cases, it may not be found.

The solution of the problem, virtually, is reduced to the identification of incorrect elements in the software and algorithmic shell, to operation monitoring and, in the case of failure, to transferring the solution to another chain of the algorithm.

The present paper describes one of the possible approaches to solving a problem consisting in the structural analysis of the algorithm and the subsequent representation of the incorrect elements of the solution in the form of set theory procedures.

The term “problem” (or task) is understood as a system consisting of elements, which ensure the reception of an answer to a posed question.

Preliminary notes

There are three ways of representing a solution to a problem: an algorithm, a block diagram and a verbal description.

1. Algorithm. It is used for formalized problems that a have a mathematical solution.

2. Block diagram. It is implemented for displaying the procedure of solving problems in the form of a hierarchical sequence of interrelated functional blocks.

Structural analysis of the block diagram allows you to mark the blocks capable of leading to an incorrect decision; however, it is impossible to determine the specific function or parameter, which is the source of this situation.

3. Verbal description. It is acceptable for the first and second variants of problem structuredness mentioned above but it is only presents a qualitative picture of the solution without the possibility to formally adjust the process.

In the second and third types of problem representation, separate elements may be of a non-formal nature, which makes it difficult to construct programs using modern programming languages and, consequently, hinders the process of automating the decision-making process.

The essence of the method lies in the use of logical conditions for determining the truth (or falsity) of the set of block elements by matching one of its parameters to a specified value. This permits to establish relationships between blocks of the solution of the problem, which is presented in structural or verbal forms.

The problem solution is divided into two stages. At the first stage, a block diagram (flowchart) of the problem solution is developed with maximum detailing of the last level of the hierarchy. At the second stage, the structural blocks are described using symbols of set theory.

Based on a sufficient number of works dedicated to the study of problem structuredness and their solution [2-6], it can be said that problems are divided by the degree of structuredness into: structured (SP), semi-structured (SSP) and non-structured tasks (NSP).

The quantitative and/or qualitative **definiteness** of the elements of the problem is used as a criterion for classifying elements of the problem. Another criterion for dividing problematic situations into structured, semi-structured and non-structured is the extent the algorithm of their solution is known [2]. Considering the abovementioned, the following descriptions can be given to each problem type.

SP: are characterized by the presence of a solution algorithm based on mathematical relationships as well as by a quantitative result.

SSP: are known for their qualitative relationships between elements of the problem, information about part of the elements may be absent.

NSP: are characterized by the absence of a mathematical solution algorithm and by the partial absence of conditions and source data.

It is appropriate to add quasi-structured problems (QSP) to the given list of problem types determined by the level of problem structuredness. These problems are characterized by the presence of quantitative and qualitative constituent elements with a probability of failure of the latter during the problem resolution process but with a possibility of finding a quantitative or qualitative solution. This problem type is situated between SP and SSP.

Graphically, this distribution of tasks by the degree of structuredness is given in Figure 1.

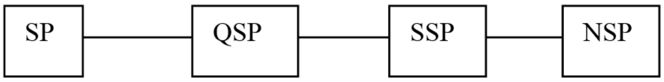


Fig.1. Distribution of tasks according to degree of structuredness.

A structured problem was taken as a baseline variant (rating) for conducting the study.

The level of structuredness of a task was determined by the ratio of correct and incorrect properties inherent in its constituent blocks.

Depending on the functional load of the elements comprising the solution, there can be three forms of representing the solution:

- software (SW; program-algorithmic);
- information computing (IC);
- hardware/technology (HT).

Each form of representing a task is comprised of the following elements:

- software and analytical (SA): statement + algorithm + program;
- information computing element: SA + database (DB) and archives;
- hardware and technology elements: IC + computer + data path for inputting measurements into DB.

Regardless of belonging to one or another form, a problem has the same set of elements given below.

Elements of a problem

Functional parts of the problem

1. method — statement, algorithm;

2. software — operating system (OS), programming language, program codes, standard subroutine library (SSL);

3. calculation — computing procedure (processor: type, frequency), random access memory.

Data part:

1) DB, archives, data exchange (communication);

2) interface — input and output data, control of intermediate computations.

Stages of solving the problem:

1) development of solution structure, method and algorithm;

2) program development and performing calculations.

Ways to solve the problem:

1) formal solution (algorithm and software development);

2) informal solution (search for an alternative solution);

3) semi-structured solution (in the case of a failure possibility) – divided into a “quality” or “quantity” solution, depending on the ratio of formalized and non-formalized elements of the problem as well as on the real conditions of the solution.

Data description formats:

1) quantitative;

2) qualitative;

3) set-format (sets and operations with them).

Description of problem solution model:

1) mathematical algorithm;

2) architecture (block diagram);

3) description of the sequence of the stages of solving the problem;

4) technological model in the form of a description of problem structural blocks and their connections using set symbols.

As an example, the article considers a solution to the problem of processing trajectory measurements of the ballistic-navigation support (BNS) with respect to the software (program-algorithmic, SW) form.

The majority of tasks concerning measurement processing are structured tasks. Yet, in the presence of three incorrect elements in the structure of such a problem, its type may change.

The specific features of such problems are:

— a large volume of measurement data saved in the source data (SD);

— the presence of a restriction on the quantity and quality of measurements in the algorithm;

— the main measurement processing method is statistical (least-squares method (ordinary least squares), OLS, Kalman filter);

— the use of a mathematical model in the algorithm for every measurement type (range, phase, Doppler frequency and etc.); and

— the implementation of the following mathematical functions: calculation of partial derivatives, composition and solution of systems of linear equations, formulas of calculated values of measured parameters.

Given below are the results of analyzing a problem for determining the correctness of solution blocks.

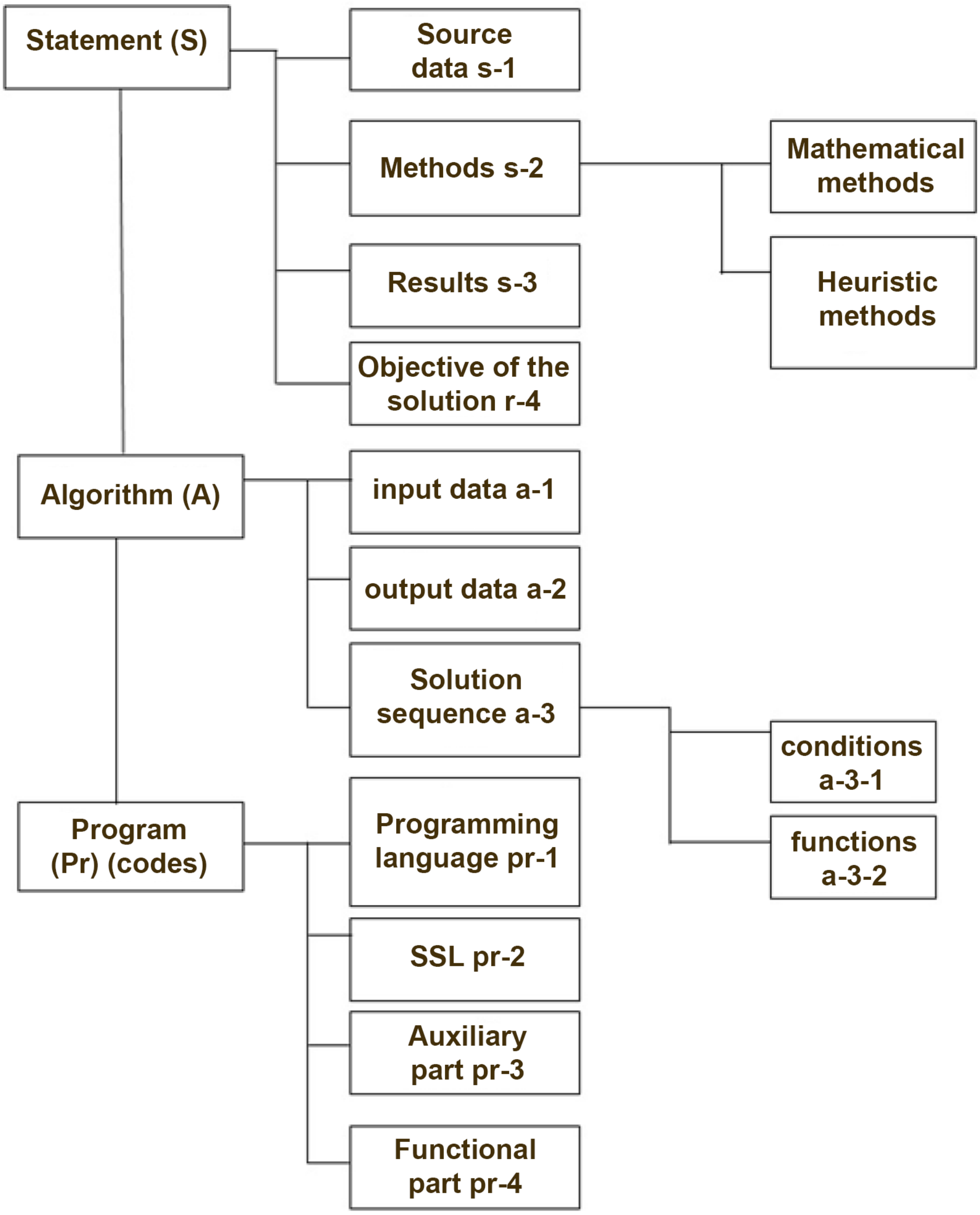


Fig.2. Block diagram of problem in SW form.

Problem structure

First, we will set forth a block diagram of a problem in SW form (Figure 2).

The auxiliary part includes:

- source data processing and generation
- access to DB
- access to archives
- access to SSL
- loop organization (looping)
- conditional and unconditional branching between program parts, and
- generation of output data.

The functional part includes:

- a mathematical model of the considered physical process;
- iterative calculation methods
- control of correspondence of mathematical formulas to the physical process being described
- methods of interpolation and extrapolation
- numerical methods for solving differential equations.

Based on the distribution of problems according to the degree of structuredness and to the block diagram (Figure 2), Table 1 gives data on the correctness (+) and incorrectness (-) of the blocks distributed across problems of varying degrees of structuredness and its constituent parts.

Table 1. Distribution of problems according to the degree of element structuredness

Parts of the problem	Symbol for blocks of problem parts	Types of problems in accordance to their structuredness			
		SP	QSP	SSP	NSP
Statement	s1	+	+	-	-
	s2	+	+	+	-
	s3	+	+	-	-
	s4	+	+	+	-
Algorithm	a1	+	+	-	-
	a2	+	+	-	-
	a3-1	+	-	-	-
	a3-2	+	-	-	-
Program	pr1	+	+	+	+
	pr2	+	+	+	+
	pr3	+	+	+	+
	pr4	+	-	-	-

Table 2. Dependence of solution on problem type

Type	Solution		
Problem type	Quantitative	Qualitative	0
SP	+	-	-
QSP	+	+	-
SSP	-	+	+
NSP	-	-	+

Table 2 gives classification data on the distribution of types of problem solutions depending on the degree of problem structuredness, where "+" is the correctness of the solution of the corresponding problem type; and "-" is the problem solution incorrectness (the absence or the non-compliance with the conditions of normal solution functioning). A quantitative solution is a result in numerical format, a qualitative solution is a result in verbal form; 0 – is the lack of options for solution continuation: obtained result does not correspond to mission in terms of physical sense.

Data from Tables 1 and 2 allow you to define the degree of structuredness of the problem, determine its place in the proposed classification, as well as refine the incorrect blocks in the main elements of the problem.

The received information makes it possible to preselect the appropriate solution method with account for the possible incorrectness of separate parts of the problem.

Block diagram of the problem

Given below is a block diagram (flowchart) of the functional blocks of the problem of measurement processing, where blocks with a probability of an incorrect solution are marked (in yellow) (Fig. 3).

The apparatus of set theory is proposed to be used for the transition from a structural display of problem elements and their relationships to a formalized form.

In this case, each block (which is a system) can be identified as a set of heterogeneous elements. The relations between problem elements (Figure 3) can be represented in the form of a corresponding relationship between the parameters of sets [7].

The top level of structural blocks:

p = s;

S = std SD = s1

A = a1

R =

Let us represent the conditions for the selected structural blocks in the form of the following relations.

Initial data:

s1 = { φ , $\boldsymbol{\gamma}$};

s1 = {$_i$, measurement limitation for $_i$<;

s1 = {k , limit on the number of measurements per session.

Formulas

a1 = { , establishing the correctness of the formulas.

Matrices:

, condition for using matrices in the algorithm

where Mt are the transition matrices of coordinate systems;

Ms are the matrices for solving systems of linear equations.

Composition of sought data:

p1 = {Q, n},

Q – is the form (type) of data representation (for example, orbit parameters are osculating elements);

n – is the number of data (for example, a kinematic vector of state – 6 values)

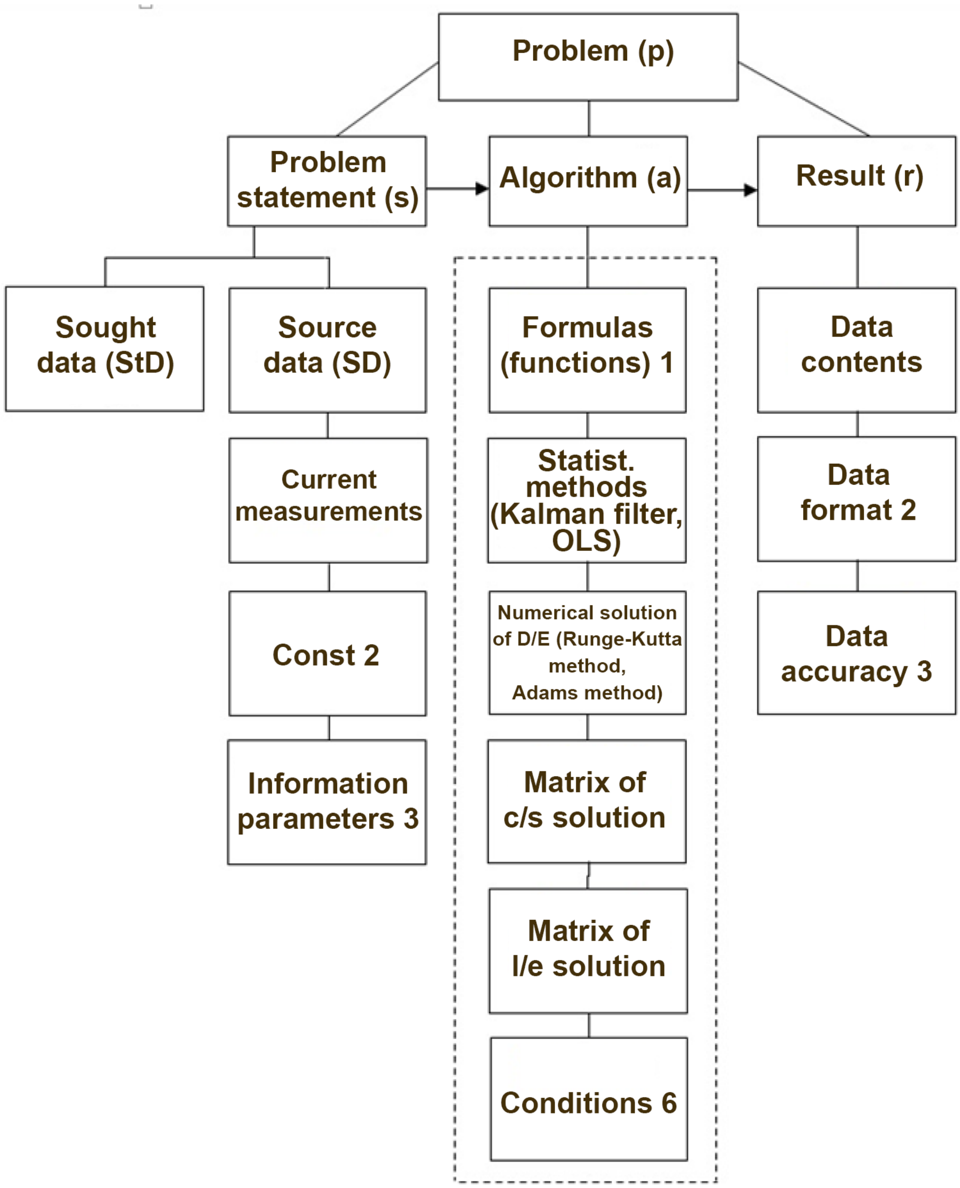


Fig. 3. Diagram of functional blocks of measurement processing: d/e – differential equation, l/e – linear equation, c/s – coordinate system.

$Q = \{Q_i : Q_j\}$, the condition for choosing the type of result corresponding to the given one.

$Q_i = \{s = n\}$, match condition of the number of data.

Accuracy of sought data:

$s3 = \{\sigma , v\}$, root-mean-square deviation (RMSD) and the number of significant digits after the decimal point.

$s3 = \{ \sigma_i : \sigma_n\}$, condition for checking the result for accuracy by RMSD.

$s3 = \{ v_i : v_n\}$, condition for checking the result by the number of digits after the decimal point.

$s = \{ s$, condition for choosing result by the fulfillment of condition p3.

The first stage of solving the problem

As an example of implementing the proposed method, we will consider the problem of refining the initial conditions of spacecraft (SC) motion using trajectory measurements. The problem solution is divided into two stages: at the first stage, the solution will be given in form of a flowchart (block diagram) (Figure 4); at the second step, it will be given in the form of an algorithm in the language of set theory.

1[st] stage of solution

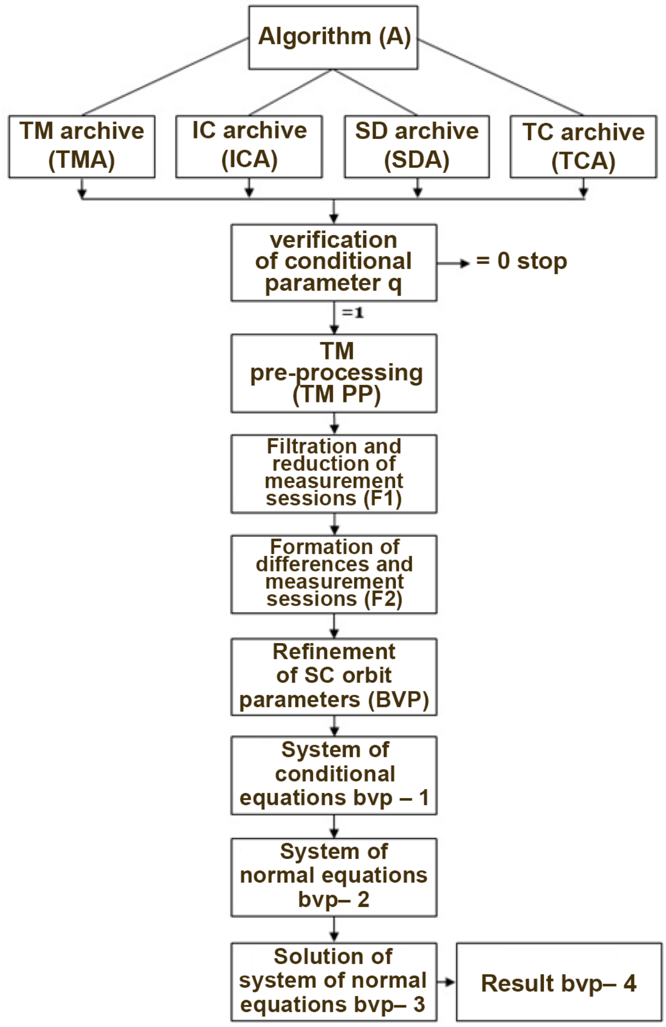


Fig. 4. Flowchart (block diagram) of solution for problem of refining initial conditions.

We will give explanatory notes to Figure 4.

TM archive – is the file archive of trajectory measurements. The principal format for TM are Rinex-files.

IC archive – is the file archive of SC initial conditions obtained at previous BNS-session. The main form: date, time and parameters of the kinematic state vector in the Greenwich coordinate system (GCS);

SD archive – is the archive of source data. The main contents: global constants of the parameters of Earth's gravitational field, Earth rotation parameters, etc.

TC archive – archive of technical characteristics of the spacecraft. The main contents: SC number, SC type, time delay of signal relay, nominal values of carrier values of navigation radio signals from onboard generators;

BVP – is the boundary value problem - the generally accepted name of the task of refining the ICs. Below, the blocks that make up the BVP belong to the main stages of processing measurements with the help of OLS;

Result – are the refined values of the components of the SC kinematic vector of state in GCS at the start of measurement.

The second stage of solving a problem

Every block of the aforementioned flowchart (block diagram) is a set with heterogeneous elements. Therefore, the transition to the form of a solution in symbols of set theory is the first approximation to the final form of the algorithm for the further development of the program. This example allows us to talk about the fundamental possibility of such an approach to the solution of semi-structured problems.

A = {TMA ;

TMA = ICA = SDA = TCA=

Conditions of the correctness (validity) of data in the archives are checked.

Let us introduce conditional parameter *q*. If the correctness conditions *q* of these archives are met, then *q* takes on the value of 1, otherwise – 0.

TMA = {}

TMA = {$_i$, limitation on measurements for $_i$<, where **φi** are measurements.

TMA = {k, limitation on the number of measurements in session N.

ICA = {, limit on the value of RMSD.

SDA = {$v_i v_n$}, limitation of data by the number of digits after the decimal point.

TCA = { correspondence of TCs to nominal values.

After checking each archive, depending on the outcome of the procedure, the conditional parameter *q* was assigned a corresponding value (0 or 1).

The value of the conditional parameter is checked.

If q = {= 1}, then the action goes to the block "TM preliminary processing (TM PP)";

If q = {= 0}, the solution comes to a **stop**.

TM PP = {f1 ;

where f1 = {y1,y2},

y1 = {$_i$, limitation of $_i$< measurements.

y2 = Δt = {(t_i- t_{i-1}) : δt}, condition for reduction of measurement sessions;

t_i , t_{i-1} – are the start times of two adjacent measurements, δt is the given (specified) time interval.

f2 = {y3,y4},

y3 = {Di = D_3}, limitation on the length of baselines *Di*, (D_3 is the given length).

y4 = {k , limitation on the number of measurements in session *N*.

We shall consider the block for refining the parameters of the orbit (BVP).

BVP = {bvp1 bvp;

The general form of the conditional equation r_{meas} – r_{est}= Δr, the number of equations is equal to the number of measurements *s*.

BVP1 = {n = s};

The system of normal equations can be represented as follows

$$a_{11}P1 + a_{12}P2 + \ldots + a_{1k}P_n = A_{11}$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{k1}P1 + a_{k2}P2 + \ldots + a_{kk}P_n = A_{kk}$$

The number of equations in the system should be equal to the number of parameters being refined.

bvp2 = {n : k};

The general form of the solution of a system of normal equations:

[p] = B, B = H^{-1} – is the product of the inverse matrix of partial derivatives for the measured values with respect to the sought parameters by the correlation matrix of the measurement vector.

bvp3 = [p] =, matrix H must not be equal to 0, otherwise we will have division by 0.

The result of the solution of the BVP is a state vector refined by the corrections received by solving a system of normal equations.

= $_0$ + p

where – is the refined state vector of the SC;

$_0$ – is the initial value of the state vector of the SC;

p – is the vector of corrections as a result of solving the BVP.

: }, is the correspondence of the obtained solution to the given (set) type (format, accuracy).

The presented two-step technological model for solving semi-structured problems cannot pretend to be a complete algorithm but is can greatly facilitate the search for a solution due to the overcoming of situations that arise from the incorrectness of separate elements in the blocks of the problem in a formalized manner.

Conclusion

The following conclusions can be made based on the materials in the present article.

1. A technological model for the solution of semi-structured problems has been developed based on set theory.

2. The technological model is implemented by sequentially performing the following steps:

- development of a block diagram (flowchart) of the problem solution;
- allocation of incorrect (invalid) blocks;
- representation of the conditions of incorrectness in symbols of set theory.

3. The creation of a complete algorithm for the solution of semi-structured problems is possible by the introduction of conditions of incorrectness for its separate blocks into the sections of the mathematical algorithm using set theory symbols.

4. The use of the basic concepts of set theory during the development of a problem algorithm allows us to formalize the structural and verbal description of the incorrectness of its separate elements, which provides the possibility to form a formal solution to the problem.

References

1. Betanov V.V., Larin V.K. Ispol'zovaniye sistemnogo podkhoda k resheniyu problemnykh voprosov funktsionirovaniya AKP BNO poletov KA GNSS [Using a Systematic Approach to solving the problematic issues of functioning of the automated complex of programs for ballistic and navigational support of GNSS spacecraft missions]. *Raketno-kosmicheskoe priborostroenie i informatsionnye sistemy* [Rocket-Space Device Engineering and Information Systems]. Moscow, 2016, Vol. 3, No. 1, pp. 3–10. (in Russian)
2. Betanov V.V., Larin V.K. Kontseptsiya postroyeniya ekspertno-diagnosticheskogo kompleksa dlya analiza informatsionnykh sistem [The concept of building of an expert-diagnostic complex for the analysis of information systems]. *Raketno-kosmicheskoe priborostroenie i informatsionnye sistemy* [Rocket-Space Device Engineering and Information Systems]. Moscow, 2018, Vol. 5, No. 2, pp. 65–72. (in Russian)

3. Betanov V.V., Larin V.K. Kontseptsiya gibridnoy tekhnologii ballistiko navigatsionnogo obespecheniya nazemno-kosmicheskoy svyazi v GAS RF "Pravosudiye" [The concept of the hybrid technology of ballistic-navigational support for ground-space communications in the GAS RF "Justice"]. Pravovaya informatika, 2018, No. 2, pp. 39–46. (in Russian)
4. Tarasenko F.P. *Prikladnoy sistemnyy analiz: Uchebnik* [Applied systems analysis: textbook]. Tomsk, Izdatel'stvo Tomskiy universitet, 2004, 186 p. (in Russian)
5. Betanov V.V., Larin V.K., Pozyaeva Z.A. Prototip ekspertnoy diagnosticheskoy sistemy poiska i korrektsii skachkov v bezraznostnykh fazovykh izmereniyakh [The Prototype of an Expert Diagnostic System for Cycle Slip Detection and Correction]. *Raketno-kosmicheskoe priborostroenie i informatsionnye sistemy* [Rocket-Space Device Engineering and Information Systems]. Moscow, 2014, Vol. 1, No. 3, pp. 73–81. (in Russian)
6. Larin V.K. Postroyeniye prototipa ekspertno-diagnosticheskoy sistemy analiza trayektornoy izmeritel'noy informatsii [Building of a prototype of expert-diagnostic system for the analysis of spacecraft flight path measurement information]. *Raketno-kosmicheskoe priborostroenie i informatsionnye sistemy* [Rocket-Space Device Engineering and Information Systems]. Moscow, 2017, Vol. 4, No. 1, pp. 53–60. (in Russian)
7. Kuznetsov O.P. *Diskretnaya matematika dlya inzhenerov* [Discrete mathematics for engineers]. Saint Petersburg, Lan', 2009, 400 p. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 61–70

Algorithm to Analyze Spectral Characteristics of Snow and Cloud Cover Based on MSU-MR/Meteor-M No. 2 Data

K.I. Zubkova, *contact@spacecorp.ru*
Joint Stock Company “Russian Space Systems”, Moscow, Russian Federation
L.A. Grishantseva, *Cand. Sci. (Phys.-Math.), contact@spacecorp.ru*
Joint Stock Company “Russian Space Systems”, Moscow, Russian Federation
T.G. Kurevleva, *contact@spacecorp.ru*
Joint Stock Company “Russian Space Systems”, Moscow, Russian Federation
A.A. Skripchuk, *contact@spacecorp.ru*
Joint Stock Company “Russian Space Systems”, Moscow, Russian Federation
A.A. Morozov, *contact@spacecorp.ru*
Joint Stock Company “Russian Space Systems”, Moscow, Russian Federation

Abstract. Prompt acquisition of up-to-date weather data, particularly in the form of snow cover maps and cloud maps, is one of the components providing information support for target use of the orbital remote sensing spacecraft constellation. The planning complex, in addition to weather forecast calculation facilities, is currently equipped with an information database containing MSU-MR/Meteor-M No. 2 images (Gorbunov et al. 2015) in the form of composite images in three shortwave channels (with no theme-based processing). In case the data is insufficient for making decisions, then an operator additionally analyses the data from foreign satellites not included in the planning complex. Performance optimization calls for the automation of weather support processes. The paper proposes an algorithm of generating a snow cover mask and cloud mask using domestic MSU-MR/Meteor-M No. 2 data. The algorithm is based on the spectral characteristics analysis of different types of underlying surface: ground, clouds, and snow. The algorithm uses the reflectance values in channel 1 of MSU-MR (green, 0.63 μm) and in channel 3 of MSU-MR (SWIR, 1.7 μm). The proposed algorithm differs from its foreign counterparts: it employs threshold values dependent on the values of reflectance in channels 1 and 3 rather than constant thresholds. The algorithm was verified across 459 survey routes. Testing confirmed the versatility of the proposed algorithm: thresholds do not depend either on geographical location or on seasons. Considering the continuity of imagery from MSU-MR, this algorithm is capable of producing snow cover and cloud maps that are automatically updated with a periodicity of 2-3 days. This will enable the planning complex to be provided with up-to-date snow cover data.
In the nearest future, the creation of the program module realizing this algorithm and the validation of the possibility of its practical application using means of the simulation stand created within the R&D “Earth remote sensing operator” is planned.

Keywords: remote sensing, snow cover, cloud cover, MSU-MR, Meteor-M

Introduction

One of the directions of modernization of the ground-based and space-based infrastructure of Earth remote sensing (ERS) of priority within the framework of the implementation of measures of the Federal Space Program for 2016-2025 is the automation of technologies of the target use of the orbital constellation (OC) of ERS spacecraft (SC).

The process of target use includes such stages of working with ERS data as the generation of an order for satellite imagery, survey planning, reception of satellite information, its standard and thematic processing as well as the issue of the final product to the customer. Planning a survey session is one of the more resource-intensive aspects of target use.

One of the important constituents of comprehensive planning of satellite surveys is meteorological support. The main purpose of weather forecasting is to assist in deciding on the appropriateness of planning a survey at a given location at a given time as well as in determining the input parameters for exposure calculations. In particular, the presence or absence of snow cover significantly affects the estimate of the time required for accumulating the signal on the planned survey route.

The operators of the planning service use RGB composites of satellite images in the form of global coverages generated at short intervals – ranging from one to two-three days. Currently, the planning complex includes the image of global coverage based on the data from the MSU-MR/Meteor-M No. 2 [1] in the daytime. Contours of the routes being planned are superimposed on the RGB images of the MSU-MR. Additional products involved in the process of assessing the presence of snow cover at the survey location are the snow cover mask provided by data from the AVHRR/MetOp-A, B [2] and daily RGB composites of MODIS images in channels 1-4-3 [3]. The operator makes a decision about the presence or absence of snow cover and/or cloudiness after conducting a visual analysis of the global images at the location of the planned survey. However, the processing of one route is time-consuming (up to several minutes). This significantly slows down the operating process (especially in conditions of a large number of orders) and increases the complexity.

The ultimate goal of working with MSU-MR data is automated generation of regularly updated masks of the snow and cloud covers, directly built into the planning complex. At this stage, an algorithm for the formation of such a mask using MSU-MR/Meteor-M No. 2 has already been created and validated.

The basis for the discrimination of snow and clouds from other types of underlying surfaces are the spectral properties of these objects. Snow and clouds are characterized by elevated values of reflectance (R-values; spectral brightness coefficient) in the visible and near IR-ranges [4-7], by a significant decrease of the reflectance in the short-wave IR-range (SWIR – Short-Wave Infra-Red) and by low values of the radiation temperature in the far IR-range (LWIR — Long-Wave Infra-Red). In the MSU-MR, channel 1 (green 0.6 μm) operates in the visible range; channel 2 (0.8 μm) operates in the near IR-range, channel 3 (1.7 μm) – in the SWIR-1 range; channel 4 (3.8 μm) – in the SWIR-2 range; two channels operate in the far IR-range, which are channels 5 (11 μm) and 6 (12 μm).

Snow and clouds are distinguished reliably from other objects by the contrast of brightness (reflectance) in the visible and near IR-ranges. Nevertheless, the task of separating these objects from each other cannot be fulfilled without resorting to the use of SWIR-range data. In the channels of the visible and near IR-ranges, the R-values of snow and clouds are very close, whereas in channels of the SWIR-range there is a contrast between the R-values of these objects [4, 5]. It is this feature that is used for calculating the NDSI snow index (Normalized Difference Snow Index): NDSI = (Green–SWIR-1)/(Green + + SWIR-1) [8]. Yet, it should be taken into consideration that the use of the NDSI threshold value does not guarantee an unequivocal differentiation between snow and clouds. High values of the NDSI may correspond not only to snow-covered areas but also to so-called ice clouds. The final decision on the type of underlying surface is made after comparing the values of the radiation temperatures in the 11 μm channel with the model average climatic values: if the difference of temperatures significantly differs from 0 K, then the object is considered to be a snow cloud.

In [6], a threshold algorithm for singling out snow according to AVHRR data is given. A block diagram of the algorithm is shown in Figure 1.

The following constant threshold values are used in the algorithm: radiation temperature in channel 4 T4, difference in radiation temperatures in channels 4 and 5 T45, NDVI [9], difference in temperatures in channels

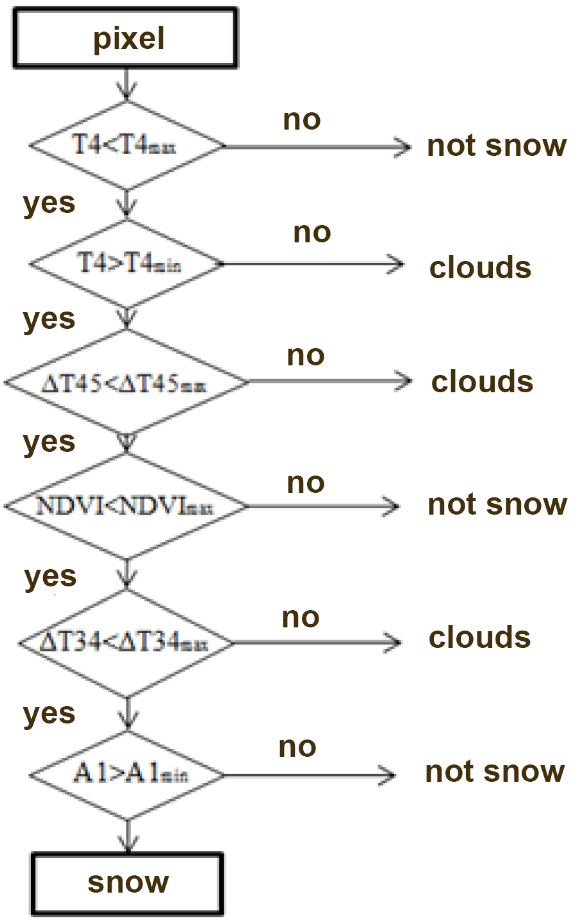


Fig. 1. Algorithm for detecting snow cover with the application of AVHRR [6] data.

Moreover, threshold values of the Normalized Difference Vegetation Index (NDVI) are used for forest-covered territories because the values of NDSI are lower for forest areas in comparison to those of a surface without vegetation.

Figure 2 gives a scattergram [5] of the distribution of R-values in channels 0.6 μm and 1.6 μm of AVHRR/NOAA-17 [11] obtained for the survey scene over the Alps. An open surface (clear), snow and clouds (cloudy) are present in the scene.

Mixed cells are present at the boundaries dividing the clusters of the abovementioned three types of surfaces in the scattergram.

The cell indicated by a yellow arrow is given as an example of such a situation: out of six points in a cell, three correspond to cloudiness, two – to an open surface and one – to snow cover. The sections of the image corresponding to the pixels in the mixed cell can be characterized as cloudless only with a probability of 50%. Note that the arrangement of clusters corresponding to the three types of surfaces (clouds, snow and open surface) indicates that any threshold values of the reflectance in the 0.6 μm and 1.6 μm will inevitably lead both to false alarms (commission errors) and gaps (omission errors) in the masks of snow or cloud cover.

3 and 4 T34, albedo (R) in channel 1 A1. The threshold values were obtained empirically only for one region (Eastern Canada) and for the time of year – early spring.

The high reliability of the results can be noted: the probability of the correct detection of snow reaches 97%. The snow mask is verified with the implementation of a large amount of ground data. Yet, the authors themselves position their algorithm as a regional one; the threshold levels are determined and work correctly solely for one region and one time of year.

The algorithm for generating a snow mask using MODIS/Terra, Aqua data is also based on constant thresholds [7, 10]. The algorithm employs NDSI, constant threshold values of reflectance in channel 2 (near IR) and reflectance in channel 4 (green). Pixels, in which NDSI values exceed 0.4, most likely, belong to the class of snow. If reflectance in channel 2 (near IR-range) exceeds 11% and reflectance in channel 4 (green) exceeds 10%, then the pixel is defined as snow with a probability of 100%.

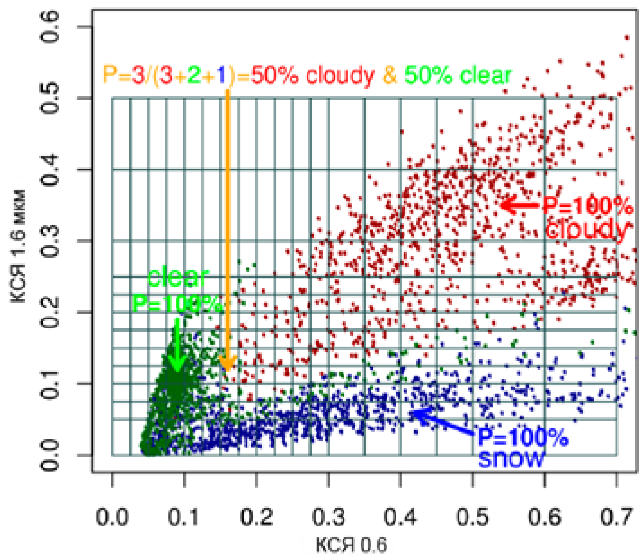


Fig. 2. Scattergram of the distribution of reflectance values in channels 1 and 3a of AVHRR/NOAA17 for various types of underlying surfaces according to a survey of the Alps [5].
* (clear) P – probability of the correct recognition of surface type.

Currently, practically all algorithms of snow mask generation use constant threshold values of reflectance, NDSI and radiation temperatures. The implementation of the abovementioned algorithms based on constant-threshold values on a global scale, Russia included, is complicated for the following reasons:

– thresholds are inconsistent and should change depending on the time of year;

– in most algorithms, thresholds are applicable only to a certain territory and their use for another territory leads to a noticeable decrease in reliability.

In addition, it should be noted that constant thresholds lead to the appearance of errors of the first and second kind, i.e. to gaps and false alarms. As a rule, errors occur at the boundaries of objects belonging to different types, which are characterized by a significant spread in R-values and radiation temperature values, for example, snow and ice inclusions in bare soil. The presence of errors is typical for any algorithm of snow and cloud mask formation but the use of constant thresholds exacerbates the situation. The algorithm set forth in the present paper for building masks of snow and cloud cover, not dependent on the time of year and territory, is primarily intended for providing exposure calculations performed by the satellite imagery planning complex. Maps of the snow and cloud covers are generated using MSU-MR/Meteor-M No. 2 data, which is received and processed on a daily basis by the Operator of Russian ERS space systems.

The main characteristics of the MSU-MR/Meteor-M No. 2 are given in Table 1. The same table also shows principal characteristics of the AVHRR, which is the foreign counterpart of the MSU-MR.

The initial data for the algorithm for detecting snow and cloud covers was the data of a daily survey of the MSU-MR/Meteor-M No. 2 of the territories of Canada, Russia and Europe obtained from 2015 to 2018. In total, 459 survey route fragments were used. The distribution of routes according to season is given in Table 2.

Table 2. Distribution of routes by season of the year and by region

	Number of routes			
Season of the year	Canada	Europe and the European part of Russia	Siberia and the Russian Far East	Total
Winter	27	28	9	64
Spring	43	32	29	104
Summer	39	24	143	206
Fall	30	46	9	85
Total	139	130	190	459

The coverage of the territory of Russia in the Far East and Eastern Siberia as well the survey of the territory of Canada are presented in Figure 3 in the form of overview images plotted on a map.

At the pre-processing stage, absolute calibration of MSU-MR data was performed, where the normalized digital readings were converted into physical quantities: into reflectance values in three short-wavelength channels and into values of radiation temperature in three long-wavelength channels.

Table 1. Main characteristics of MSU-MR / Meteor-M No. 2 and AVHRR

	MSU-MR		AVHRR	
Number of spectral channels	6		5 (working simultaneously)	
Spectral ranges	channel 1 channel 2 channel 3 channel 4 channel 5 channel 6	0.50 – 0.70 μm 0,70 – 1.10 μm 1.60 – 1.80 μm 3.50 – 4.10 μm 10.50 – 11.50 μm 11.50 – 12.50 μm	channel 1 channel 2 channel 3a channel 3b channel 4 channel 5	0.58 – 0.68 μm 0.725 – 1.10 μm 1.58 – 1.64 μm 3.55 – 3.93 μm 10.30 – 11.30 μm 11.50 – 12.50 μm
Swath, km	2800		2900	
Spatial resolution, km	1.0		1.1	

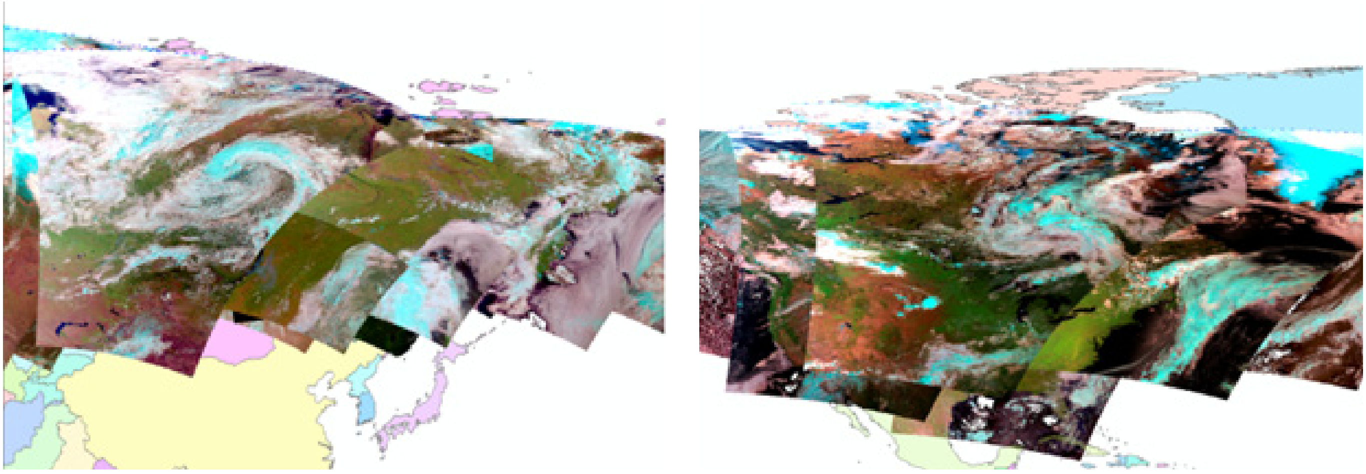


Fig. 3. Scheme of coverage by MSU-MR/Meteor-M No. 2 data of the territories of Eastern Siberia, Russian Far East (left) and Canada (right).

At the first stage, 17 survey routes were selected where all three types of surfaces are present: snow, clouds, "land" (in the given context, "land" means forest, sand and soil). The total area of test routes in the samples exceeds 130 million km^2. Areas known to belong to the each of the three surface types in every one of the seventeen routes are visually identified and outlined. The total area of the sampled territories attributable to clouds amounted to over 10 million km^2, the area of the sampled territories with snow cover was over 1 million km^2, the area of the sampled "land" – more than 500 000 km^2.

Figure 4 gives an example illustrating the result of this work: areas covered by clouds are marked by red contours; cloudless snow-covered areas are given with blue contours; open surface areas with "land" without snow and clouds – with green contours.

Data from MSU-MR measurements for the given areas were used to analyze the separability of the three surface types.

Considering that the most informative indicator for snow recognition is the NDSI, histograms of NDSI value distribution in the selected areas were built. In Figure 5, lines in shades of green represent NDSI histograms for "land"; red and lilac lines represent the NDSI values for clouds and blue, dark blue lines represent the NDSI histograms for snow. Histograms in Figure 5 are based on data from a survey conducted over a period of five days in May of 2016 over the territory of Canada.

According to the relative position of the histograms in Figure 5, the threshold value of NDSI = 0.4 (precisely this value of the NDSI index is used by the authors [7] when detecting the snow cover using MODIS data) allows

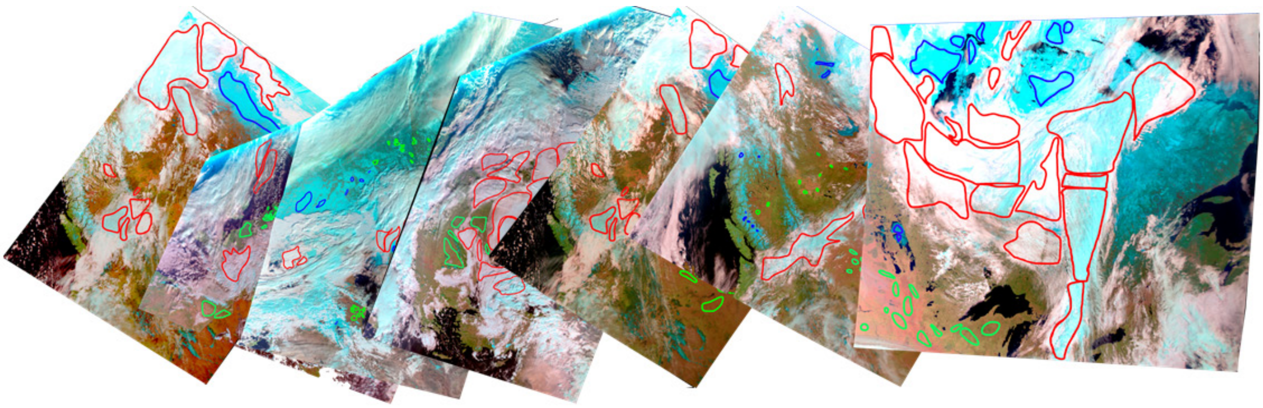


Fig. 4. Highlighting territories belonging to three types of surface types in images obtained in the period from autumn 2016 to spring 2017.

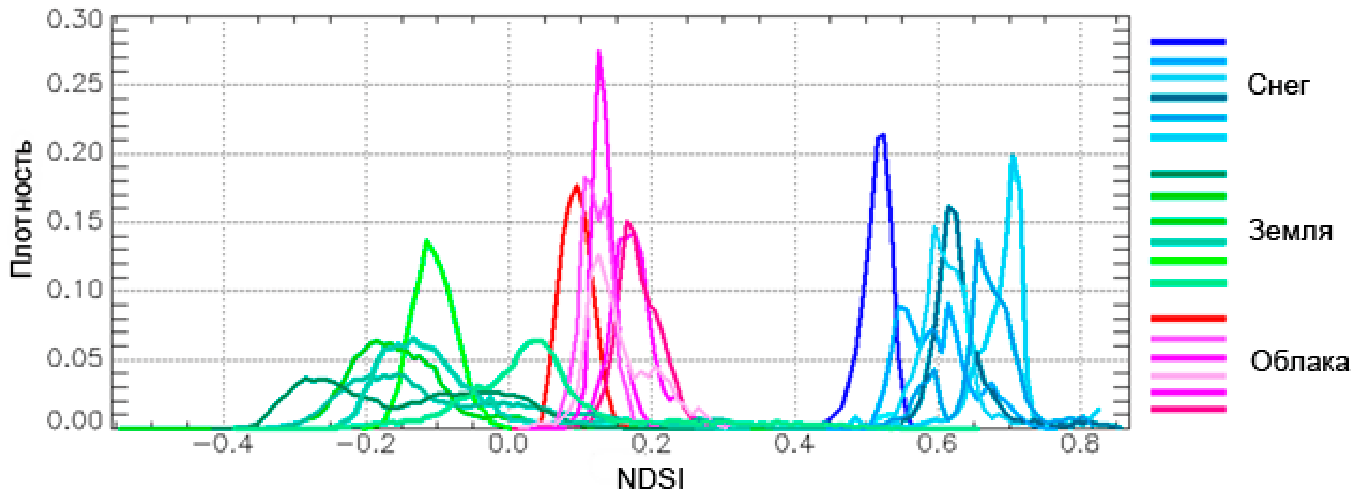


Fig. 5. Distribution histograms for routes over the territory of Canada.

to reliably discriminate areas containing snow from areas containing clouds and "land". Yet, no threshold value of NDSI provides a reliable differentiation between clouds and "land".

The next stage of the analysis consisted in building scattergrams for two indicators: NDSI and NDVI. Figure 6 demonstrates a scattergram of the distribution of NDSI-NDVI values for the same data sets as in Figure 5 obtained over the territory of Canada.

The distribution density of the pairs of NDSI-NDVI values is presented in the rainbow color palette: red corresponds to the highest frequency of occurrence of value pairs; yellow – to a lower frequency, blue and black – to the lowest occurrence frequency.

The scattergram in Figure 6 allows us to conclude that it is possible to reliably separate areas corresponding to snow and areas corresponding to "land" in the space of two indicators NDSI–NDVI. For this purpose, it is sufficient to determine a variable threshold, which can be represented as a function y=x. The graph of this function reliably cuts off snow from "land" but practically divides the cluster of values corresponding to clouds in half. Thus, in the space of two indicators – NDSI–NDVI even

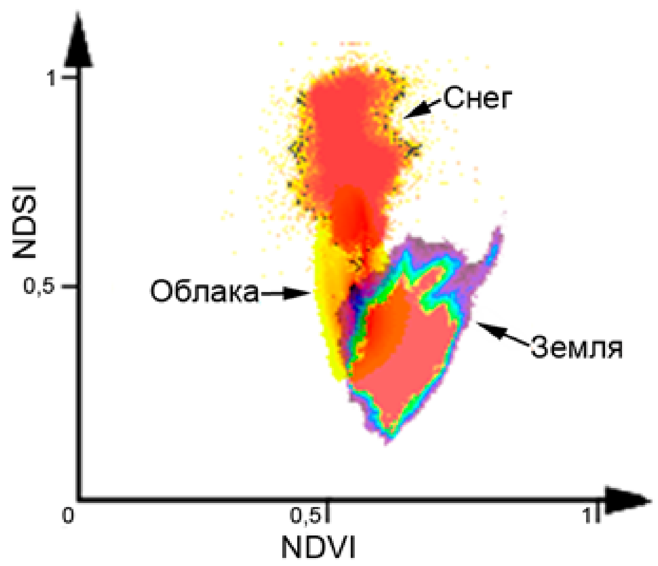


Fig. 6. Scattergram of the distribution of values in the NDSI-NDVI space for the selected areas corresponding to three surface types: cloudiness, snow, land – for five survey routes over the territory of Canada.

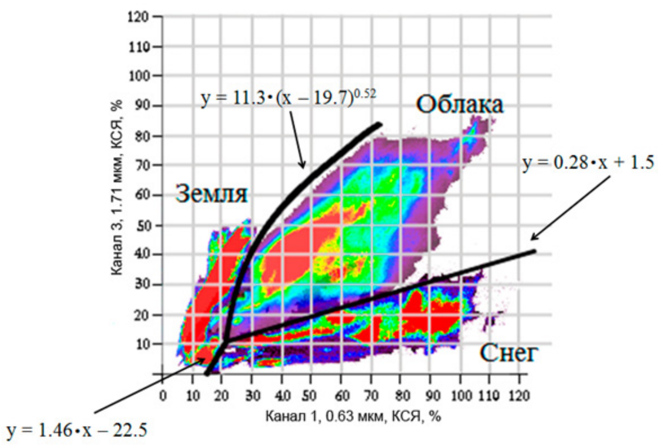


Fig. 7. Scattergram of the distribution of reflectance in channels 1 and 3 of MSU-MR/ Meteor-M No. 2.

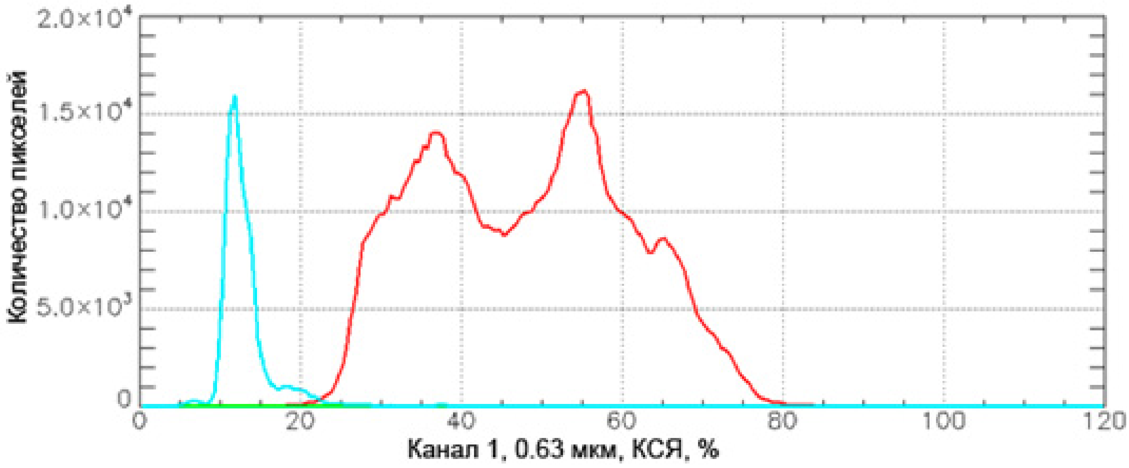


Fig. 8. Section of histogram in channel 1 (0.63 μm) within the limits of albedo values from 65% to 70%.

a variable threshold of the form y = x does not permit to solve the problem of reliably separating snow, cloudiness and an open surface completely.

Nevertheless, a more detailed study of the spectral properties of “land”, clouds, snow with the use of all of the channels of MSU-MR allowed us to identify the possibility of their separation in the space of two indicators: reflectance, % (albedo) in channel 3 (SWIR-1, 1.7 μm) and reflectance, % in channel 1 (green, 0.63 μm).

Figure 7 shows a scattergram of the distribution of the indicated R-values for the previously selected areas (plots of “land”) of land, cloudiness, and snow. Yet, the separation must be performed strictly with the use of variable thresholds; constant thresholds will not solve the problem. Note that the idea of the possibility of such separation of classes also arises during the analysis of the scattergram given in Figure 2 [5].

It is necessary to note that some R-values, % in channel 1 (0.63 μm), corresponding to bright objects exceed 100%. This is conditioned by the inaccurate calibration of the channels of the MSU-MR/Meteor –M No. 2 equipment noted in [12]. This also explains the presence of NDSI values exceeding 1 in Figure 6. Yet, it should be borne in mind that the inaccuracy of the absolute calibration does not affect either the algorithm or the result – the mask of cloud cover and mask of snow cover.

Here, as in Figure 6, the density of the distribution of R-values in 1.7 μm channel – R-values in channel 0.63 μm is given in the “rainbow” color palette (rainbow, ENVI environment)

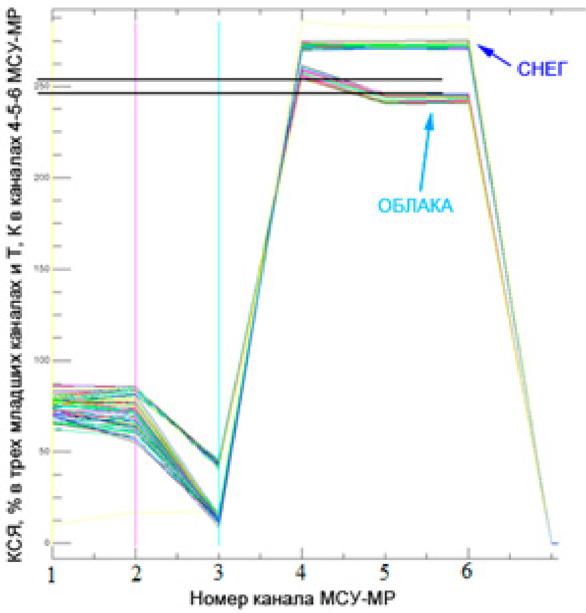


Fig. 9. Spectral images of objects pertaining to clouds or snow: difference in radiation temperatures in channels 4 and 5 of cloudy objects exceeds 3 K.

Variable threshold values are given in the form of functions y = f(x), where x is the value of the reflectance, %, in channel 1 of the MSU-MR (0.6 μm) and y is the value of reflectance, %, in channel 3 of the MSU-MR (1.7 μm). The graphs of the functions are presented in Figure 7 with thickened black lines. The expressions for function of the form y = f(x) are also given in Figure 7. Two functions out of three are represented as a linear function: $y = a \cdot x + b$; the third variable threshold is given in the form of a power function: $y = a \times (x+b)^c$.

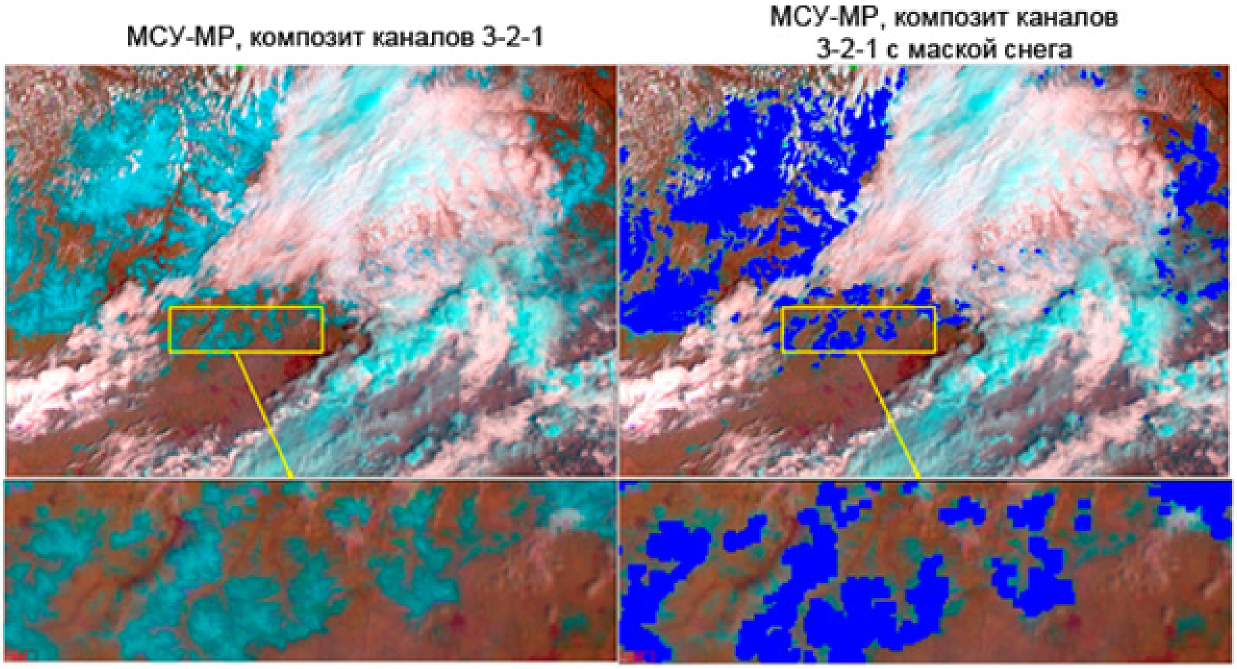


Fig. 10. Russian Far East, May 18, 2016, snow cover mask.

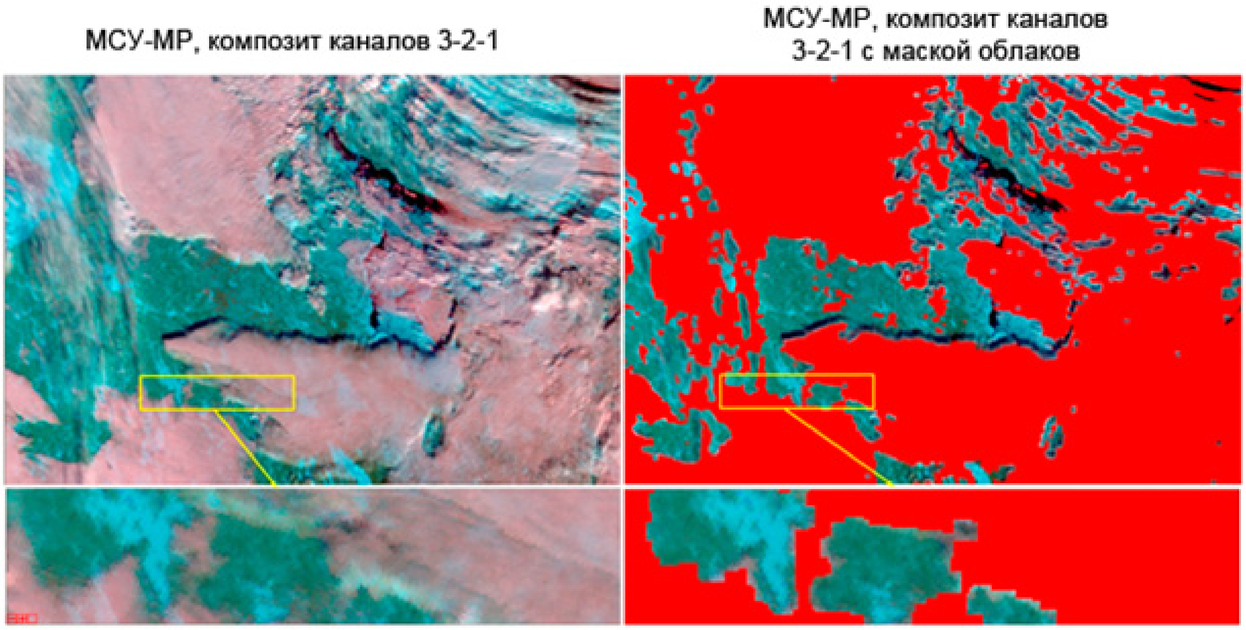


Fig. 11. European territory of Russia, February 27, 2017, cloud cover mask.

The evaluation of parameters a, b and c of variable threshold functions was performed with the help of the operation of equalizing the minimum standard error for a set of points, which is a set of particular threshold values that separate three types of surfaces. Particular threshold values were determined as follows. The entire space of the scattergram was divided along the x-axis into sections with a width of 5% in the range of low R-values and 10% – in the range of high R-values; for every segment with a width of 5% or 10% particular one-dimensional histograms were built. As an example, Figure 8 shows the particular histograms (land, snow, clouds) for the scattergram slice for R-values in channel 0.63 μm from 65% to 70%.

The middle parts of the sections, which the entire range of possible R-values was divided into in the 0.6 μm channel along the x-axis of the scattergram given in Figure 7, correspond to the x coordinates of points included in the set of particular threshold values. The y coordinates of the points in the set of threshold values were determined as R-values in the 1.7 μm channel, for which there is an intersection of particular histograms. For example, according to Figure 8, the particular threshold value for the range of R-values in the 0.63 μm from 65% to 70%, which discriminates snow from clouds, has the following coordinates: x = 67.5%, y = 22.8%.

In the process of algorithm validation, types of underlying surfaces were discovered, for which the application of additional conditions is required. Such objects include, for example, "land" with thin snow, which occurs, in particular, in late fall. Snow is classified erroneously as clouds on such a surface. The additional condition for correcting the error consists in the analysis of the values of the difference in radiation temperatures in the fourth and fifth channels of the MSU-MR: objects with a difference exceeding 3 K (Fig. 9) can be classified with confidence as clouds.

The variable threshold algorithm with an additional verification for early winter formed based on the data from 17 surveys of various territories was validated using survey data from Eastern Siberia, Russian Far East, the European territory of Russia as well as on Canada survey routes.

Testing confirmed the efficiency of the algorithm. The most important positive results are: 1) independence of variable thresholds from the time of year and 2) independence from the region where the survey is performed.

As an example, Figures 10 – 11 demonstrate the results of algorithm operation for two survey routes of the MSU-MR. Fig. 10 shows the Russian Far East in spring of 2016 and the mask of snow cover; Fig. 11 gives the European part of Russia in winter of 2017 and the cloud cover mask.

Note that the selected threshold values depend on the specific absolute calibration of the equipment. In order to maintain the correctness of the results over time or in case of a launch of a new SC, it is necessary to recalculate the coefficients of threshold functions or to take measures to ensure the constancy of the absolute radiometric calibration of MSU-MR. The latter, however, is true for any ERS data processing product.

It should also be noted that the developed algorithm of building snow and cloud masks does not provide pixel accuracy corresponding to the resolution of MSU-MR images of 1000 m. This is not conditioned by a peculiarity of the proposed algorithm but this is due to the general problem of surface classification: in transition zones – at the border, which divides objects of various types in the image, the probability of detecting objects of one type is reduced. This is caused by the simultaneous presence of signals incoming from objects of both types in the pixels of the given zone. Yet, it should be borne in mind that the snow and cloud mask generated with the help of the given algorithm is intended for use as a constituent part of the information base of the planning complex, in which data is given on a global scale (resolution – worse than 1 km). Consequently, the proposed algorithm ensures the generation of a snow mask and cloud mask, which are quite acceptable for planning (and processing) satellite surveys. The probability of detecting areal objects is close to 95 – 100%.

Conclusion

An algorithm is proposed for creating masks of snow and cloud cover using MSU-MR/Meteor-M data, which are a part of the planning complex for the target use of the Russian orbital constellation of ERS satellite and for processing the obtained data. The algorithm employs the values of the reflectance in channels 1 (green, 0.63 μm) and 3 (SWIR-1, 1.6 μm) of the MSU-MR. The significant difference of the present algorithm from analogues is the divergence from the practice of using constant thresholds. An analysis conducted on an extensive test set of source data demonstrated that the variable thresholds, which were determined by the present study, do not depend on the time of year or geographic location of the territory being surveyed.

Considering the continuous nature of MSU-MR surveys, the developed algorithm is capable of providing maps of snow cover and clouds in automatic mode with an update frequency of 2-3 days.

In the future, the introduction of the proposed algorithm in terms of snow cover mask formation directly into the technological complex of survey planning is anticipated. In addition, the cloud mask may be used to automate the process of searching for cloudless routes in archives of ERS data in accordance with consumer requests.

References

1. Gorbunov A.V., Il'ina I.Yu., Saul'skiy V.K. Sostoyaniye i perspektivy razvitiya kosmicheskikh kompleksov "Kanopus-V" i "Meteor-M" [State and perspectives of space complexes "Kanopus-V" and "Meteor-M" development]. *Raketno-kosmicheskoye priborostroyeniye i informatsionnyye sistemy* [Rocket-Space Device Engineering and Information Systems], Vol. 2, No. 4, 2015, pp. 14–19. (in Russian)
2. Stowe L.L., Jacobowitz H., Ohring G., R.K. Knapp, Nalli N.R. The Advanced Very High Resolution Radiometer (AVHRR) Pathfinder Atmosphere (PATMOS) Climate Dataset: Initial Analyses and Evaluations. *Journal of Climate*, Vol. 15, No. 11, June 2002, pp. 1243–1260.
3. *Portal distantsionnogo zondirovaniya* [Worldview Earth data]. Available at: https://worldview.earthdata.nasa.gov/ (accessed 23 September 2018). (in Russian)
4. Romanov P. Gutman G. Csisar I. Automated monitoring of snow cover over North America with multispectral satellite data. *Journal of Applied Meteorology*, Vol. 39, 2000, pp. 1866–1880.
5. Musial J.P., Hüsler F. Probabilistic approach to cloud and snow detection on Advanced Very High Resolution Radiometer (AVHRR) imagery. *Atmospheric Measurement Techniques*, 7, 2014, pp. 799–822.
6. Chokmani K., Bernier M., Royer A. A Merging Algorithm for Regional Snow Mapping over Eastern Canada from AVHRR and SSM/I Data. *Remote Sensing*, 5, 2013, pp. 5463–5487.
7. Hall D.K., Barton J.S. *Algorithm Theoretical Basis Document (ATBD) for the MODIS Snow and Sea Ice-Mapping Algorithms.* Available at: http://modis-snow-ice.gsfc.nasa.gov (accessed 04 April 2018).
8. Boschetti M., Nutini F., Manfron G., Brivio P.A., Nelson A.D. Comparative Analysis of Normalised Difference Spectral Indices Derived from MODIS for Detecting Surface Water in Flooded Rice Cropping Systems. *PLoS ONE*, Vol. 9, No. 2, 2014. Available at: https://www.researchgate.net/publication/260447511_Comparative_Analysis_of_Normalised_Difference_Spectral_Indices_Derived_from_MODIS_for_Detecting_Surface_Water_in_Flooded_Rice_Cropping_Systems (accessed 06 June 2019).
9. Gandhi M.G., Parthiban S., Thummalu N., Christy A. Vegetation Change Detection Using Remote Sensing and Gis – A Case Study of Vellore District. *Procedia Computer Science*, Vol. 57, 2015, pp. 1199–1210.
10. Xiong X., Barnes W., An overview of MODIS radiometric calibration and characterization. *Atmospheric Environment*, Vol. 37, No. 17, June 2003, pp. 2403–2412.
11. Ignatov A., Sapper J., Laszlo I., Nalli N., Kidwell K. Operational Aerosol Observations (AEROBS) from AVHRR/3 On Board NOAA-KLM Satellites. *Journal of atmosphere and Oceanic Technology*, Vol. 21, No. 1, January 2004, pp. 3–26.
12. Filei A.A., Rublev A.N., Kiseleva Yu.V. Otsenka stabil'nosti radiometricheskoy kalibrovki korotkovolnovykh kanalov mnogokanal'nogo sputnikovogo ustroystva KA "Meteor-M" No. 2 [Estimation of the stability of radiometric calibration of Multi-Channel Scanning Unit shortwave channels on board Meteor-M No. 2]. *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa* [Current problems in remote sensing of the Earth from space], Vol. 15, No. 42018, pp. 71–77. (in Russian)

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 71–80

Ways to Improve the Efficiency of the Spacecraft Flight Control System

I.N. Panteleymonov, *panteleymonov_in@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation

Abstract. The article is devoted to the problem of increasing the efficiency of the spacecraft flight control system. Various options for the organization of the spacecraft flight control systems are con-sidered. Network solutions for the transmission of control information in a single stream with target information using the TCP/IP Protocol stack will automate the process of receiving and transmitting information as well as lead to the reduction in the number of onboard radio equipment. The use of the SCADA-system will serve to improve ergonomics. Combining relay satellites into a single network of inter-satellite communication lines will enable receiving information from the spacecraft even when it is in the Western Hemisphere without placing command and measurement and gateway stations there. The decision to employ a low-orbit constellation as relay satellites would reduce delays in the trans-mission of information and reduce the energy budget of the radio link between the spacecraft and re-lay satellite.

Keywords: communication, spacecraft, orbital constellation, flight control system, radio link, com-mand and measurement station, onboard equipment, antenna system

Introduction

One of the main spacecraft flight control system requirements is efficiency, which is defined by the time required for receiving information on the state (telemetry) of the object being controlled τ_c (system response time), decision time τ_d and by the time required to communicate the controlling actions to the object being controlled τ_r:

$$\tau_c=\tau_s+\tau_d+\tau_r. \quad (1)$$

The following (and no less important) requirement for a flight control system is reliability, which is defined by the guarantee of a timely and reliable reception of telemetry information (TMI) q_t from the object being controlled, by the quality of decision-making q_d and the guarantee of the timely and reliable communication of control actions to the object being controlled q_r:

$$\mathbf{q_c=q_s+q_d+q_r}. \quad (2)$$

The reliability of the flight control system q_c is also dependent on the flexibility of the communication and information processing system.

Consequently, the main aim of building a control system is to reduce the time of response (τ_c) and increase the reliability q_c.

Note that for commercial systems an important factor in boosting the profitability is the reduction of the cost of the control system.

Implementation of geostationary satellites for increasing the efficiency of a SC flight control system

The traditional approach to increasing the efficiency of a flight control system of a spacecraft (SC) is to organize a two-tier control and communication system with the implementation of three-four geostationary relay satellites (GRS) [1-4].

The use of two active phased antenna arrays (APAA) functioning as antenna systems (AS) of GRSs is proposed with the aim to increase the number of simultaneously established communication links with a SC as well as to decrease the transition time from one SC to another and to boost the reliability of antenna systems, which is conditioned by the absence of mechanical parts. These antennas are a receiving and transmitting one; and each consists of three-four modules positioned at an angle generating a multitude of pointed (pencil) beams.

It is possible to ensure a round-the-clock communication link between a single command and measurement station (CAMS) or gateway station (GS) simultaneously with all the spacecraft located at all parts of the trajectory if there is an intersatellite communication link (ICL) between GRSs.

The implementation of interconnected GRSs by intersatellite radio links as repeaters will make it possible to:

1) simultaneously receive information from all the SC at any point on the flight path, which will lead to an increase in control efficiency;

2) reduce the number of CAMS and decrease the cost of the control system.

This direction of development of SC flight control systems is present-day and promising, yet, it has such a disadvantage as large information transmission delays conditioned by high altitudes of GRS orbits (circa 36 000 km) and, consequently, by large distances between orbital slots.

A generalized scheme of the organization of communication with satellites via GRS is shown in Figure 1. Note: GLONASS SC are used to transmit ballistic navigation information and synchronization signals to other spacecraft.

Implementation of a satellite communication system on low Earth orbit relay satellites for boosting the efficiency of SC flight control systems

An orbital constellation (OC) of a satellite communication system (SCS) based on low orbit relay satellites (LRS) is a certain amount of spacecraft located in several orbital planes (OP). Neighboring (adjacent) spacecraft belonging to one constellation should be connected using ICLs in such a manner so that every spacecraft is connected to four adjacent satellites in a single or in adjacent OPs [5]. Such architecture is employed for the Iridium mobile personal satellite communication network (PSCN) [6]. It has also been proposed in research papers [7] and in patents for the creation of a low-orbit telecommunication system based on small spacecraft [7, 8].

The implementation of the TCP/IP protocol stack will help transmit control information in a single stream with the target information to a unified GS concurrently performing the function of a CAMS.

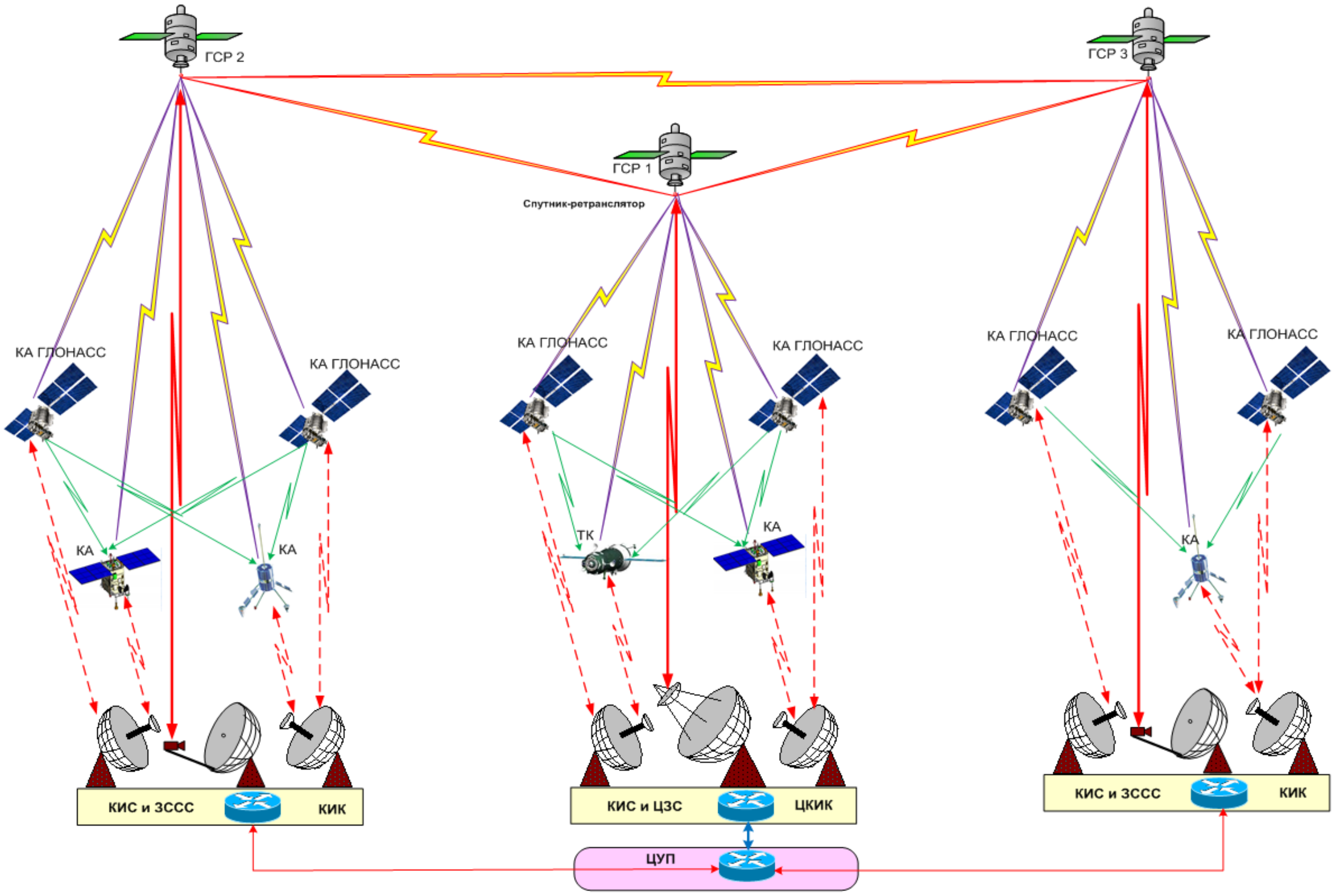


Fig. 1. Generalized scheme of link configuration for satellite communication via GRS.

Thus, an OC forms a fully connected satellite network for data transmission with the use of TCP/IP protocol stacks, where every spacecraft is a satellite router with its own information input/output ports:

– 1 global network port – connection to a CAMS – radio uplink (radio link, RL) Earth – SC in S-band;

– 1 global network port – connection to a GS, Earth – SC radio link in V-band or in the optical band;

– 4 global network ports – for retransmitting data to neighbor satellites using ICL;

– 1 local network port – for transmitting control data to spacecraft's own onboard equipment (OE).

A fully connected architecture of a satellite communication network, which is made up of SC connected via intersatellite communication links, has the following advantages [5, 8, 9]:

1) it supports the creation of a flexible network, where adaptive routing protocols can be implemented to build any data transfer route:

– with a minimum length path, which is especially critical for delays;

– with optimal throughput with account for onboard relay system (ORS) loading – for broadband traffic;

– with routes bypassing faulty SC or SC that are located in special areas (e.g., on the dark side of the orbit, in disaster areas and war zones);

2) demonstrates high survivability and adaptivity;

3) allows a single command and measurement station (CAMS) to have real-time access to any of the constellation SC.

The application technology of SCS based on LRS for SC flight control was implemented for the first time and perfected when controlling the flight of the TNS-0 No. 1 nanosatellite, developed and built at JSC "Russian Space Systems" [10]. This system used the SC of the Globalstar orbital constellation as relay satellites. Currently, the TNS-0 No. 2 flight program is being executed.

Unlike the existing SCS based on Globalstar LRS, the relay satellites of the proposed SCS are additionally equipped with four or six sets of receiving and transmitting wide-beam antenna (WBA) systems, located along two or three axes, for example: 1st variant: -Y and +Y, -Z and +Z;

2nd variant: -Y and +Y, -Z and +Z, -X and +X. For the needs of intersatellite communication, the Radio Regulations distributed a bandwidth in the S band the following way: 2025-2110 MHz – for the direct communication link and 2200–2290 MHz – for the feedback channel. In this frequency range, it seems most optimal to use a horn or spiral antenna system or a system of spiral antenna systems with a gain of 6-8 dB as wide-beam antennas.

Connected via intersatellite communication links, orbital constellation spacecraft form a global satellite network for data communication. A GS or CAMS, by establishing communication with a single SC from the OC within its radio coverage zone, gains access to any of the SC in the orbital constellation. By switching from one SC to another, 24-hour access to any spacecraft can be achieved with the use of one or several gateway stations.

In relation to another SC, a SC is a subscriber of the mobile satellite communications network (MSCN). Unlike the MSCN user terminal, each SC is assigned a basic and backup frequency on a regular basis as well as primary and backup code sequences, which are required for code-controlled access to the ICL between the LRS and SC. The spacecraft constantly receives pilot signals from various LRS via the service channel using wide-beam antennas. The decision-making device of the SC chooses the antenna that receives signals with the maximum amplitude. The pilot signals with a signal-to-noise ratio exceeding allowable threshold values are chosen. Then, using the Doppler frequency shift, the LRS, which is approaching the SC rather than moving away from it, is determined [10]. Having identified the optimal LRS for registration, the SC sends a registration request. After receiving the registration request, it transfers data via the LSR. If the value of the signal-to- noise ratio of the pilot signal worsens, the SC selects a different LRS using the abovementioned criteria and sends a registration request. After receiving a registration request, it transmits data via the newly chosen LRS, having completed the communication session with the previous LRS.

The repeaters of the SCS on the LRS with spacecraft, which are the objects being controlled, operate in multiple access mode in the LRS-SC radio link (forward channel) using MCPC technology (many stations on a single carrier) and in the SC-LRS link (reverse channel) – the SCSC is used (one station on one carrier)

Each SC is provided with two (2) fixed frequencies in the Ka-band and two spread-spectrum code combinations for reception and transmission. When the SC passes relative the LRS, a token-passing procedure (handover) between beams of one LRS and between LRSs is performed. During connection to LRS: authentication, registration and establishment of VPN-tunnels are carried out for information protection.

Every SC monitors the LRS-SC radio link, reads the headers of IP packets. If they identify their IP address, then they send the IP packets addressed to them into processing. The LRS receives and monitors every reverse frequency communication channel and then relays the IP packets via IRL and FRL (in the presence of a GS) to the MCC.

The architecture of SC flight control with the implementation of a SCS based on LRS is shown in Figure 2.

A functional diagram of a LRS satellite communication system is given in Figure 3 [2].

Link and control organization using one LRS is demonstrated in Figure 4.

An organizational chart of communication and control with the implementation of two LRSs is shown in Figure 5.

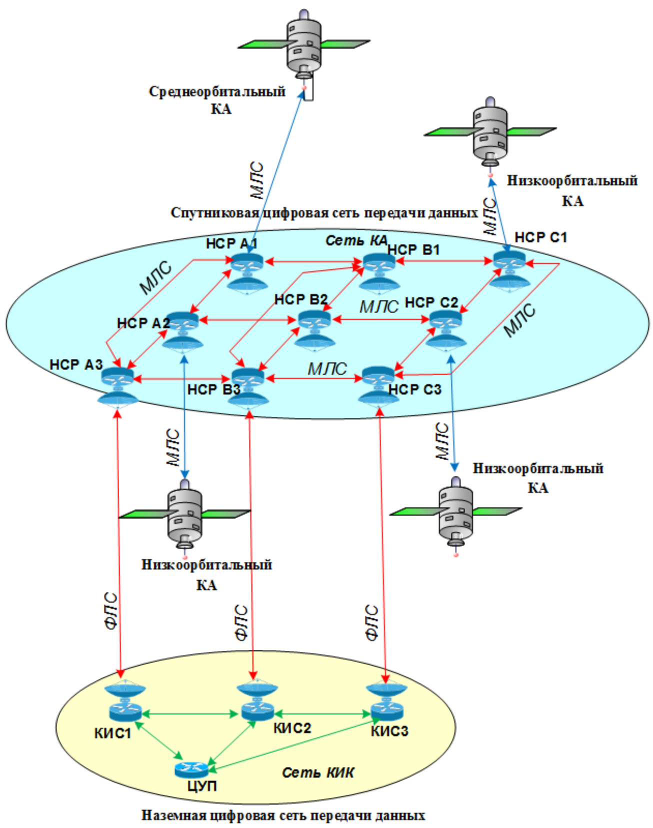


Fig. 2. Architecture of SC flight control with the implementation of a SCS based on LRS.

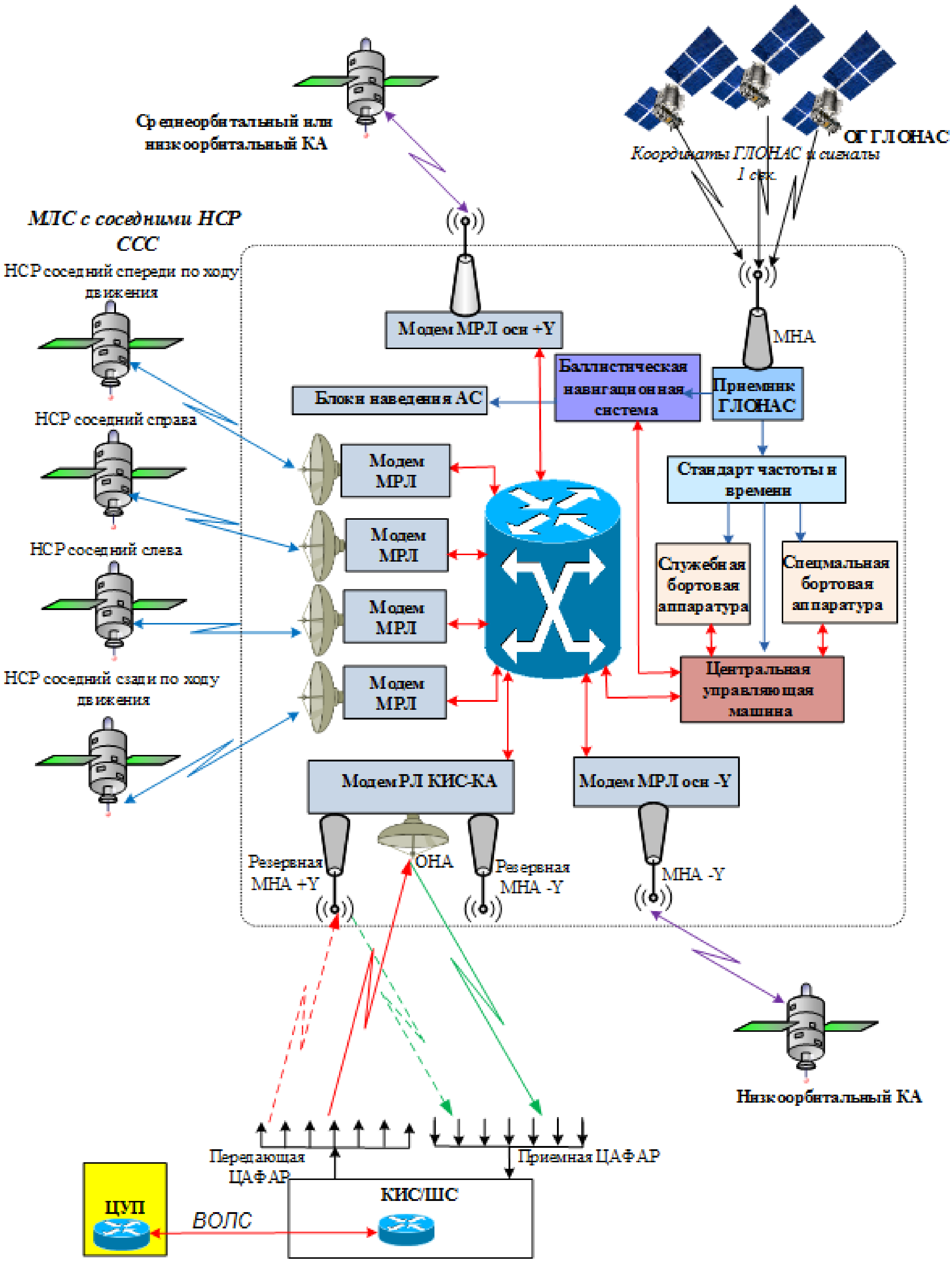


Fig. 3. Functional diagram of a SCS using LRS.

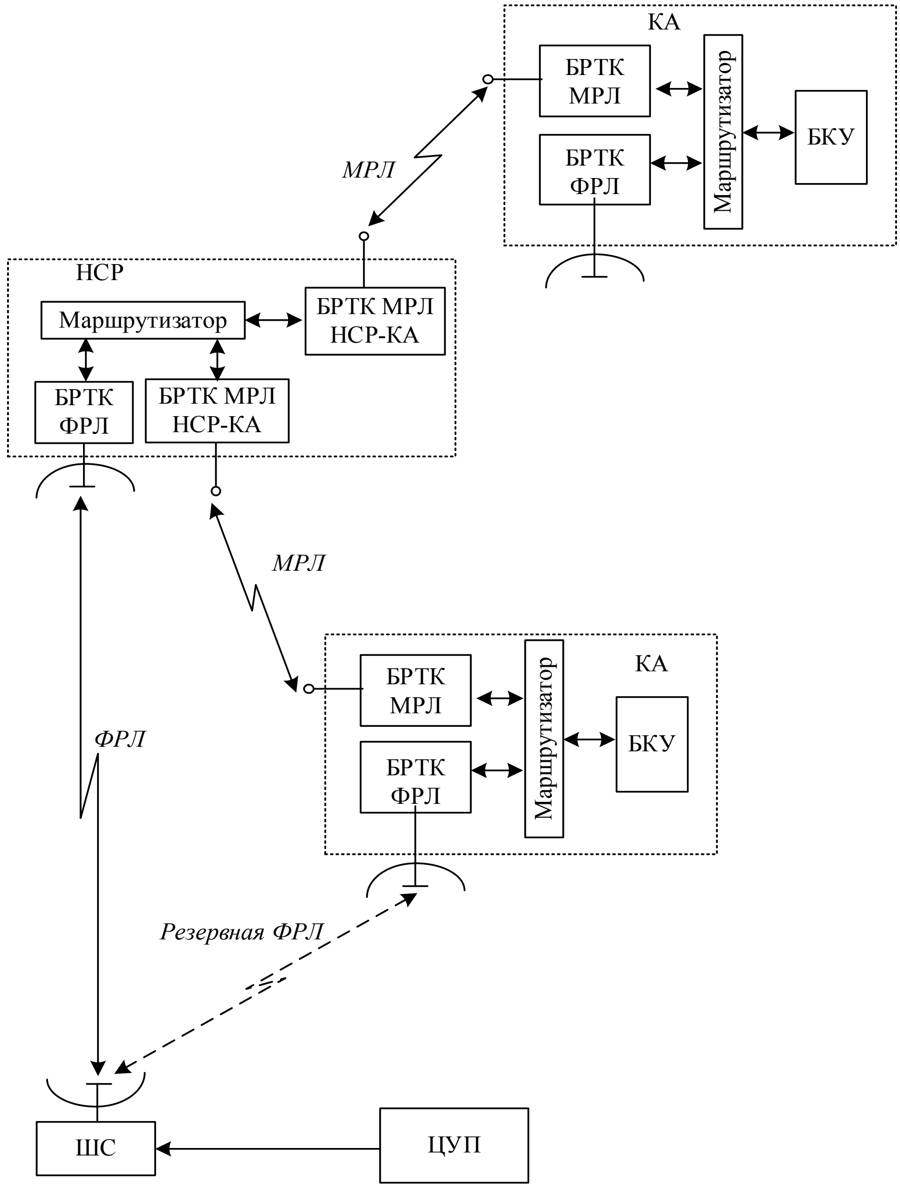


Fig. 4. Organizational chart of linkage and control using one LRS. Notation key: FRL – feeder radio link; IRL – intersatellite radio link; CRL – command radio link; GS – gateway station; ORU – onboard radio unit; OCU – onboard control unit

The highlighted advantages boost the efficiency of spacecraft flight control τ_c and the reliability of the control system q_c.

Thus, an LRS-based SCS is a satellite-based digital transport network for data transfer intended for relaying information of the information channel from the Mission Control Center (MCC) via GS and LRS to the SC and vice versa. The use of an LRS-based satellite communication system with a single GS will allow having simultaneous access to every SC of different OC, which will ensure high reliability and efficiency of the control system.

For communication between LRSs of the SCS through the ICL, it is most sensible to use the V-radio band and, in the future, switch to transmitting data in the optical band of radio waves.

The application of the V-band [5, 9] will allow us to:

– reduce the dimensions of antenna-feeder devices (AFD) and of the waveguide path of the microwave equipment;

– reduce the energy expenditures required to power the guidance systems of the AFDs and microwave equipment;

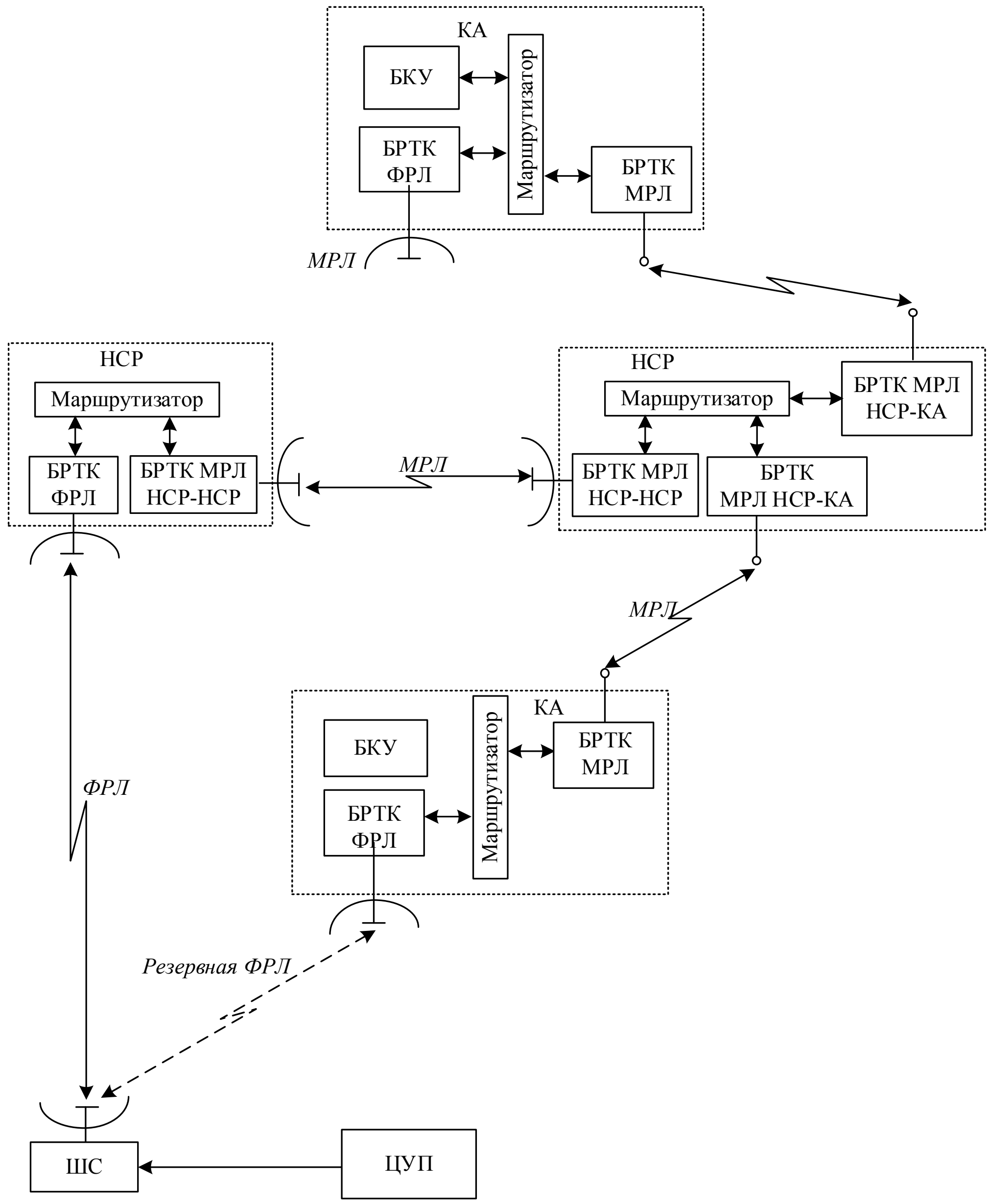


Fig. 5. An organizational chart of communication and control with the implementation of two LRSs.

– increase the bandwidth and the speed of information transmission (equal to and exceeding 10 GGbps).

Thus, for communication through an intersatellite radio link in V-band is most reasonable to use four pencil beam mirror antenna systems with a small diameter (up to 0.3 – 0.6 m) [7, 11], located along the X and Y axes of symmetry of the SC.

The implementation of the optical band in the ICL between LRS of the SCS [5, 6] will allow us to:

– reduce the dimensions by two to four times and the energy consumption of the onboard equipment of the intersatellite communication link in the future;

– significantly increase the bandwidth and information transmission rates (up to 10 GGbps).

Optical communication systems in the ICL, at the first stage of system operation, can be used alongside communication systems implementing the V-band radio ranges.

Optical receiving and transmitting antenna systems must be located along axes X and Z of the SC.

For communication via the Earth – SC RL (feeder communication link, FCL) it is most reasonable to resort to using the C, X, Ku or Ka-band, and it will be possible in the future to duplicate the radio channel by transmitting

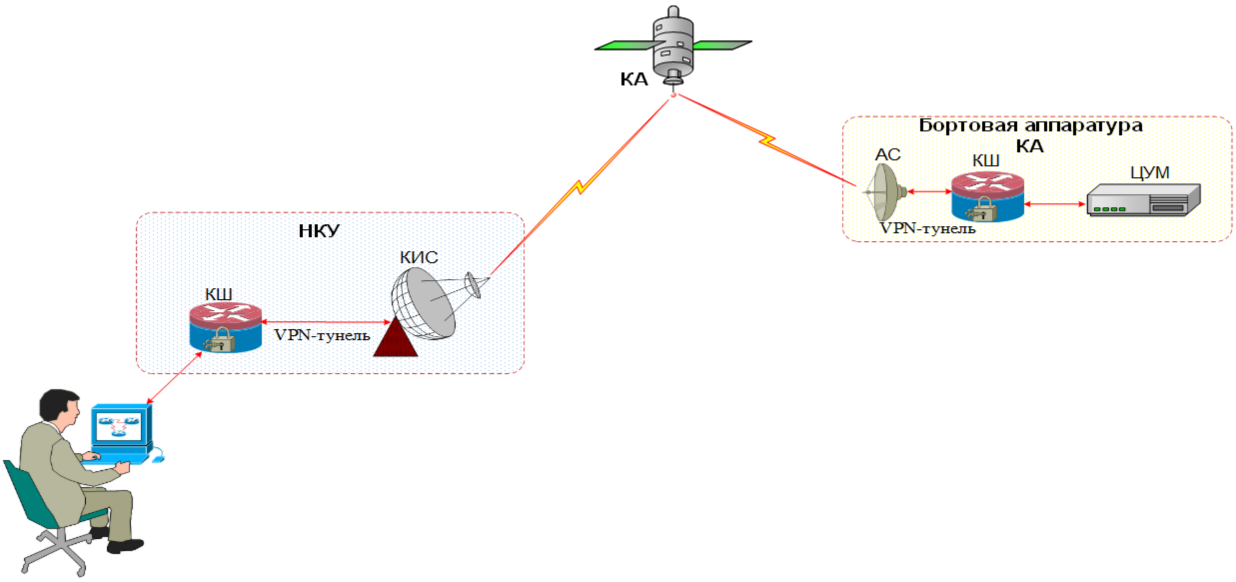


Fig. 6. Logical scheme of MCC operator access to the OE of the SC.

data in the optical range of radio waves [5, 6]. Due to the fact that every SC has a substantial radio coverage zone (RCZ), it is rare to have zone-wide poor weather conditions influencing atmospheric light transmission. Optical data traffic will be transmitted to a CAMS situated in favorable atmospheric light transmission conditions. Even if optical communication proves to be impossible in the given RCZ, it is possible to connect to a CAMS located in other RCZs via an ICL.

Therefore, for communication through the radio uplink (Earth-SC), it is most reasonable to install 1–2 pencil beam mirror antenna systems of a small diameter (up to 0.6 – 1 m) or a receiving and transmitting active phased array antenna (APAA) [5, 9]. Mirror antenna systems, APAA as well as optical receiving and transmitting antenna systems are to be located along the –Y axis of the SC and should be Earth-oriented.

The main flight control mode is remote access to the central controlling machine (CCM) of the SC onboard equipment (OE) achieved by setting up VPN-tunnels between the local area network (LAN) of the CCM and the LAN of the SC via wired and intersatellite communication links [1, 2, 5, 9]. Thus, MCC personnel are provided with remote access from their computers to the SC system control servers (controllers) and are able to promptly manage the OE of the SC by using specific software. A convenient windowed interface displaying images, graphs and tables will simplify the control system; improve its clearness, ergonomics and controllability along with reducing the decision-making time [5, 9].

Since the maximum information transmission rate in the flight control channel does not exceed 64 –128 kbit/s, it is possible to ensure the control of a large number of SC of a low-orbit OC via one or several CAMS at an objectively possible information transmission rate in the Earth-to-SC communication links and ICL.

The logical scheme of MCC operator access to the OE control system is given in Figure 6. [1, 2, 5, 9].

The main problem of flight control with the implementation of SCS on LEO RS, which are connected via ICL, is maintaining SC orientation accuracy during flight and pencil beam antenna array (AA) pointing accuracy. In the event that one or several SC in the OC fail, a way around them can be found with the use of dynamic routing protocols [5, 9], yet, in this case, arises the problem of gaining access to the SC, which had lost its orientation. For these purposes, at least two-four-six wide-beam AAs are available onboard, which ensure a low-speed emergency radio link with a CAMS of the ground-based control unit.

In the case of emergencies, for example loss of SC orientation, as well as during orbital insertion and disorbit, the CAMS of the ground-based control unit is capable of establishing a connection with the SC via wide-beam AAs on the SC through the ICL or a ground-satellite uplink.

The implementation of SCS based on low-orbit relay satellites, interconnected by ICLs, has all of the advantages as the methods described in part 1 but, unlike SCS based on geostationary relay satellites, it demonstrates additional advantages:

1) a smaller distance from the CAMS and GS to the SC and, as a result, data transmission delays are shortened and, thus, control efficiency is boosted;

2) apart from this, for this precise reason, the energy budget of the SC-SC radio link is reduced, which allows the use of wide-beam antenna systems based on LRS and SC. Thus, opens up the possibility of relaying control commands to a large quantity of SC without having to track them by pencil beam antenna systems.

3) has a high level of redundancy and, as a result, is highly reliable because the failure of several LRS, which can be easily be bypassed via ICL, cannot in any manner affect the timely delivery of the control channel information.

Conclusion

The comprehensive use of all of the abovementioned methods for solving the problem of boosting the efficiency of the SC flight control system provides the following advantages:

– control flexibility and agility;
– flight control system high reliability;
– provides a highly ergonomic and modern approach to solving the problem of control.

The implementation of data transmission protocols, which are standard for computer networks, (TCP/IP) will allow the implementation of standard network equipment (custom-made) and standard software for both building spacecraft onboard equipment and building onboard and ground-based control units. This will significantly simplify the flight control system and the architecture, and lower the production costs.

The flight control algorithms and architectural solutions for ground control and onboard control systems described in the paper make it possible to create a universal and reliable, dynamic and efficient communication and flight control system for any SC.

References

1. Panteleymonov I.N. Perspektivnyye algoritmy upravleniya poletom kosmicheskogo apparata [Perspective Algorithms for Spacecraft Missions Control]. *Raketno-kosmicheskoye priborostroyeniye i informatsionnyye sistemy* [Rocket-Space Device Engineering and Information Systems]. Moscow, FIZMATLIT, Vol. 1, No. 4, 2014, pp. 57–68. (in Russian)
2. Panteleymonov I.N. Korniyenko V.I. Arkhitekturnyye resheniya postroyeniya bortovoy apparatury kosmicheskogo apparata i perspektivnaya metodika upravleniya poletom kosmicheskogo apparata s primeneniyem setevykh tekhnologiy [Architectural concepts of creation of the onboard equipment of the spacecraft and perspective technique of spacecraft flight control using network technologies]. *Raketno-kosmicheskoye priborostroyeniye i informatsionnyye tekhnologii. 2015. Sbornik trudov VII Vserossiyskoy nauchno-tekhnicheskoy konferentsii "Aktual'nyye problemy raketno-kosmicheskogo priborostroyeniya i informatsionnykh tekhnologiy"* [Rocket-Space Device Engineering and Information Technologies. 2015. Proceedings of the VII All-Russian scientific and technical conference "Current problems of rocket-space device engineering and information technologies]. June 2–4, 2015, Ed. Romanov A.A. Moscow, JSC "RSS", 2015. (in Russian)
3. Bulgakov N.N., Alybin V. G., Krivoshein A.A. Osobennosti postroyeniya bortovoy apparatury komandno-izmeritel'noy sistemy kosmicheskogo apparata dlya upravleniya im kak v zone ego radiovidimosti s nazemnoy stantsii, tak i vne eye [Features of creation of the onboard equipment of a command and measurement system of the spacecraft to control it both in its radio visibility area from the ground station and out of it]. *24-ya Mezhdunarodnaya krymskaya konferentsiya "SVCH-tekhnika i telekommunikatsionnyye tekhnologii"* [The 24th International Crimean conference "Microwave Technique and Telecommunication Technologies"]. September 7–13, 2014, Sevastopol, Veber, 2014, Vol. 1, pp. 6–9. (in Russian)
4. Bulgakov N.N., Alybin V.G., Krivoshein A.A. Osobennosti postroyeniya dvukhkonturnoy bortovoy apparatury komandno-izmeritel'noy sistemy dlya upravleniya kosmicheskim apparatom na etape ego vyvoda na GSO [Building Peculiarities of Two-Channel On-Board Equipment of a Command-Measuring System for a Spacecraft Control During Its Placing into GEO]. *Raketno-kosmicheskoye priborostroyeniye i informatsionnyye sistemy* [Rocket-Space Device Engineering and Information Systems]. Moscow, FIZMATLIT, Vol. 1, No. 2, 2014, pp. 74–80. (in Russian)

5. Panteleymonov I.N. Perspektivnaya metodika upravleniya poletom kosmicheskikh apparatov odnoy orbital'noy gruppirovki s primeneniyem mezhsputnikovykh radioliniy [Advanced Technique of Spacecraft Flight Control of One Orbital Constellation Using Intersatellite Radio Links]. *Raketno-kosmicheskoye priborostroyeniye i informatsionnyye sistemy* [Rocket-Space Device Engineering and Information Systems]. Moscow, FIZMATLIT, Vol. 5, No. 2, 2018, pp. 73–83. (in Russian)
6. Kamnev V., Cherkasov V., Chechin G. *Sputnikovyye seti svyazi* [Satellite communication networks]. Moscow, Alpina Publisher, 2004, 536 pp. (in Russian)
7. Panteleymonov I.N. Sputnikovaya sistema, upravlyayemaya po mezhsputnikovoy radiolinii [The satellite system operated in the intersatellite radio link]. *Decision on issue of the patent for the invention dated April 25, 2019 for the application No. 2018125659/11(040586) dated July 12, 2018*. (in Russian)
8. Urlichich Yu.M., Grishmanovskiy V.A., Selivanov A.S., Stepanov A.A. Kosmicheskaya sistema global'noy sluzhebnoy sputnikovoy svyazi [Space system of global service satellite communication]. *Patent for the useful model No. 47600 dated March 24, 2005.* (in Russian)
9. Panteleymonov I.N. Kontseptsiya sozdaniya sistemy personal'noy sputnikovoy svyazi na nizkoorbital'nykh sputnikakh – retranslyatorakh dlya shirokopolosnogo dostupa k setyam peredachi dannykh [The concept of creation of a system of personal satellite communication on low earth orbit relay satellites for broadband access to data transmission networks]. *Sbornik trudov XXI Mezhvedomstvennoy nauchno-prakticheskoy konferentsii "Nauchno-prakticheskiye aspekty sovershenstvovaniya upravleniya kosmicheskimi apparatami i in-formatsionnogo obespecheniya zapuskov kosmicheskikh apparatov" (26–27 oktyabrya 2017 g.)* [Proceedings of the XXI of the Interdepartmental scientific and practical conference "Scientific and Practical Aspects of Improvement of Control of Spacecraft and Information Support of Spacecraft Launches" (October 26–27, 2017)]. Krasnoznamensk, Titov Main Test and Space Systems Control Centre, 2017, pp. 206–223. (in Russian)
10. Selivanov A.S. Razrabotka i letnyye ispytaniya pervogo rossiyskogo nanosputnika TNS-0 No. 1 [Development and Flight Testing of First Russian Technological Nanosatellite TNS-0 No. 1]. *Raketno-kosmicheskoye priborostroyeniye i informatsionnyye sistemy* [Rocket-Space Device Engineering and Information Systems]. Moscow, FIZMATLIT, Vol. 1, No. 2, 2015, pp. 74–90. (in Russian)
11. Panteleymonov I.N. Sputnikovaya sistema, upravlyayemaya po mezhsputnikovoy radiolinii [The satellite system operated in the intersatellite radio link]. *Patent for the invention No.* 2690966 dated June 06, 2019.

ROCKET-SPACE DEVICE ENGINEERING AND INFORMATION SYSTEMS
Vol. 6, Iss. 2, 2019, pp. 81–87

Calculated Experimental Evaluation of the Active Life of Microelectronic Devices for Space Purposes

V.P. Bezmen, *Cand. Sci. (Engineering), contact@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
A.A. Zhukov, *Dr. Sci. (Engineering), associate professor, contact@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
S.M. Ilyin, *contact@spacecorp.ru*
Joint Stock Company "Russian Space Systems", Moscow, Russian Federation
A.A. Stepanov, *Cand. Sci. (Chemistry), info@rjmpozit-mv.ru*
Joint Stock Company "Kompozit", Korolyov, Russian Federation
A.I. Vyalov, *info@rjmpozit-mv.ru*
Joint Stock Company "Kompozit", Korolyov, Russian Federation

Abstract. The article presents a calculation-experimental estimate of the active lifetime (ALT) of microelectronic devices (MED) for space purposes. One of the structural elements that determines the ALT of the MED is an adhesive, which is used to mount the MED on the base of the main body. In the calculations, it is assumed that the degradation of the adhesive at elevated temperature is described by the Arrhenius equation. Three batches of prototypes of the MED with test structures were made, in which the mounting of the crystal of the test structure was carried out using the OTPK–P and VK–26M adhesives as a reference sample. Accelerated testing of test structures showed that the use of OTPK brand adhesives for mounting MEA could provide ALT of 15 years.

Keywords: adhesive, degradation, microelectronic device, lifetime, temperature

Introduction

A spacecraft is a complex technical system, which, during space flight, must fulfill its functions under conditions with destabilizing factors of the outer space: deep vacuum, large temperature difference, radiation, charged particle fluxes, etc. [1, 2]. Therefore, the onboard equipment (OE) of the spacecraft (SC) should have a mean time between failures of at least 130 thousand hours (15 years) and a storability of at least 25 years [3]. One of the possible ways to assess the active life of the equipment is to conduct accelerated testing of the devices that make it up, and, in particular, microelectronic devices (MED). In this regard, the calculation and experimental assessment of the active life of microelectronic devices for space applications is an extremely urgent task.

The purpose of this work is to develop an estimate of the active life of microelectronic devices that make up the OE of SC.

To achieve this goal it is necessary to solve the following tasks:

- choose the structural element of the MED, which determines the period of its active existence. According to [4], in some cases, this element is a structural material — adhesive used in the MED for attaching the crystal to either the base of the ceramic-metal casing or substrate. During the operation of the MED, chemically active compounds (for example, water, halogens, etc.) can be released from the adhesive, leading to its failure due to corrosion of aluminum coating. In addition, like any polymer material, the adhesive is subject to uncontrolled natural aging;

- choose a test method by which it is possible to obtain the data on the time of active existence of MED.

As a rule, acceleration of lengthy tests is achieved by increasing the temperature compared to the normal operation. In this case, the assumption is made that both at increased temperatures and under normal conditions, the same degradation processes are accelerated, leading to a failure of the device. The rate of degradation processes and failures of semiconductor devices, including MED, at an elevated temperature obeys the Arrhenius equation [5, 6], which allows linear extrapolation of test results from high temperatures to normal MED operating conditions.

For the MED to function, the adhesive must have a specific volumetric electrical resistance of no less than 10^{14} Ohm · cm and adhesion strength at separation of at least 2.0 MPa in the range of operating temperatures [7]. Degradation processes in the MED adhesion joint occur both during its functioning and downtime, and are manifested either in the loss of adhesive or insulating properties, or both of them. As a rule, the manifestation of the loss of adhesive properties in polymer coatings is due to the breaking of chemical bonds in the organic material and adhesive bonds with the substrate and the crystal of the device [8]. The emergence of currents leaks through the adhesive joint is the result of breaking bonds in the polymer matrix of the material with the formation of uncompensated valencies.

For mounting the MED crystal on the base of the metal ceramic casing, a one-component phenol-rubber adhesive VK–26M is used. It completely wets the bonded surfaces of the MED crystal and the ceramic base of the body, ensuring the formation of joints with a minimum thickness of the adhesive joint. The small thickness of the adhesive joint allows avoiding the destruction of adhesive joints due to internal stresses in the adhesive joint caused by the difference in coefficients of linear thermal expansion (CTE) of the materials and the adhesive. According to the authors of [9], after it is cured at a temperature of 523 K for at least 3 hours, it practically does not emit products capable of condensation, ensuring a minimum content of water vapor (up to 0.5% vol.) in the casing of the device with the adhesive connection.

It is known that not all MED elements withstand the temperature of 503–523 K during the curing of the adhesive. We also consider a one-component, heat-conducting, low gas emission, epoxy adhesive of the OTPK-P brand developed at the Kompozit JSC, which has high electrical insulation characteristics and a curing mode of 1 hour at 393 K or 2 hours at 423 K [10].

To conduct accelerated tests, a test structure was developed, which is a KDB – 7.5 <100> silicon crystal measuring 5.1 × 5.1 mm in size, coated with thermal oxide 0.5–0.6 μm thick, on which aluminum strips of various widths: 8, 16 and 24 microns are formed. The metallization thickness is 1 μm. The size of the contact pads of the strips is 120 × 100 microns. Aluminum wire outlets 30 μm in diameter was welded to them by the method of ultrasonic welding. Figure 1 shows a silicon wafer with a diameter of 76 mm with test structures. The total number of test structures on the plate was 88.

For testing, twelve samples of crystals of the test structure were made, the installation of which on the ceramic base of the H14.42–1B body was carried out using the VK-26M adhesive and two modified adhesives

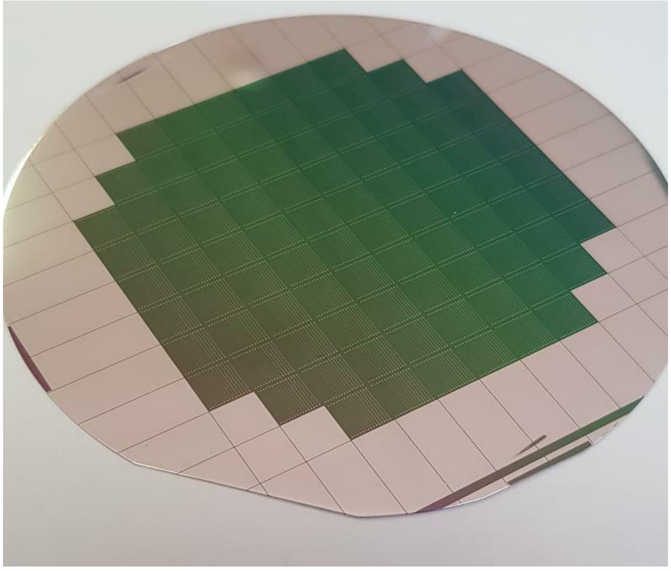


Fig. 1. Appearance of a silicon wafer with test structures

of the OTPK-P brand. Five samples used the OTPK– 5–C adhesive, and four used the OTPK– ED–3 adhesive. Three samples in which the VK–26M adhesive was used were comparison samples. After mounting the crystal, the case was sealed using the existing technology.

The calculation of the accelerated testing duration to confirm the active life of 15 years was carried out according to the technique presented in [11]. The mean activation energies of the failure process are taken from the Table 1 [11].

The activation energy of the failure process was selected based on the geometric dimensions of the gate of the CMOS IC transistor installed in the N14.42–1V casing. The gate size of the CMOS IC transistor is more than 1 μm. Therefore, the calculations used the data for the CMOS row ≥ 1.

According to [12], adhesive joints made with the VK-26M adhesive can withstand temperatures of 523 K for 500 hours, and temperatures of 573 K for 100 hours. The OTPK–P adhesives should have temperature stability no worse than the VK–26M adhesive. Therefore, the temperature of accelerated tests was taken equal to 473 K.

The calculation of the accelerated short-term (ASTT) and long-term (ALTT) failure tests at a temperature of 473 K was carried out using the following data:

- the value of power dissipation P_{dis} = 1.5 W;
- thermal resistance of the crystal – case $R_{cr\text{-}cas}$ = 20°C/W;
- test temperature (ambient) T_{amb} = 100° C.

The value of the temperature of the crystal T_{cr} when testing at a temperature T_{amb} is determined, according to [11], by the formula:

$$T_{crSTT} = T_{crLTT} = T_{amb} + R_{cr-cas} \times P_{dis} \qquad (1)$$

where: $R_{cr\text{-}cas}$ is the crystal–casing thermal resistance ºC/W;
P_{dis} *is* the dissipation power in the forced mode, W;
T_{amb} *is* the ambient temperature in the forced mode, K.

Table 1. The activation energy values of the failure process for groups of integrated circuits at various temperatures of the crystal[11].

Integrated circuit group	The value of the generalized activation energy at different temperatures of the crystal (transition), eV			
	E_{a1}	E_{a2}	E_{a3}	E_{a4}
	25 - 70 ºC	71 - 150 ºC	151 - 200 ºC	201 - 250 ºC
Bipolar digital TTL, ECL	0.3	0.4	0.5	0.6
Bipolar digital TTL–S on p–MOS structures	0.3	0.5	0.6	0.7
Bipolar digital n–MOS structures, CCD	0.35	0.55	0.65	0.75
Bipolar Digital I^2L	0.4	0.6	0.7	0.8
CMOS ≥ 1.0 μm	0.45	0.65	0.8	0.9
CMOS 1.0–0.5 μm	0.55	0.75	–	–
CMOS 0.5–0.09 μm	0.6	0.8	–	–
Analog	0.45	0.65	0.8	0.9

where: TTL is transistor–transistor logic; ECL is emitter-coupled logic; TTL–S is transistor–transistor logic with Schottky diodes; CCD is device with charge coupling; I^2L is integral injection logic; n– and p – MOS are metal–oxide semiconductor structure with n and p type channels; CMOS is complementary metal-oxide semiconductor.

Substituting the numerical values in the formula (1), we obtained the value of the crystal temperature in the normal mode:

$$T_{STT} = T_{LTT} = 100\ °\mathrm{C}\ + 1.5W \times 20°C/W = 130°\mathrm{C}$$

The temperature of the crystal T_{cr} during the accelerated (forced) tests at $T_{amb.f}$=473 K was determined by the formula:

$$\mathrm{T_{crASTT}} = \mathrm{T_{crALTT}} = \mathrm{T_{amb.f}} + \mathrm{R_{cr-cas}} \times \mathrm{P_{dis}} \quad (2)$$

where: $R_{cr\text{-}cas}$ is the crystal–casing thermal resistance ⁰C/W;

P_{dis} is the dissipation power in the forced mode, W;

$T_{amb.f}$ is the ambient temperature in the forced mode, K⁰C.

Substituting the numerical values in the formula (2), we obtained the value of the crystal temperature in the forced mode:

$$\mathrm{T}_{\mathrm{cr}ASTT} = \mathrm{T}_{\mathrm{cr}ALTT} = 200\ °\mathrm{C} + 1{,}5W \times 20\ °\mathrm{C}/W = 230\ °\mathrm{C}$$

The activation energy of the failure process was selected based on the Table 1 data for CMOS≥ 1.0 μm.

In accordance with [11], the value of the acceleration coefficients for the long-term and short-term testing is determined by the formula:

$$\mathrm{C_{ASTT}} = \mathrm{C_{ALTT}} = \exp\left[\frac{\mathrm{E_a}}{\mathrm{k}} \times \left(\frac{1}{\mathrm{T_{cr}}+273} - \frac{1}{\mathrm{T_{cr.f}}+273}\right)\right] \quad (3)$$

where: E_a is the activation energy of the failure mechanisms, eV;

k is the Boltzmann constant, $8.6 \cdot 10^{-5}$, eV/K;

T_{cr} is the temperature of the crystal (transition) in the normal mode, °C;

$T_{cr.f}$ is the temperature of the crystal (transition) in the forced mode, °C;

Since T_{cr} and $T_{cr.f}$ lie in different temperature ranges (for which different values of activation energy are shown in Table 1), the total acceleration coefficient is equal to the product of acceleration coefficients calculated for each temperature range according to the formula (3).

Substituting the numerical values for the above temperature ranges, we obtained the numerical values of the acceleration coefficients K_1, K_2, and K_3, respectively, equal to 2.489, 10.227, and 3.525.

Then according to [11], the total acceleration coefficient is equal to:

$$\mathrm{K_a} = \mathrm{K_1} \times \mathrm{K_2} \times \mathrm{K_3}$$

$$\mathrm{K_a} = 2{,}489 \times 10{,}227 \times 3{,}525 = 89{,}729. \quad (4)$$

The duration of accelerated short-term reliability tests was determined by the formula:

$$\mathrm{t_{ASTT}} = \frac{\mathrm{t_T}}{\mathrm{K_A}}\ , \quad (5)$$

where: t_T = 1000 h is the duration of long-term reliability tests (LTT);

K_{ASTT} is the acceleration coefficient for the accelerated short-term reliability tests [11].

Substituting the values of the duration of the long-term reliability test and the acceleration coefficient for the accelerated short-term reliability test in the formula (5), we obtain:

$$t_{\mathrm{ASTT}} = \frac{t_{\mathrm{T}}}{\mathrm{K_A}} = \frac{1000}{89.729} = 11.15\ \mathrm{h}.$$

The duration of accelerated long-term reliability tests is determined by the formula:

$$\mathrm{t_{ALTT}} = \frac{\mathrm{t_{LTT}} - \mathrm{t_T}}{\mathrm{K_A}}\ , \quad (6)$$

where: t_{LTT} = 131400h is the duration of long-term reliability tests;

t_T is the LTT duration;

K_{ALTT} is the acceleration coefficient for the accelerated short-term reliability tests.

Substituting the numerical values in the formula (6), we obtain:

$$\mathrm{t_{ALTT}} = \frac{\mathrm{t_{LTT}} - \mathrm{t_T}}{\mathrm{K_A}} = \frac{131400 - 1000}{89.729} = \frac{130400}{89{,}729} = 1453{,}27\ \mathrm{h}.$$

Then the total time for accelerated testing, according to [11], is:

$$\mathrm{t_{AT}} = \mathrm{t_{ASTT}} + \mathrm{t_{ALTT}} = 11.15 + 1453.27 = 1464.42\ \approx 1465\ \mathrm{h}.$$

Before testing at elevated temperatures, all elements of the test structures were measured for their resistance R_{0i}. The measurements were carried out using the 4-probe method using a GOM–802 digital milliometer.

The test structures were exposed at a temperature of 473 K in a vertical heat chamber PV–212. Every 100 hours, the test structures were removed from the chamber, cooled to room temperature for 30 to 40 minutes. After that, the resistance of the elements of the test structures was measured using the 4-probe method. The first measurement was carried out after 48 hours in the chamber.

The results of changes in the resistances of the elements of the test structure are presented in the graphs plotted on a semi-logarithmic scale, where the X axis is the heat exposure time τ, in hours, on a logarithmic scale, the Y axis is $R_{rel.i}$, the relative change in the resistance value of the i element in%, calculated by the formula:

$$R_{\mathrm{rel}.i} = \frac{R_i - R_{0i}}{\mathrm{R}_{0\mathrm{i}}} \times 100\% \qquad (7)$$

where: i is the number of the element of the test structure, i varies from 1 to 9;

$R_{0\,i}$ is the resistance value of the i element before the start of testing.

Figures 2–4 show the results of testing test structures at a temperature of 473 K for 1788 hours. Figures 2 and 3 show the results of changes in the resistance of the elements of the test structures, the crystals of which were mounted with the OTPK–ED–3 (structure No. 1) and OTPK–5–C (structure No. 6) adhesives.

Similar dependences were obtained for the seven remaining samples with the OTPK–P adhesives. As can be seen from the graphs presented, at the initial stage of the test (up to 248 hours), a decrease in the resistance of the resistive elements of the test structures from 5 to 8% is observed, then with a further continuation of the test, a monotonic increase in the resistance of the elements from 10 to 40% is observed (up to a test duration of 748 hours), then follows the same decrease in resistance (up to a test duration of 1088 hours) from 15 to 45%. With continued testing up to 1788 hours, a gradual decrease in the resistance of the elements of the test structures was observed.

The same dependences were obtained on comparison samples in which the crystals of test structures were mounted using the VK–26M adhesive. Data on the accelerated testing of the test structure No. 2, in which the VK-26M adhesive was used, is presented in Figure 4.

As a result of the accelerated testing, all the elements of the test structures showed a decrease in the resistance of resistive elements by 10–15% compared to the initial resistance and none of the elements failed during the testing. Thus, the test structures with both OTPK–P adhesives and VK–26M adhesive passed the accelerated tests at 473 K for 1465 h, which should correspond, according to the above calculations, to 15 years.

The decrease in the resistance of the resistive elements of the test structures during the initial testing is due to annealing of defects in the aluminum film produced by ion-plasma spraying. In addition, it is known that the films obtained are polycrystalline in structure, while the crystal grains are coated with a

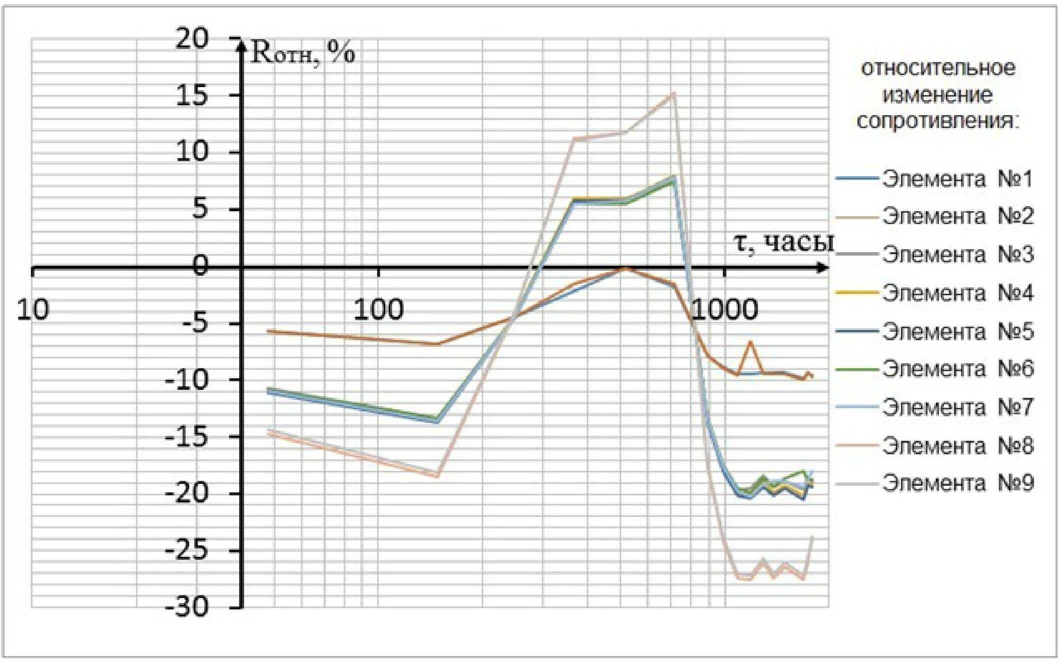


Fig. 2. Relative change in the resistance of elements of the test structure No. 1 with the OTPK–ED-3 adhesive during the accelerated tests at a temperature of 473 K: No. 1, No. 2 — changes in R_{rel} of resistive elements with a track width of 8 μm; No. 3 – No. 7 — changes in R_{rel} of resistive elements with a track width of 16 μm; No. 8, No. 9 — changes in R_{rel} of resistive elements with a track width of 24 μm.

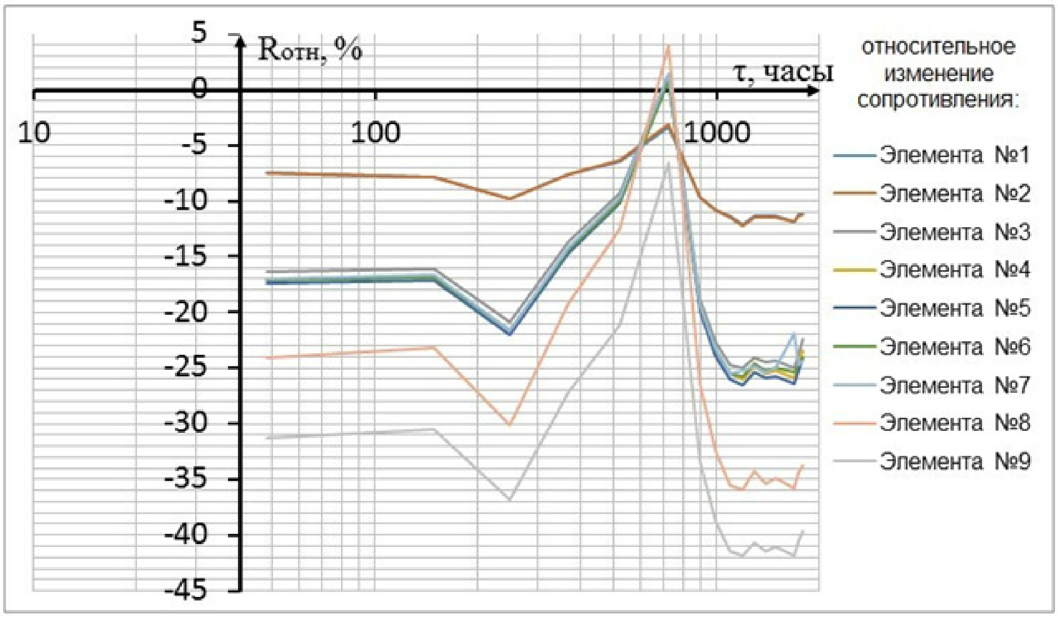


Fig. 3. Relative change in the resistance of elements of the test structure No. 6 with the OTPK-5-S adhesive during the accelerated tests at a temperature of 473 K: No. 1, No. 2 — changes in R_{rel} of resistive elements with a track width of 8 μm; No. 3 – No. 7 — changes in R_{rel} of resistive elements with a track width of 16 μm; No. 8, No. 9 — changes in R_{rel} of resistive elements with a track width of 24 μm.

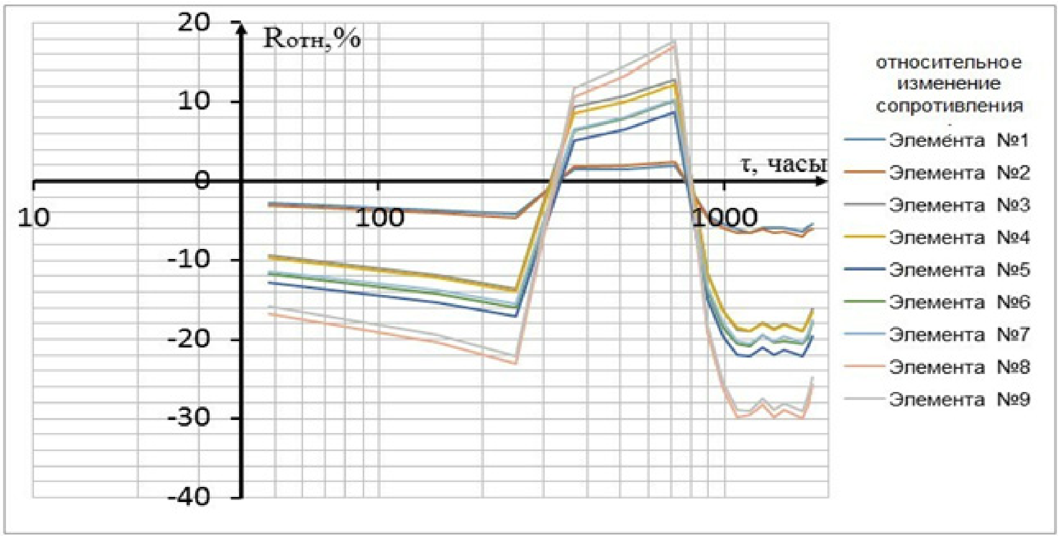


Fig. 4. Relative change in the resistance of elements of the test structure No. 2 with the VK-26M adhesive during the accelerated tests at a temperature of 473 K: No. 1, No. 2 — changes in R_{rel} of resistive elements with a track width of 8 μm; No. 3 – No. 7 — changes in R_{rel} of resistive elements with a track width of 16 μm; No. 8, No. 9 — changes in R_{rel} of resistive elements with a track width of 24 μm.

natural oxide with an amorphous structure [13]. The oxide thickness is about 6 nm and it is the $\gamma - Al_2O_3$ phase [13]. During the test (during heating), a phase transition occurs - the transition of the amorphous $\gamma - Al_2O_3$ phase to the crystalline phase in the form of $\alpha - Al_2O_3$. When an amorphous modification transforms into a crystalline one on the surface of alumina particles, chemically active centers disappear, which may be sites with structural defects or weak Al – O chemical bonds. We observed this phenomenon with a test duration of more than 248 hours — an increase in the resistance of the resistive elements of the test structures as the amorphous phase transitions to crystalline. As is known, the transition of the amorphous phase to the crystalline phase will be accompanied by a change in volume due to a denser packing of atoms, and an oxide film will crack [13]. In addition to the annealing of defects in the aluminum film during the test, its crystals will grow, moreover, the growth occurs due to a decrease in the fraction of the small ones. Changing the size of the crystals should lead to a decrease in the resistance of the aluminum film, which is observed with a test duration of more than 1000 hours Most likely, the transition of the amorphous phase to the crystalline phase ends at the test site from 700 to 1000 h, the growth process begins to play an increasingly important role, as a result of which a monotonic decrease in the resistance of the resistive elements of the test structures is observed.

As a result of accelerated testing of test structures with crystals, the installation of which was carried out using the OTPK– ED–3 and OTPK–5–C adhesives, it was found:

- the use of these adhesives for gluing microelectronic devices can provide an active life cycle of 15 years. However, the obtained calculations and experimental data require additional verification and confirmation regarding the mechanical strength of the adhesive joints;

– the variance in the tested parameter, the resistance of the resistive element made of aluminum film, during the test does not exceed 10 to 15%.

References

1. Model' kosmosa: nauchno–informatsionnoe izdanie [Model of the cosmos: scientific and informational publication]. In 2 volumes. Ed. by Panasyuk M.I., Novikov L.S. Vol. 1. Fizicheskie usloviya v kosmicheskom prostranstve [Physical conditions in outer space]. Moscow, 2007, KDU 872 p. (in Russian)
2. Model' kosmosa: nauchno–informatsionnoe izdanie [Model of the cosmos: scientific and informational publication]. In 2 volumes. Ed. by Panasyuk M.I., Novikov L.S. Vol. 1. Vozdeystvie kosmicheskoy sredy na materialy i oborudovanie kosmicheskikh apparatov [The impact of the space environment on the materials and equipment of spacecraft]. Moscow, 2007, KDU 1145 p. (in Russian)
3. Sevastyanov N.N., Andreev A.I., Osnovy upravleniya nadezhnost'yu kosmicheskikh apparatov s dlitel'nymi srokami ekspluatatsii [Fundamentals of the Reliability Control of Spacecraft with Long Service Life], Tomsk, TSU Publishing House, 2015, p. 266. (in Russian)
4. Kharlamov M.S., Zhukov A.A., Didyk P.I., Bezmen V.P. Degradatsionnye protsessy ustroystv mikrosistemnoy tekhniki [Degradation Processes in Microsystem Technology]. Raketno–kosmicheskoe priborostroenie i informatsionnye sistemy [Rocket-Space Device Engineering and Information Systems], 2018, Vol.5, No. 3, pp. 87–96. (in Russian)
5. Strogonov A.V. Otsenka dolgovechnosti BIS po rezul'tatam uskorennykh ispytaniy, [Assessment of the durability of LSI based on the results of accelerated tests]. Tekhnologii elektronnoy promyshlennosti [Electronic industry technology]. 2007, No. 3, pp. 10–16. (in Russian)
6. Kolpakov A. O termotsiklakh i termotsiklirovanii [On thermocycles and thermocycling]. Silovaya elektronika [Power electronics]. 2006, No. 2, pp. 6–11. (in Russian)
7. Gora Z.Yu. Tekhnologiya mikroelektronnykh ustroystv: Spravochnik [Technology of microelectronic devices: a Handbook]. Moscow, Radio i svyaz', 1991, 528 p. (in Russian)
8. Voronin I.V., Kondrashov E.K. Dolgovechnost' adgezionnykh svyazey polimernykh pokrytiy [Durability of adhesive bonds of polymeric coatings]. Available at: www.viam.com/public (in Russian)
9. Lukina N.F., Petrova A.P. Svoystva i primenenie kleev v pribornoy tekhnike [Properties and use of adhesives in the instrument technology]. Klei. Germetiki. Tekhnologii [Adhesives. Sealants. Technology]. 2005, No. 11, pp. 11–16. (in Russian)
10. Gladkikh S., Stepanov A., Antipin S. Teploprovodyashchie elektroizolyatsionnye klei dlya sborki izdeliy mikroelektroniki [Thermally conductive electrical adhesives for the assembly of microelectronic products]. Pechatnyy montazh [Printed circuitry], 2017, No. 3, pp. 183–187. (in Russian)
11. GOST R 57394-2017 Metody uskorennykh ispytaniy na bezotkaznost' [Methods of accelerated tests for reliability]. (in Russian)
12. Kalisevich V.P., Antipova T.I., Petrova A.P. Opyt primeneniya kleev pri izgotovlenii magnitoprovodov [Experience in the use of adhesives in the manufacture of magnetic circuits]. Available at: www.viam.com/public. (in Russian)
13. Zenin V.V., Spiridonov B.A., Zemlyansky A.I., Bokarev D.I. Korrozionnaya stoykost' alyuminievoy metallizatsii kristallov korpusov poluprovodnikovykh izdeliy [Corrosion resistance of aluminum metallization of semiconductor products case crystals]. Vestnik Voronezhskogo gosudarstvennogo tekhnicheskogo universiteta [Bulletin of the Voronezh State Technical University], 2015, vol.11, No.6, pp. 34–37. (in Russian)